

The formation of social conventions in real-time environments

Robert X.D. Hawkins¹², Robert L. Goldstone²,

1 Department of Psychology, Stanford University, Palo Alto, CA, USA

2 Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

* rxdh@stanford.edu

Abstract

Why are some behaviors governed by strong social conventions while others are not? We experimentally investigate two factors contributing to the formation of conventions in a game of impure coordination: the continuity of interaction within each round of play (simultaneous vs. dynamic) and the stakes of the interaction (high vs. low differences between payoffs). To maximize efficiency and fairness in this game, players must coordinate on one of two equally advantageous equilibria. In agreement with other studies manipulating continuity of interaction, we find that players who were allowed to interact continuously within rounds achieved outcomes with greater efficiency and fairness than players who were forced to make simultaneous decisions. However, the stability of equilibria in the real-time condition varied systematically and dramatically with stakes: players converged on more stable patterns of behavior when stakes are high. To account for this result, we present a novel analysis of the dynamics of continuous interaction and signaling within rounds. We discuss this previously unconsidered interaction between within-trial and across-trial dynamics as a form of social canalization. When stakes are low in a real-time environment, players can satisfactorily coordinate ‘on the fly,’ but when stakes are high there is increased pressure to establish and adhere to shared expectations that persist across rounds.

Introduction

Many everyday activities are governed by strong societal conventions: the side of the road we drive on, the meaning of “red” and “green” on traffic lights, the currency we use to pay for our meals, and the way we greet one another. These are all *self-sustaining*, in the sense that we will continue to conform to the convention as long as we expect others to, and also *arbitrary*, in the sense that at least one alternative regularity exists and would be equally acceptable as long as everyone coordinated on it [?].¹

Equally interesting, however, are the many activities not governed by conventions. Consider a pedestrian deciding what path to take through a busy marketplace. There is still a coordination problem to solve, because we do not want to keep running into one another, but we solve it ‘on the fly’ every time and there is no mutual expectation of

¹Some of these behaviors have also come to be regulated by the *law*, but this does not undermine their status as conventions. If we expected everyone to drive on the left, the cost of driving on the right (and crashing) would be much greater than the cost of following the convention (and breaking the law).

conformity. A broad literature on self-organization in crowd dynamics demonstrates that the spatiotemporal patterns formed by pedestrians are driven more by reactive or subconscious factors than strategic or conventional considerations [?]. Similarly, there are no societal conventions governing who gets the last bite of food at a meal or what kind of music to put on afterward.

First, note that these examples have a hierarchical structure: coordination must be achieved *within* a single interaction, but we repeat the interaction many times, so it may also be maintained or adjusted *across* interactions. Second, there is in some sense *more at stake* in the first set of examples than the second. In this paper we formalize this intuition in a game theoretic setting, present an experiment explicitly manipulating these qualities, and propose a mechanism by which new coordination problems may come to be governed by conventions. In order to obtain the resolution of data needed to answer this question, however, we first introduce a class of repeated games for which each interaction unfolds in real-time.

Across 70 years of game theory research, many properties of games have been fruitfully manipulated: the payoff structure, the number of players, the number of rounds played, the information players are given regarding the others' decisions, and the opportunity for players to communicate [?, ?, ?]. Until recently, however, most research has used games that unfold in discrete stages. These stages may be arranged simultaneously or sequentially, but players are typically given a fixed period of time to make their decision, which is recorded at the end of the period.

For many real-world decision-making tasks, there are no such timing restrictions. In the stock market, for example, brokers are free to buy or sell shares at any moment in time, not just on the hour, and their decisions immediately affect the payoffs received by the rest of the market. Thus, the trading industry, as well as a range of other economic, ecological, and social situations noted by Charness et al [?], feature *real-time* (or *continuous*) decision-making environments, as opposed to traditional *staged* environments.

The distinction between these different decision-making environments is not only useful for appropriately formulating the above scenarios, but has been shown to critically affect outcomes. Friedman and Oprea created a real-time version of the prisoner's dilemma game, in which players could toggle between 'cooperate' and 'defect' options at any point in time [?]. As soon as either player changed their action, the continuous flow of payoffs adjusted to a new level, which was maintained until the next change of action occurs. Total payoffs were computed by integrating these payoff curves over time. Median cooperation rates in the real-time environment increased to 90%, as opposed to 50% for the same game in an environment with 8 short discrete stages spanning the same total time as the real-time game, and close to 0% for the one-shot version. When the time period was split into finer grids of discrete stages (e.g. 60 discrete games, with each lasting 1 second), cooperation rates linearly approached those found in real-time games. Additionally, theoretical models in biology showed that when agents can respond to each other in real-time, a cooperative population evolves more easily and robustly than when interactions are restricted to discrete stages [?]. Again, this benefit increases on a continuum as the delay between stages decreases to zero.

We argue that the inclusion of real-time dynamics into a coordination game also introduces novel possibilities for exploring interactions across temporal scales. Instead of one continuous period of game play, as found in the above studies, we introduced a hierarchical design. If a game is repeatedly played across multiple rounds, and each round incorporates real-time play culminating in the awarding of a payoff, then within-round and across-round player interactions can potentially shape each other.

One hypothesis is that learning across rounds makes within-round coordination more efficiently achieved, as players adopt conventions learned not only from their general

experience with a community [?] but also from within the immediate, local interaction [?, ?, ?]. Another hypothesis, not mutually exclusive to the first, is that within-round conflict incentivizes the creation of conventions across rounds. Players experiencing conflict in the real-time interactions of a particular round of play may strategically create coordination patterns to avoid repeating these failures to efficiently coordinate in future rounds. Both of these patterns of influence across temporal scales were explored in a novel game that rewards dyads that can coordinate their game play across trials to avoid conflict within single trials.

In games like the prisoner's dilemma or the stag hunt, the most efficient (i.e. highest collective payoff) outcomes also happen to be fair. To the contrary, many interesting real-world situations create tension between efficiency and fairness: the most efficient outcome on a given round requires one party to come out of the interaction with a higher payoff than the other, inducing a power dynamic. We formulated a variation on the classic "Battle of the Sexes" game with this property, which we call "Battle of the Exes". Suppose there are two coffee shops in a town, one with better coffee than the other. Both you and your ex want to go out for coffee during your simultaneously occurring coffee breaks, but if you pick the same place and run into one another, it will make both of you unhappy. In game theoretic terms, if players choose different locations, the better coffee shop gives a payoff of a and the worse coffee shop gives a payoff of b , with $a > b > 0$. However, if the two players choose the same location, neither player is given a reward (i.e. given a payoff of 0). The matrices for the payoff settings (a, b) used in our experiment are formally specified in ??.

On a theoretical basis, we are interested in the "three levels of priority made necessary by social contract theory" proposed by game theorist Ken Binmore: efficiency, fairness, and stability [?]. Given the increased capacity for signaling and immediate reaction, it is perhaps unsurprising that real-time environments lead to more efficient solutions when everyone benefits equally. It is less obvious how a pair of players will implicitly negotiate the power dynamic induced by the payoff structure of impure coordination games. There are three Nash equilibria for the one-shot version of the game: two (unfair) pure strategies where one player always gets the high payoff and the other player always gets the low payoff, and one (inefficient) mixed strategy where players randomly pick between the two with an optimal probability determined by the payoffs (see ?? for a proof).

In the repeated version of the game, however, there exist many equilibria that are both fair and efficient. For example, players may take turns getting the high payoff target [?, ?, ?], or may implicitly agree to follow a reliable public signal, if one exists [?, ?]. These equilibria are efficient and fair, but require more complex patterns of self-organization on the part of the players. Our results indicate that when all other properties are held constant, allowing interactions to take place in a real-time environment solves many of the inefficiencies and inequalities that have frequently been observed, but it does not necessarily support stable equilibria. To explain the strong interaction between payoff discrepancy and real-time environments on stability, we use the rich millisecond-resolution data provided by our real-time paradigm to illuminate the relationship between micro-level social interactions within rounds and macro-level outcomes across rounds.

Materials and Methods

Experimental Framework

Experiments were performed using a suite of recent web technologies centered around Node.js. A full technical account of this experimental framework, including open-source

code to replicate the task, can be found in Ref. [?].

Ethics Statement

This manuscript reports experimental data from human subjects. Written informed consent was obtained after the nature and possible consequences of the studies were explained. The research contained in this submission was approved by the Indiana University Institutional Review Board

Participants

We recruited 568 participants on Amazon Mechanical Turk, yielding data for 284 dyads, which were spread across the four conditions of our factorial design. We posted 200 HITs for each condition, in separate batches. Because Mechanical Turk workers wish to participate in the experiment at unpredictable rates and times, and two participants had to be present to start a game at the same time, many players accepted \$0.10 for time spent waiting and left the virtual environment before being paired with another player. Hence, we collected uneven sample sizes for each condition.² As with all web experiments, there was nothing to prevent participants from dropping out halfway through the experiment, or simply ignoring the experiment throughout. ?? summarizes the number of participants who completed the full set of rounds in each condition, as well as information about how many dropped out or were excluded due to inattention.

Procedure

Participants were assigned to one of four conditions in a 2×2 experiment design. The real-time ('dynamic') decision-making environment and the simultaneous ('ballistic') environment formed one dimension. The \$0.01 vs. \$0.02 ('low') payoff discrepancy and \$0.01 vs. \$0.04 ('high') payoff discrepancy formed the other dimension. If no other players were waiting to begin a game when a participant entered the environment, they were placed in a waiting room where they could click to navigate around an empty world; if another player was already in the waiting room, a new game was immediately started. In this way, players were automatically paired into dyads. To avoid penalizing participants simply because others failed to join in time, players were eligible to submit their HIT for \$0.10 after completing a pre-test. Whatever amount of money players earned from payoffs in the game was rewarded in real money as a bonus. If the game terminated before completion due to one player dropping out, players were paid whatever they had earned up to that point as a bonus.

For the 'dynamic' condition, players were placed at opposite ends of the virtual world (approximately 120 pixels from both targets at the top and bottom), as shown in Fig. ?. The display gave the instruction "Click where you want to go," along with information on the amount of money earned so far, the number of games remaining, and the current speeds of both players, which were held constant in this experiment. One target was always assigned the value of \$0.01 and the other target always assigned \$0.02 ('low' condition) or \$0.04 ('high' condition), but which values appeared at the top and bottom were randomly selected at the beginning of each round. Players were given a 3 second countdown during which they had the option of secretly registering an initial destination. Once the countdown was completed, players started moving toward their destination, with full freedom to change that destination at any time. Changes in angle were immediately registered and updated in real-time within both players' displays, but movement took place in small increments of 10 pixels every 0.67 seconds. This interface

²Because we used non-parametric statistical methods to obtain our results, these differences did not pose a problem and we did not need to run more dyads.

Figure 1. Screenshot of the first round of the experiment. The colored triangles represent the players' avatars, the white circles represent targets (labeled by their corresponding payoffs), the outer shell gives the 'tie radius,' and the colored cross shows a player their current 'destination'

was chosen in order to minimize any effects due to lag or unequal proficiencies in using movement controls, as may be found when using arrow keys or an on-screen slider to move continuously.

For the 'ballistic' condition, all settings were the same, except for the initial stage. Instead of a countdown, players were first asked to click on one of the two targets, with the other player's choice hidden from them. Once both players made a valid selection of one target, no more input was registered and the avatars began moving toward the selected destination at the same rate as the dynamic condition, but without being able to change their courses. Clicks in the vicinity of a target were auto-corrected to the center of the target so that players did not accidentally miss the target. In order to bring total payoffs closer together across the 'low' and 'high' conditions, players in the 'low' condition played 60 rounds, while players in the 'high' condition only played 50 rounds. All other settings were held constant across this manipulation.

The rules governing assignment of payoffs were as follows: around each target (represented by a solid white circle), there was a visible, thin outer ring. If a player reached the inner target while any pixel of the other player's avatar was inside this outer ring, it was counted as a tie and the round ended without awarding either player a payoff. If a player reached the inner target while the other player's avatar was *not* within the ring, they were awarded the amount of money associated with the target, movement stopped, the round ended, and the other player was automatically awarded the amount associated with the other target, regardless of their location. This mechanism prevented a single player from trying to get both payoffs. The players were notified of the amount they earned via a message in the center of the screen saying "You earned 4¢[2¢]" At each step within a round, data on the player's locations, angles, current earnings, and payoff assignments was written to the file. Example play for the ballistic and dynamic conditions are shown in ?? and ??, respectively. The data and code used in the following analysis are available online³.

Results

Efficiency and Fairness

Our first major result concerns the efficiency and fairness of strategic interaction. Efficiency is defined as the sum $\rho_1 + \rho_2$, where ρ_i is player i 's total payoff. It quantifies the total amount of money the players were collectively able to earn. We divide by the total amount it was *possible* to earn in order to normalize all efficiency scores to the $[0, 1]$ interval and compare across conditions with different payoff structures. If a pair of players achieves the maximum efficiency of 1, they are optimally efficient. Note that there are many different sets of outcomes that achieve the same level of efficiency, some more fair than others.

To distinguish among these different outcomes, we introduce a measure of fairness, defined as the normalized payoff ratio

$$\text{Fairness} = \frac{\min(\rho'_1, \rho'_2)}{\max(\rho'_1, \rho'_2)}$$

³<https://github.com/hawkrobe/socialConventions>

Figure 2. Main results at Binmore’s three levels of analysis. Bars reflect standard errors. Note that the results for the two dynamic conditions were the same by efficiency measures (a) and fairness measures (b), yet differed markedly in stability (c). Higher payoff differences increased stability (i.e. decreased surprisal) within the dynamic condition, but had no effect on stability in the ballistic condition. ?? shows the same results using the less interpretable but more technically correct measure of “mean rank” on the y-axis.

where ρ'_i is the number of rounds that player i earned the higher payoff. This normalization maps the fairness of all conditions to the same $[0, 1]$ interval. If one player gets the higher payoff every round, this measure of fairness will be zero; if the players finish the experiment with an equal number of times of earning the high payoff, it will be one.

Given that the four distributions of fairness scores featured varying degrees of bimodality and the efficiency scores were similarly non-Gaussian (??), we used non-parametric techniques to test the null hypothesis of stochastic equality. For efficiency, a Kruskal-Wallis H -test showed a significant difference in the mean ranks of the four different distributions ($H(3) = 30.07, p < .0001$). Post-hoc Mann-Whitney U -tests at the Bonferroni corrected α value of $.05/4 = .0125$ showed that the dynamic conditions, ($M = 0.84, M = 0.84$), are significantly more efficient than ballistic conditions ($M = 0.69, M = 0.70$) both within the ‘high’ condition ($U = 1056, n_1 = 56, n_2 = 69, p < .0001$) and the ‘low’ condition ($U = 735, n_1 = 52, n_2 = 46, p < .001$). There was no main effect associated with the payoff manipulation, either at the ‘dynamic’ or ‘ballistic’ levels ($p = .32$ and $p = .47$, respectively; see Fig. ??a). Given that ties are the sole mechanism through which efficiency can be lowered in the game, a more specific restatement of this result is that ties are significantly more frequent in the ballistic conditions than the dynamic conditions.

Turning to fairness, a Kruskal-Wallis H -test showed a significant difference in the mean ranks of the four different distributions ($H(3) = 9.08, p = 0.030$). Post-hoc Mann-Whitney U -tests at the Bonferroni corrected level of $.05/4 = .0125$ showed that the ‘high’ ballistic condition ($M = .51$) was significantly less fair than both the dynamic ‘high’ ($U = 1378, n_1 = 56, n_2 = 69, p = 0.003$) and 1 v. 2 ($U = 953, n_1 = 56, n_2 = 46, p = 0.009$) conditions, both of which had a mean fairness of 0.69. The fairness scores observed in the ‘low’ ballistic condition ($M = .62$) were not significantly different from the ‘low’ dynamic ($p = 0.198$) or the ‘high’ ballistic ($p = 0.095$) conditions.

Stability

The third property in Binmore’s ‘social contract theory’ is *stability*, capturing the regularity of outcomes over multiple games. Unlike efficiency and fairness, which are summaries aggregated over all rounds, stability concerns the dynamic patterns of actions from one round to the next, thereby giving insight into how the overall efficiency and fairness was achieved. Intuitively, a strategy is stable if both players’ behavior is predictable and persistent. Knowing the outcome of a social interaction at one point in time should reduce one’s uncertainty about what will happen at other points in time. We capture this quantitatively using the information-theoretic measure of *surprisal*, which Shannon [?] defined as the negative logarithm of the probability of an event. In particular, we compute the average surprisal of a Markov Chain encoding the transition probabilities between events on successive rounds. An unlikely event will have low probability and an observer would therefore be highly surprised to see it happen, given knowledge of other events. This formulation is related to the model of Markov fictitious play introduced by Vanderschraaf and Skryms [?], though we only use it as a measure

for data analysis rather than a cognitive model of how agents are reasoning about one another within the game (see ?? for technical details of the measure).

Of course, the apparent stability of a pattern of events depends upon the encoding used to define events. An obvious choice in our experiment is the outcome – who, if anyone, received the high payoff at Round t . However, there turns out to be another encoding that better reflects the strategies used by some participants, especially in the ballistic condition. This alternative encoding uses the direction a participant chooses to go, regardless of the payoff associated with it. Dyads that settle on each player reliably moving in the same, unique direction (top or bottom) every round can be fair and efficient due to the randomization of payoffs to spatial locations on each round.

If players implicitly agree to trust that the high payoff has an equal probability of turning up in the top or bottom position, one player can choose the top, the other can choose the bottom, and the players can go to these pre-decided locations every round – they use an indifferent public signal to determine who gets the high payoff. This shows up as a highly random (though still efficient and fair) pattern in the outcome encoding, but a highly predictable pattern in the direction encoding. To be charitable, we computed stability under both encodings and recorded the more stable of the two (this measure was also robust across a). This biases all conditions equally toward higher stability.

First, it is interesting to note that dyads in the ballistic condition have a higher prevalence of adopting a public signal convention than those in the dynamic condition, $\chi^2(4) = 17.18, p = 0.002$ (see ?? for further details of this analysis). Second, plotting the surprisal distribution, we observe that the high-discrepancy, dynamic condition has strictly higher surprisal values than the ballistic conditions and the low-discrepancy, dynamic condition has strictly lower surprisal values (see ??). We formally checked this observation using a Kruskal-Wallis test, which showed that there are differences between the groups ($H(3) = 98.62, p < .0001$). Post-hoc Mann-Whitney tests at the Bonferroni corrected level of .0125 demonstrated that within the ‘low’ condition, the dynamic condition ($M = 0.71$) is significantly less stable than the ballistic condition ($M = 0.61$; $p < .0001$) and within the ‘high’ condition, the dynamic condition ($M = 0.56$) is significantly more stable than the ballistic condition ($M = 0.62$; $p < .0001$). The two dynamic conditions were significantly different from one another ($p < .0001$), but there was no significant difference between the two ballistic conditions ($p = .06$). This interaction, using surprisal as the y axis, is depicted in Fig. ??c. Note that lower surprisal implies higher stability.

Peel-off Times

Why is the high payoff difference version of the dynamic condition so much more stable than the low when they achieve the same efficiency and fairness? The answer to this question requires us to analyze the real-time dynamics within each round, which is the most unique feature of our data. When both players move toward the high payoff on a given trial, a dispute naturally arises over who should get the high payoff. Because getting the low payoff is better than a tie, one player will often ‘peel-off’ from their course toward the high payoff to concede that round’s dispute to the other player. We quantify the *degree of conflict* within a round i using the ‘peel-off time’ $t_C^{(i)}$, the percentage of round i ’s length before one player peels off. If neither player concedes, $t_C^{(i)} = 1$. If the players begin moving toward opposite targets at the outset of the round, $t_C^{(i)} = 0$.

The way peel-off time $t_C^{(i)}$ varies as a function of round number i conveys information about the formation of conventions. High peel-off times indicate a lack of consensus over which player should earn which payoff. Conventions by definition reduce

Figure 3. (a) Non-parametric local regression curve fitting peel-off times on each trial, broken down by ‘high’ and ‘low’ payoff conditions. Envelopes surrounding curves denote ± 1 (bootstrapped) standard deviation. Note that as conventions form, we predict a decrease in peel-off times, because there is a pre-determined action for both players to take in a given round. (b) Interaction plot comparing the mean peel-off time at the beginning of the experiment against mean peel-off times at the end (with intervals indicated by the arrows in (a)). We find a crossover effect, with the ‘high’ payoff condition beginning with higher peel-off times and ending with lower peel-off times. ?? demonstrates the robustness of this effect to different windows.

this uncertainty because players conform to their prescribed actions. Regardless of which convention is being chosen (e.g. turn-taking, public signal based on location, pure dominance by one player), peel-off times are predicted to decrease as the convention is adopted. Fig. ??(a) shows the average peel-off time for each round, smoothed using non-parametric local regression (lowess).

Note that the curve for the ‘high’ condition starts with higher peel-off times than the curve for the ‘low’ condition, and then by 1/3 of the way through the experiment has obtained lower conflict. In other words, higher initial conflict resulted in a quicker and longer-lasting adoption of conventions. Note also that the ‘low’ curve shows an uptick in conflict (i.e. a horizon effect) for the final several turns of the game, while the ‘high’ curve does not. We focus now on the first of these observations. To make the comparison rigorous, we conduct statistical tests directly on the data in the time intervals of interest, rather than smoothed curves.

We compared the peel-off times in the early interval (to the left of the first vertical line in Fig. ??(a)) against the peel-off times in the late interval (to the right of the second vertical line). A Kruskal-Wallis test showed a significant difference in mean rank across the four groups ($H(3) = 189.19, p < .0001$). Pairwise Mann-Whitney tests at the Bonferroni corrected α level of $.05/2 = .025$ show that the ‘high’ condition had significantly larger peel-off times than the ‘low’ condition at early times ($U = 71190, n_1 = 378, n_2 = 417, p = .0062$) and that the ‘high’ condition had significantly smaller peel-off times at late times ($n_1 = 1951, n_2 = 2433, p < .0001$).

We set our interval boundaries at $20\% \pm 6\%$ of the total number of rounds, where 20% is the crossover point of the curves. This is the smallest reasonable gap to place between early and late times, given that we do not expect values to differ across groups closer to the crossover point, but the result is robust across a range of larger spreads around this point (see ??).

Discussion

Social Canalization and Conventional Assimilation

Our results demonstrate that while players in a real-time environment achieve higher efficiency and fairness than players in a discrete-stages environment, the stakes determined by the payoffs of individual games are crucial in determining whether a stable convention will emerge across games. We conjecture a theory of *social canalization* to explain these results, by analogy to the mechanisms of genetic assimilation and canalization in biology.

In biological evolution, genetic assimilation occurs when a phenotypic characteristic that is elicited from an environmental condition becomes genetically encoded. A classic example involves exposing drosophila embryos to ether, producing a second thorax in some proportion. If the flies are selectively bred when they produce a second thorax, by

20 generations some flies develop a second thorax even without exposure to ether [?].

A fruitful analogy can be drawn between genetic change across generations and the formation of conventions across rounds of our task, and also between the development of a single organism and the dynamics within a single round of Battle of Exes. Similar to the genetic assimilation over generations of characteristics acquired within the lifetime of an organism, conventions arise from within-round dyadic interactions over rounds of game play. The end result in both cases is that the longer temporal scale of accommodation (e.g. genetic change and dyadic convention formation) depends for its existence on the shorter temporal scale of accommodation (e.g. individual development and within-trial dynamics), but once acquired, this longer-term accommodation eliminates some of the need for the shorter temporal scale accommodation.

In our task, after many rounds of interaction, dyads interacting in real-time are typically well coordinated without the any further need for inefficient within-trial dynamic interactions, particularly when the high discrepancy between payoffs places a premium on finding a coordinated solution. The greater efficiency, fairness, and stability of the (high discrepancy) dynamic condition compared to the ballistic condition indicate that the real-time dynamics within a round was instrumental for achieving these coordinating conventions across rounds.

A second interaction across temporal scales observed in biology is “canalization,” which refers to the robustness of a population’s phenotype with respect to changes in its environment and genotype [?]. The ability of organisms to learn or develop within a lifetime widens the basin of attraction surrounding an important trait, allowing a much broader range of genetic starting points to access it [?]. If there is an efficient, reliable mechanism for some trait to form during an agent’s development, there is diminished pressure to encode it in the genome.

Our game also demonstrates a social analog of biological canalization. When a dyad interacts in a dynamic, rather than ballistic, environment, then it can rely on real-time accommodation to achieve coordination and avoid ties. Experimental evidence for this comes from the greater stability of the 1 v. 2 ballistic, relative to dynamic, condition. When stakes are relatively low (i.e. there is not a large discrepancy between the payoffs), then dynamic dyads are not pressured to create an across-game social convention to assure their coordination. They can rely on within-round interaction to avoid ties, just as highly canalized organisms can rely on development within a lifetime to make up for genetic variation.

It is possible to unify these two apparently contradictory interactions across temporal scales. Sometimes, as in our high payoff discrepancy condition, real-time dynamics facilitate the creation of conventions to avoid inefficient and repetitious coordination negotiations on every round. Other times, as in our low discrepancy condition, real-time dynamics can be employed to prevent ties and promote coordination without requiring overarching conventions. Our experiment identifies *stakes* as one factor that governs whether real-time interactions will promote or obviate social convention formation. The same within-round interaction that establishes across-round conventions when the stakes are relatively high provides an acceptable low-effort alternative to establishing conventions when the stakes relatively low. Other factors that will likely affect whether the shorter temporal dynamic promotes or impedes the longer dynamical adjustment include the difficulty involved in establishing conventions, the number of rounds of iterated play, and the existing social structures for promoting coordination.

Conclusion

This theoretical framework may be fruitful in exploring other conventionalized behaviors in cognitive science. One area of particular interest is language, where some communication goals are handled ‘on the fly’ by pragmatic inference and others are

encoded in the lexicon [?, ?]. The factors that determined this division and continue to shape it remain an open question. Children face the same problem in language learning; they must infer which utterances they should encode and reproduce as linguistic conventions and which should be left to pragmatics. A substantial literature in social psychology documents the cultural norms that shape our values and attitudes toward others [?], and our results suggest a high-stakes, real-time pathway through which new norms may emerge. Our analysis also contributes to the broader groundswell of interest in cognitive science [?, ?, ?], economics [?, ?], and biology [?] in exploring the trajectory of decision processes as they unfold in real time.

Supporting Information

S1 Video

Demonstration of game play in ballistic condition

S2 Video

Demonstration of game play in dynamic condition

S1 Text

Proof of the three Nash equilibria for the one-shot “Battle of the Exes” game.

S2 Text

Technical details for computing stability.

S3 Text

Chi-Squared Test for whether different conventions tend to emerge in different conditions.

S1 Fig

Payoff Matrix for “Battle of the Exes” There are two coffeeshops in town, one with better coffee than the other. Both individuals would prefer to go to the coffeeshop with better coffee, but only if the other will not be there. If they run into either other, they are unhappy and get nothing.

S2 Fig

Empirical distribution of Efficiency, Fairness, and Stability measures for each condition. Note that our measure of fairness is not normal and does not keep the same shape across conditions, hence we must use non-parametric Kruskal-Wallis and Mann-Whitney tests to compare their stochastic ordering.

S3 Fig

Efficiency, Fairness, and Stability results using mean rank. Figure 2 in the main text uses interpretable means on the y axis, but because the Kruskal-Wallis and Mann-Whitney tests are based on mean rank rather than the mean of the sample

distribution, it is technically more correct to visualize the differences using mean rank. Note, however, that the qualitative patterns visible in Figure 2 are identical with the patterns here, so the visualization remains reliable.

S4 Fig

Surprisal CDFs for all conditions. Note that the ‘low’ dynamic condition lies above the other curves over the entire range of values and that the ‘high’ dynamic condition lies below the other curves.

S5 Fig

Surprisal analysis pipeline. Our pipeline of analysis from outcome time series (top row) to surprisal time series (middle row) to surprisal distributions (bottom row). The left column demonstrates what is intuitively a stable equilibrium, with some initial struggle converging into an alternation pattern. The right column demonstrates what is intuitively a less predictable or more unstable equilibrium, which has a much more erratic surprisal time series. The mean surprisal for the left column is consequently much smaller than the mean surprisal for the right column.

S6 Fig

Parameter robustness for estimating the Markov Chain and computing surprisal values. m determines how many steps back the Markov Chain looks when estimating the probability of transitions. The results shown in the main text are robust across many choices for this parameter.

S1 Table

Breakdown of dyads by condition. Dyads counted in “# included” were included in all analyses. Dyads in “# uncompleted” were excluded from the analyses because one player dropped out before completing the experiment. Note that all four conditions had roughly the same drop-out rate. Dyads in “# not paying attention” were excluded from the analyses because one or more players in the game allowed 5 or more rounds to pass without providing any input, indicating that they stopped paying attention.

S2 Table

Robustness of peel-off time analysis to different window sizes. All Kruskal-Wallis tests are significant at the $\alpha = .001$ level, and all post-hoc Mann-Whitney tests are significant at the Bonferroni corrected level of $\alpha = .05/2 = .025$. Note that the ‘high’ condition has greater peel-off times for early round, but lower peel-off times for later rounds.

Acknowledgments

This work was supported in part by a National Science Foundation Graduate Research Fellowship and Stanford Graduate Fellowship (to RXDH) and National Science Foundation REESE grant 0910218 (to RLG)

References

1. Lewis D. *Convention: A philosophical study*. John Wiley & Sons; 2008.
2. Helbing D, Buzna L, Johansson A, Werner T. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation science*. 2005;39(1):1–24.
3. Von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior*. Princeton University Press; 1944.
4. Kreps DM. *Game theory and economic modelling*. Clarendon lectures in economics. Clarendon Press; 1990.
5. Rasmusen E. *Games and Information: An Introduction to Game Theory*. 4th ed. Wiley-Blackwell; 2006.
6. Oprea R, Charness G, Friedman D. Continuous Time and Communication in a Public-goods Experiment. *Journal of Economic Behavior & Organization*. 2014;108:212–223.
7. Friedman D, Oprea R. A continuous dilemma. *The American Economic Review*. 2012;102(1):337–363.
8. van Doorn GS, Riebli T, Taborsky M. Coaction versus reciprocity in continuous-time models of cooperation. *Journal of theoretical biology*. 2014;356:1–10.
9. Jackson MO, Xing Y. Culture-dependent strategies in coordination games. *Proceedings of the National Academy of Sciences*. 2014;111(Supplement 3):10889–10896.
10. Frey S, Goldstone RL. Cyclic game dynamics driven by iterated reasoning. *PloS one*. 2013;8(2):e56416.
11. Garrod S, Anderson A. Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*. 1987;27(2):181–218.
12. Brennan SE, Clark HH. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 1996;22(6):1482.
13. Binmore K. *Natural justice*. Oxford University Press; 2005.
14. Vanderschraaf P, Skyrms B. Learning to take turns. *Erkenntnis*. 2003;59(3):311–347.
15. Helbing D, Schönhof M, Stark HU, Holyst JA. How individuals learn to take turns: Emergence of alternating cooperation in a congestion game and the prisoner's dilemma. *Advances in Complex Systems*. 2005;8(01):87–116.
16. Lau SHP, Mui VL. Using turn taking to mitigate coordination and conflict problems in the repeated battle of the sexes game. *Theory and Decision*. 2008;65(2):153–183.
17. Aumann RJ. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics*. 1974;1(1):67–96.

18. Gintis H, Helbing D. Homo socialis: An analytical core for sociological theory. Available at SSRN 2362262. 2013;.
19. Hawkins RXD. Conducting real-time multiplayer experiments on the web. Behavior research methods. 2014;p. DOI 10.3758/s13428-014-0515-6.
20. Shannon CE. A mathematical theory of communication. Bell system technical journal. 1948;27.
21. Waddington CH. Genetic assimilation of the bithorax phenotype. Evolution. 1956;p. 1–13.
22. Waddington CH. Canalization of Development and the Inheritance of Acquired Characters. Nature. 1942 Nov;150:563–565.
23. Hinton GE, Nowlan SJ. How learning can guide evolution. Complex systems. 1987;1(3):495–502.
24. Grice HP. Logic and conversation. In: Cole P, Morgan J, editors. Syntax and Semantics. New York: Academic Press; 1975. p. 43–58.
25. Garrod S, Doherty G. Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. Cognition. 1994;53(3):181–215.
26. Bicchieri C. The grammar of society: The nature and dynamics of social norms. Cambridge University Press; 2005.
27. Spivey MJ, Dale R. Continuous dynamics in real-time cognition. Current Directions in Psychological Science. 2006;15(5):207–211.
28. Koop GJ, Johnson JG. Response dynamics: A new window on the decision process. Judgment and Decision Making. 2011;6(8):750–758.
29. Beer RD. Dynamical approaches to cognitive science. Trends in cognitive sciences. 2000;4(3):91–99.