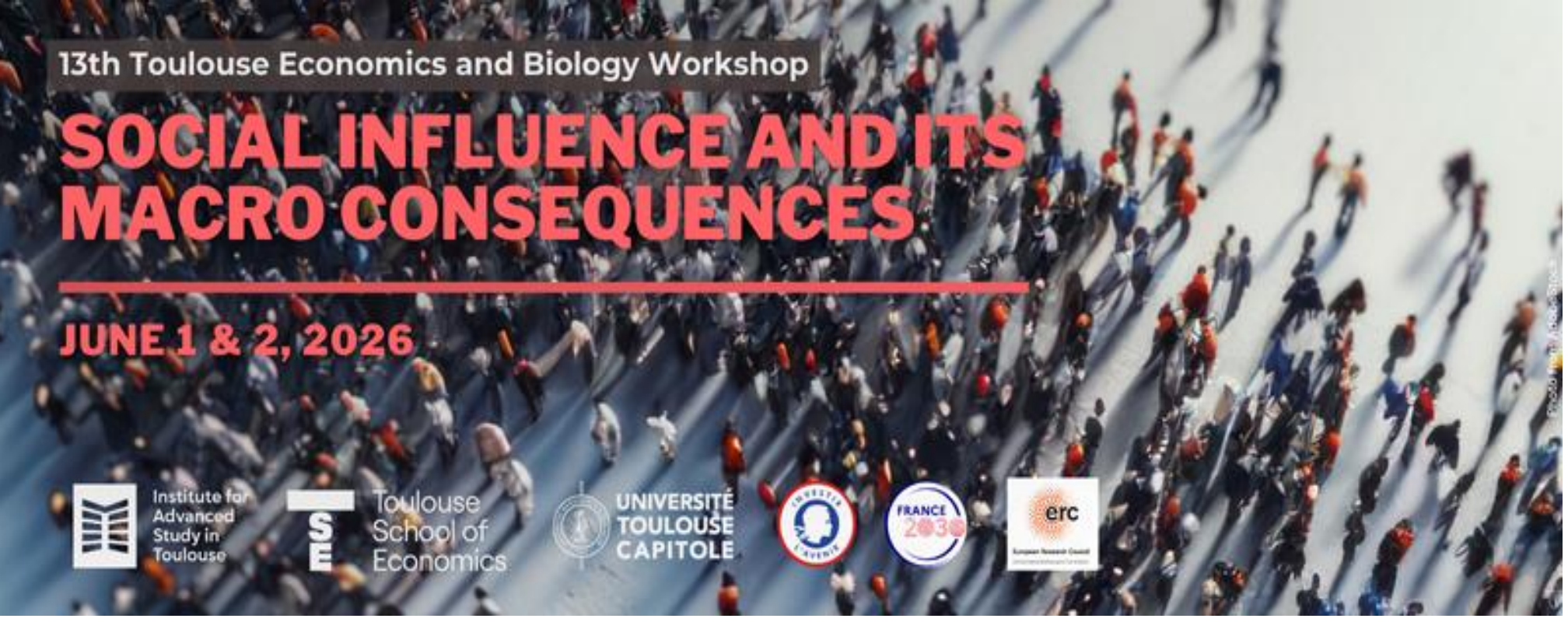


13th Toulouse Economics and Biology Workshop
Social Influence and its Macro-Consequences

June 1-2, 2026

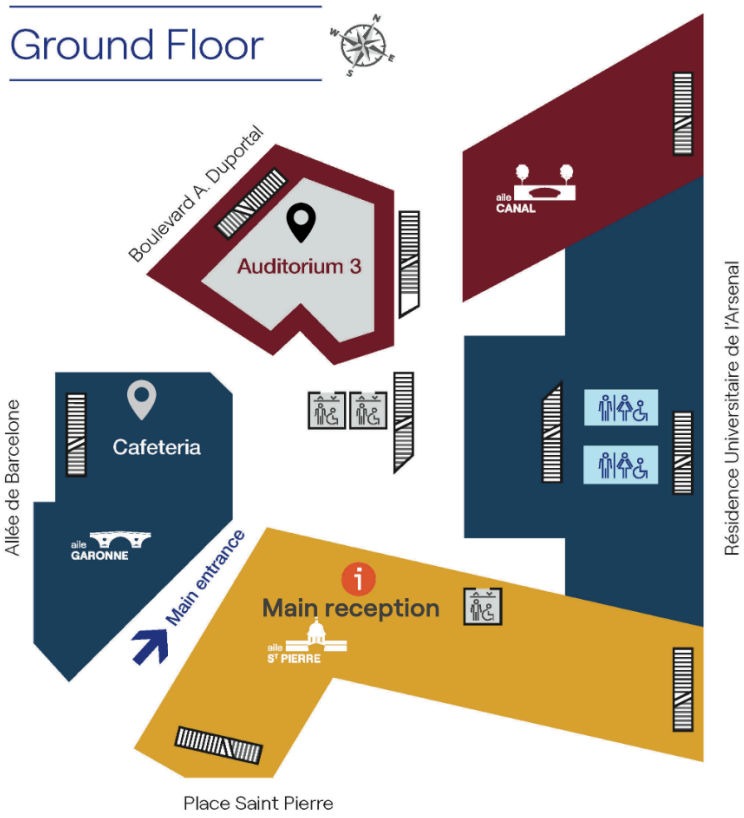
BOOKLET
Program & Abstracts

Conference venue
Institute for Advanced Study in Toulouse (IAST)
1, Esplanade de l'Université - 31 000 Toulouse, France
Auditorium 3 – Jean-Jacques Laffont (Ground Floor)

Organizers
Alexandros Gelastopoulos (IAST/TSE)
Marijn Keijzer(IAST/TSE)
Sergey Gavrilets (U Tennessee)

Conference secretariat
Aline Couratier & Valérie Nowaczyk
econbio@iast.fr

Conference Venue



Auditorium 3
All presentations

Cafeteria
Registration, breaks and lunches

Schedule

Monday June 1 - Morning

09:00–09:30 *Registration & Coffee*

09:30–09:40 *Welcome remarks*

Panel 1: Culture

09:40–10:30 **Charles Efferson**, University of Lausanne
The Strategy Space Problem in Gene-Culture Coevolution

10:30–10:55 **Sergey Gavrilets**, University of Tennessee **(short talk)**
Cultural tightness and the democracy premium in peer-enforced cooperation

10:55–11:25 *Coffee break*

11:25–12:15 **Jenna Bednar**, University of Michigan
[TBA]

12:15–13:30 *Lunch*

Monday June 1 - Afternoon

13:00–14:25 ***Poster session***

Panel 2: Norms

14:25–15:15 **Cristina Bicchieri**, University of Pennsylvania
Asymmetric Inferences from Norm Information: Behavior Valence and Social Consequences Shape the Direction of Is–Ought Inference

15:15–15:40 **Ingela Alger**, IAST/TSE **(short talk)**
Norms and norm change - driven by social preferences and Kantian morality

15:40–16:10 *Coffee break*

16:10–17:00 **Vincenz Frey**, University of Groningen
Norm perception and public goods production in networks

17:00–17:05 *Closing remarks & preview of Day 2*

19:30 *Dinner (by invitation only) - at Terra Tolosa*

Tuesday June 2 - Morning

09:00–09:30 *Welcome Coffee*

Panel 3: Macro-consequences

09:30–10:20 **Kaleda K. Denton**, Santa Fe Institute
Conformity to popular, not average, opinions

10:20–10:45 **Alexandros Gelastopoulos**, IAST/TSE **(short talk)**
The marginal majority effect: When social influence produces lock-in

10:45–11:15 *Coffee break*

11:15–11:40 **Marijn Keijzer**, IAST/TSE **(short talk)**
From preference to dominance: The reinforcing and regressing effects of social influence

11:40–12:30 **Pantelis Analytis**, University of Southern Denmark
Social influence and collective wisdom in sequential estimation

12:30–12:40 *Group photo*

12:40–14:00 *Lunch*

Tuesday June 2 – Afternoon

Panel 4: Learning

14:00–14:50 **Andrea Baronchelli**, City University of London
Conventions in Human and AI Populations: Emergence, Stability, and Collective Bias

14:50–15:40 **Sara Constantino**, Stanford University
[TBA]

15:40–16:10 *Coffee break*

16:10–16:35 **Maxime Derex**, IAST/TSE **(short talk)**
Social learning preserves both useful and useless theories by canalizing learners exploration

16:35–17:25 **Thomas J. H Morgan**, Arizona State University
Prestige and the Emergence of Adaptive Inequality in Social Influence

17:25–17:30 *Closing remarks*

19:30 *Dinner (by invitation only) – at Reflets*

Abstracts

Charles Efferson, University of Lausanne

The Strategy Space Problem in Gene-Culture Coevolution

When translating from social learning to aggregate dynamics, a core principle is that the details of how we learn from each other shape cultural evolution. The challenge is that, even for simple learning problems, the scope for complexity is spectacular. We could rely on empiricism to tame the complexity by asking, What social learning strategies do people actually use? I review behavioral experiments showing that, though strategies are somewhat systematic, they also tend to be wildly heterogeneous, and by extension existing models are wide the empirical mark. We could also rely on theory to tame the complexity by asking, What strategies evolve in a gene-culture model? This approach runs aground in analogous fashion. In particular, the strategies that evolve are sensitive to arbitrary choices the modeler makes about which strategies can or cannot arise via mutation, and this remains true even if the modeler is only adding or subtracting strategies that are bad strategies in a well-defined sense. I conclude by arguing that we need an integrated approach to experimental work and modeling with a clear emphasis on the strategy method in experimental studies. I further argue that for some questions, especially those related to applied cultural evolution or cultural group selection, we may simply want to jump to analyses at the aggregate level. The empirical challenge in this case is isolating cultural variation from other sources of group-level variation.

Sergey Gavrilets, University of Tennessee

Cultural tightness and the democracy premium in peer-enforced cooperation

Many prosocial norms arise not from authorities but from collective decisions among peers, from local commons governance to online communities. Yet the procedural effect of peer-designed rather than externally imposed rules—the democracy premium—does not operate uniformly across settings. I develop a mathematical model in which individuals choose whether to cooperate and punish based on material payoffs, cognitive dissonance, and social influence. Long-term outcomes depend strongly on the weight of nonmaterial motives relative to material payoffs, which determines how strongly early actions and observations feed back into later attitudes and behavior. Cultural tightness shapes these dynamics by changing the strength of social coordination, making change relatively gradual in loose settings but increasingly threshold-like and path dependent in tighter ones. Peer design enters through two channels: it can raise early buy-in and expectations, and it can reshape social feedback by altering how observed compliance and sanctioning influence incentives and expectations. While the first channel has positive effects across parameter values, the effects of the second channel are often confined to limited parts of the parameter space and are near zero elsewhere. Where they do matter, they are typically positive in loose cultures but can become negative in tight ones. Increasing the weight of nonmaterial motives makes the

democracy premium more negative. The framework yields testable predictions about when peer-designed rules promote cooperation and when they may backfire.

Cristina Bicchieri, University of Pennsylvania

Asymmetric Inferences from Norm Information: Behavior Valence and Social Consequences Shape the Direction of Is–Ought Inference

Social norm interventions – messages conveying what people do or what people approve of – are widely deployed to promote prosocial behavior. These interventions assume that recipients will update their beliefs in the intended direction. We demonstrate that peoples inferences from norm information are systematically asymmetric and depend on behavior valence. For positive behaviors, learning that others engage in such a behavior leads people to strongly infer that others also approve of it: common is moral. For negative behaviors, learning that others approve of a behavior leads people to strongly infer widespread practice: approved means practiced. Critically, this double asymmetry is amplified by perceived social consequences: the very behaviors with the greatest social stakes: the ones interventions most urgently target are precisely those for which the inference asymmetry is most pronounced. Additional moderation by cost type (moral and material) provides further nuance. These findings, robust across 22 behavioral domains in a large preregistered sample, reveal that the same norm message can strengthen or undermine intervention goals depending on how it interacts with receivers inferential processes.

Ingela Alger, IAST/TSE

Norms and norm change - driven by social preferences and Kantian morality

Norms indicate which behaviors are commonly expected and/or considered to be morally right. We propose a theoretical model of how such norms may come about and change, in a setting where a group of individuals face a collective action problem, formalized as a linear public goods game, where contributions are collectively rational but individually irrational from a purely material point of view. Each individual has (1) an idea of what is the "right thing to do"; (2) an attitude towards making a greater material sacrifice than others; (3) an attitude towards making a smaller sacrifice than others; (4) an attitude towards deviating from the "right thing to do". We show that these preferences uniquely determine each individual's thresholds for collective behavior, whereby an individual will contribute to the public good if and only if sufficiently many others do so. Depending on the importance attached to the said attitudes, some individuals, however, prefer not to contribute regardless of what the others do, while some prefer to contribute regardless of what the others do. The latter are leaders. We show that spontaneous norm change can occur if novel information reaches some such leaders.

Vincenz Frey, University of Groningen

Norm perception and public goods production in networks

The capacity of human collectives to produce public goods critically contributes to

human wellbeing and prosperity. Public goods are often global while individuals observe others only locally. For instance, in work teams, everyones effort affects the outcome, but individual team members observe only the effort of their immediate colleagues. We investigate how the observation network influences public goods production in a repeated game. Our model builds on two insights. First, people often adhere to social norms, contributing to the public good if they think that many others do so as well. Second, people often misperceive social norms when inferring global norms from local observation. Specifically, perceptions of the average contribution (the social norm) will be biased toward the contribution level of those who hold central positions in the observation network. Computational simulations, calibrated using existing behavioral data, predict that when cooperative individuals occupy central positions in the observation network, public good production levels are higher than when uncooperative individuals are central. Furthermore, simulation results suggest that levels of public goods production are higher in the scenario cooperators central (lower in the scenario freeriders central) than in a baseline with a degree-homogeneous observation network. Despite the theorys plausibility, results of an online experiment align relatively poorly with these predictions.

Kaleda K. Denton, Santa Fe Institute

Conformity to popular, not average, opinions

A continuous trait contains infinitely many variants on a spectrum, such as a spectrum of behaviors or ideologies. "Conformity" to such traits has been defined as the preference for the mean variant, even if this mean is not close to any individual variant (e.g., if half of the population falls on the far right and far left of a spectrum, respectively, the mean is in the center). Here, we define conformity as preference for clusters of popular variants, not average variants. Compared to trait-averaging models, this conformity model provides a better fit to empirical data on human decision-making under many of the conditions we analyze, and in simulations, it often produces different population-level outcomes such as faster shifts toward poles of a spectrum and reduced "wisdom of the crowd."

Alexandros Gelastopoulos, IAST/TSE

The marginal majority effect: When social influence produces lock-in

People are influenced by the choices of others, a phenomenon observed across contexts in the social and behavioral sciences. Social influence can lock in an initial popularity advantage of an option over a higher quality alternative. Yet, several experiments designed to enable social influence have found that social systems self-correct rather than lock in. Here, we identify a behavioral phenomenon that makes inferior lock- in possible, which we call the marginal majority effect: a discontinuous increase in the choice probability of an option as its popularity exceeds that of a competing option. We demonstrate the existence of a marginal majority effect in several recent experiments and show that lock- in always occurs when the effect is large enough to offset the quality effect on choice but rarely otherwise. Our results reconcile conflicting past empirical evidence and connect a behavioral phenomenon to the possibility of social lock- in.

Marijn Keijzer, IAST/TSE

From preference to dominance: The reinforcing and regressing effects of social influence

When adopting a behaviour, humans often rely on social information and are more likely to copy behaviours that are demonstrated by many other individuals. Despite numerous empirical demonstrations of such effects, the literature on social influence in binary choice remains disjoint, and there is a lack of a unifying framework for mapping the micro- and macro-level consequences of social influence. We propose a parsimonious and interpretable mathematical model of social influence and use it to reanalyse 294,703 choices by 20,480 individuals from an exhaustive collection of experiments on binary choice under social-influence. In addition, we define two new metrics that predict the possible macro-level consequences of social influence: preference amplification, expressing the reinforcement of the dominant behaviour, and reversal potential expressing the degree of bistability in the system, and the propensity of the inferior option to prevail. We show that in almost all experiments, social influence can lead to substantial amplification of the dominant preferences, whereas bistability appears to be possible in a subset of settings. Our approach can be used to contrast old and new work on social influence and to better understand the reinforcing or regressing effects of behaviour change interventions.

Pantelis Analytis, University of Southern Denmark

Social influence and collective wisdom in sequential estimation

In this paper, we develop a theoretical framework for studying how different aggregate information signals may affect the accumulation of collective wisdom in sequential estimation settings. We show that systems in which people have access to the mean previous response converge to a unique equilibrium that is the same as the mean of the response distribution without any social influence. Systems where people have access to the median response, by contrast, are path-dependent and quickly converge to any value within an interval, the size of which depends merely on the proportion of people that copy the median. As a result, path dependence increases and collective wisdom deteriorates as a function of the proportion of people copying the median. We assess these two new theoretical predictions about the median using the data from a recent large-scale experiment on crowd wisdom and show that they are both corroborated by the data.

Andrea Baronchelli, City University of London

Conventions in Human and AI Populations: Emergence, Stability, and Collective Bias

I present a unified framework based on the naming game to study how conventions emerge, stabilise, and change. I start by discussing controlled experiments with human participants, showing how shared conventions can arise from local interactions, and how small committed minorities can trigger abrupt transitions once a critical threshold is reached.

I then extend the same framework to populations of large language model (LLM) agents, showing that these systems also converge on shared conventions without central coordination. Strikingly, this process can generate systematic collective biases even when individual agents display no discernible bias in isolation.

These results reveal a general mechanism linking microscopic coordination to collective behaviour across human and artificial systems. As AI systems interact at scale, they highlight the need for population-level evaluation, with implications for alignment, governance, and socio-technical design.

Maxime Derex, IAST/TSE

Social learning preserves both useful and useless theories by canalizing learners exploration

In many domains, learning from others is crucial for leveraging cumulative cultural knowledge, which encapsulates the efforts of successive generations of innovators. However, anecdotal and experimental evidence suggests that reliance on social information can reduce the exploration of the problem space. Here, we experimentally investigate the extent to which cultural transmission fosters the persistence of arbitrary solutions in a context where participants are incentivized to improve a physical system across multiple trials. Participants were exposed to various theories about the system, ranging from accurate to misleading. Our findings indicate that even under conditions conducive to exploration, the transmission of cultural knowledge canalizes learners' focus, limiting their consideration of alternative solutions. This effect was observed in both the theories produced and the solutions attempted by participants, irrespective of the accuracy of the provided theories. These results underscore the significant role of cultural transmission in shaping human knowledge and technologies.

Thomas J. H Morgan, Arizona State University

Prestige and the Emergence of Adaptive Inequality in Social Influence

Who individuals choose to learn from shapes population-level influence hierarchies. Combining a cultural evolutionary model, laboratory experiments, and evolutionary simulations, we show that "prestige" – the tendency to defer to successful or esteemed individuals – generates highly skewed influence hierarchies. Small initial differences in prestige can be amplified through positive feedback, producing outcomes ranging from egalitarian to near-autocratic structures depending on a population's sensitivity to prestige. Experimental evidence with human participants reveals strong prestige sensitivity, sufficient to generate substantial inequality in influence, even in small groups. Moreover, this sensitivity appears adaptive: when direct information about others' competence is limited, collectively generated prestige provides a superior indirect signal of quality. Crucially, the resulting hierarchies are non-coercive and can improve information aggregation yet still produce large disparities in influence. These findings identify a micro-level psychological mechanism through which social learning can generate persistent macro-level inequality in influence.

Papers

Some authors have shared a paper that is relevant to their presentation. These papers can be found in **this google drive folder**.

Posters

Aleix Nicolás Olivé, Pompeu Fabra University

Network Segregation and News Propagation: A Computational and Experimental Framework

Social influence and network structure play a critical role in the macro-level diffusion of information and the formation of polarized beliefs. We propose a novel experimental framework, currently supported by multi-agent simulations, to investigate how echo chambers affect individual news diets and the collective spread of media. Participants, whose baseline opinions on issues such as climate policy are assessed via survey, are mapped to a pre-generated network featuring two highly connected communities linked by structural bridges. We manipulate the social environment across two regimes: an integrated setting (random node assignment) and a segregated setting (assortative assignment by opinion). In the proposed multi-round, asynchronous experimental design, nodes are seeded with news articles of varying leanings. To simulate realistic attention constraints, users receive a curated feed capped at a maximum of 4 items per round, with items shared by multiple neighbors having a higher probability of appearance. At each step, participants choose to share or dismiss items. With this setup we can study the news diet of the participants in different conditions as well as the virality of news with different leanings. By tracking the average and max feed size, alongside the accumulating backlog size of unseen items, we capture the dynamics of information overload and visibility constraints. Currently, the structural mechanics of this diffusion process have been evaluated through multiple simulations. We are expanding this framework to human agents in order to validate the results of our simulations. This poster will present findings from our simulation models alongside the experimental pipeline, offering a methodological bridge to study the interplay between network topology, cognitive constraints, and path-dependent information cascades.

Cameron Curtin, IAST/TSE

Traditional political institutions, cooperation, and public goods in Oaxaca, Mexico

The cross-cultural variation in human cooperation is an evolutionary puzzle. Cultural evolution researchers have proposed that this variation can be traced to differences in social norms and institutions. However, the role of norms and institutions in human cooperation remains a topic of debate. Here, we leverage a unique natural laboratory to study the relationship between institutions and cooperation. In the highly culturally diverse state of Oaxaca, Mexico, there is broad variation in the traditional indigenous political institutions by which communities self-govern. Drawing on cultural evolutionary theory, we preregistered the hypothesis that communities with institutions that (1) comprise more powerful norms about service to the community, (2) levy harsher sanctions against defectors, and (3) build more interdependence are better able to mobilize collective action and provide public goods. To test

this hypothesis, we applied a preregistered econometric approach to a new dataset that combines secondary data and newly quantified ethnographic data for 418 Oaxacan communities. Results reveal that communities with stronger institutions mobilize more cooperation for the group benefit, such as migrant remittances for public works projects. However, we find little evidence that this cooperation translates into tangible public goods outcomes. This study provides support for the cultural evolutionary hypotheses that link institutions to cooperation and raises new questions.

Felipe Perilla Reyes, University of Zurich

How Dynasties Win:Network Social Capital and Electoral Outcomes in Colombia

Who becomes a candidate, who gets elected, and how? The answers to these questions largely determine outcomes in substantive and descriptive representation. Existing research has causally identified a dynastic advantage in terms of who runs and wins elections, yet it typically defines dynasties only through kinship ties between senior and junior politicians. This approach obscures how social positions within elite networks influence resource accessibility and, in turn, electoral outcomes. Using newly collected longitudinal data on Colombian elite kinship networks, I analyze how network social capital (measured through closure and brokerage) affects both candidacy and electoral success in presidential elections (1833–1942) and congressional elections (1888–1951). Leveraging accidental deaths as exogenous shocks to the network structure, I identify how changes in social capital shape electoral performance. I find preliminary evidence that both closure and brokerage increase selection and election outcomes. I thus explain how dynastic advantages arise and become a two-stage barrier to political entry that underlies persistent and suboptimal political and economic development outcomes.

Filippo Zimmaro, Institute of Cognitive Sciences and Technologies - CNR Italy

Cultural tightness and social cohesion under coevolving beliefs, behaviors and preferences

Successful collective action, from climate change mitigation to the maintenance of democratic institutions, depends critically on societal properties such as cultural tightness and social cohesion. Yet how these properties emerge and evolve remains poorly understood, because they arise from a complex, interdependent interplay between beliefs and behaviors, that are usually modeled in isolation. We address this challenge by developing a game-theoretic framework that integrates norm-utility models to study the coevolutionary dynamics of cooperative behavior, expressed belief, and norm-utility preferences. Our model expands the action space of classical games (Snowdrift and Prisoner's Dilemma) so that each player simultaneously chooses a material behavior and an expressed belief, with utility combining both material payoffs and a psychological component arising from preferences for conformity (aligning with others' beliefs or behaviors) and coherence (aligning one's own belief with one's own behavior). A key feature of our approach is that these preferences are not fixed exogenously, but evolve endogenously through social learning alongside beliefs and behaviors. This leads to a proliferation of evolutionary stable equilibria, each with distinct societal signatures. Most strikingly, we find that

shifts toward a more demanding material environment, such as higher costs of cooperation, can increase cultural tightness while simultaneously reducing social homogeneity and cohesion. The latter occurs via evolutionary branching, whereby an initially homogeneous population fragments into two distinct groups differing not only in their willingness to cooperate, but also in their expressed beliefs and in their preferences for conformity versus coherence. These results illuminate a fundamental tension in the dynamics of social norms: the conditions that tighten a culture around a shared norm can also be the conditions that fracture society into groups with irreconcilable worldviews. Our framework offers a unified, tractable approach to connecting individual-level social influence with macro-level collective outcomes.

Jonathan Diez, University of Barcelona

Assessing infrastructural resilience under extreme weather events: a case study of the Valencian DANA

On October 29, 2024, the DANA floods hit Valencia in an unprecedented manner. Rains in the region accumulated and led to large floods that resulted in physical damage, displacement and isolation, and the death of 230 people. The flooding lasted only a couple of weeks, but the effects are noticeable more than a year later. The disaster resulted in collapse of mobility, information flows, and social coordination. Roads became impassable, communities were isolated and fake news spread leading to the allocation of resources based on false information. At the same time, a huge solidarity effort emerged. Thousands of people mobilized to remove debris and formed networks of mutual aid using digital platforms to coordinate and stay informed. In this project we analyze the collapse and the following recovery of these resources and networks during the DANA, that we conceptualize as three interdependent infrastructures and analyze them using satellite data, mobility data, and social media analysis. Physical infrastructure: Reconstructing road network and travel times using satellite images and phone mobility data. Civic infrastructure: Coordination and communication between volunteers through telegram groups and channels. Information infrastructure: Rapid share of public understanding through TikTok and Twitter. We show how and when areas with damaged physical infrastructure are discussed in coordination channels of volunteers and how information is shared about these areas in social media. This allows us to understand how these different infrastructures are correlated. With this we want to lie the basis for a general pipeline to analyze civic response to extreme weather events.

Julian Berger, Max Planck Institute for Human Development

Are large language models rewriting history?

Large language models (LLMs) are currently altering how humans retrieve and update knowledge. While recent studies suggest LLMs may homogenize human expression, their consequences on collective memory and the production of historical knowledge remain largely unexplored. This project investigates how reliance on LLMs affects the recall and transmission of historical and political narratives. We propose a human subject experiments comparing individuals recalling historical events

with LLM assistance, search engine assistance, or no assistance. We test three primary hypotheses: LLM-assisted recall will yield a greater breadth of recalled entities, exhibit a stronger bias toward Western actors and events, and result in a less diverse set of entities globally compared to traditional search methods. Currently, we quantify how humans and LLMs respond differently to a set of 720,000 questions asked on the subreddit r/AskHistorians. From here, we will select a subset of topics where LLMs are either aligned with humans but also where we detect especially divisive or biased responses. These topics will fuel our experiments. Importantly, to understand how the extent to which LLMs impact individual and collective opinion-making evolves, we introduce a transmission chain setup to our experiment. In this setup, participants sequentially revise and expand upon previous answers, mirroring the cumulative effort underlying knowledge commons like Wikipedia. By integrating humans and LLMs in these transmission chains, we measure how LLM assistance alters sequential contributions to knowledge commons and shapes what populations collectively accept as historical knowledge. This research directly aligns with the workshop's theme by treating LLMs as novel agents of macro-level social influence. We aim to explore whether human collectives run the risk of adopting biased or homogenized historical narratives through interactions with LLMs, providing new insights into the cultural evolution of knowledge and its production in the digital age.

Marcus Vinícius de Sá Torres, Federal University of Pernambuco

Family Ties How political dynasties shape federal transfers in Brazil

What is the effect of political dynasties on federal transfers to municipalities? In Brazil, most public resources are concentrated at the federal level, and ministries exercise considerable discretion in allocating funds to municipalities through voluntary transfers. This research investigates whether mayors belonging to political families, that is, those with relatives who have held elected office, secure more federal resources than mayors without such ties. The central hypothesis is that dynastic mayors obtain more federal funding through voluntary transfers because they benefit from political capital that broadens their access to the federal state apparatus, whether through prior knowledge of bureaucratic rules or through political networks that facilitate resource acquisition. In particular, mayors with relatives in the federal legislature are expected to secure larger amounts, given their proximity to actors who shape the distribution of federal funds. To identify political families, we employ a novel data-collection strategy that uses OCR-based web scraping to extract background certificates from the TSE Candidacy System. Using the family declarations from the 2020 elections, we trace kinship ties to candidacies from 1994 to 2024. We combine this dataset with information on federal voluntary transfers from 2021 to 2023 and estimate the effects of political lineage using linear regressions with year fixed effects. Results suggest that, although the effect is not universal, dynastic mayors obtain fewer resources on average than their non-dynastic counterparts within municipalities.

Paul Bergmann, University of Marburg

The Fortunate Few Misleading the Rest: Social Influence Undermines Perfor-

mance and Increases Inequality When Decision Outcomes are Path-Dependent

Individuals often gauge the risks and benefits of choice options from both their own experiences and the behaviours and outcomes of others. In scenarios where outcomes are independent and accumulate additively over repeated choices, social influence has been shown to improve collective decision-making relative to individual learning alone. However, across many key real-life contexts, ranging from financial investments, occupational decisions, or social networks, outcomes do not independently add up over time but depend on how many resources agents already have. Such multiplicative situations are characterized by potentially escalating "the-rich-get-richer"-dynamics creating immense disparities in wealth. The individual and collective consequences of social influence under path-dependent growth remain unknown. Using agent-based simulations relying on social reinforcement learning, we examine the effects of five social learning strategies: unbiased and conformist copying, relative and absolute payoff-biased copying, and an aggregate wealth-bias, across both independent (additive) and path-dependent (multiplicative) decision environments. We show that while all forms of social influence improve decision-making when outcomes accumulate additively, in multiplicative systems, performance declines when agents can learn from each other. Large positive outcomes occurring by chance reinforce otherwise suboptimal risk-taking among a fortunate few and guide others towards more risk taking, thereby producing worse outcomes for most, and accelerating population-level inequality. This effect is strongest when agents learn from choices of wealthier group members or those most successful in the last round but occurs under all forms of social influence with the exception of hyper-conformity. Social influence even undermines performance when both private and social payoff information are fully log-transformed to adjust for the heightened sensitivity to variance inherent in multiplicative accumulation. Our findings reveal fundamental difficulties in escaping maladaptive collective learning under stochastic multiplicative growth and highlight that wealth inequality is exacerbated under nearly all forms of social influence when success is partially due to chance.

Petr Krautwurm, Prague University of Economics and Business

Evidence on Team Composition and Performance: Changes Keep Performance High

This paper studies team performance trajectories after a change to team composition. A large data set from a competitive team based online game is analysed using semi-parametric generalised additive mixed-effects models. The data show that teams (five members) overperform directly after a change, after which performance dips and does not recover. The effect is independent of the size of the change, i.e. of how many members are already known to each other from the previous stage. Moreover, over the average lifespan, team performance declines by about 16 percent of the interquartile range of overall team performance. From an applied perspective, the data supports the view that small teams targeting tasks for which performance is easy to evaluate are best newly put together on a case by case basis.

Stavros Anagnou, University of Hertfordshire

In search of machine normativity: towards a cognition-agnostic approach for studying norms in AI collectives

As large language models and other forms of artificial intelligence are increasingly integrated into institutions and society, understanding their emergent behaviour becomes critical, whether in the context of political bots influencing elections or pricing algorithms tacitly colluding with one another. Norms are a key ingredient for understanding and controlling the emergent behaviour of AI (and human-AI) collectives. However, existing norm definitions are insufficient for this task. As Westra et al. (2024) have argued, prevailing definitions assume a human norm psychology, including capacities such as shared intentionality and advanced belief attribution. This human-centred framing risks overlooking norm-like processes arising in non-human entities through different cognitive means. Drawing on work from comparative normative psychology, we propose a cognition-agnostic framework that does not presuppose any particular psychology. Our definition is behavioural, focusing on group-level behavioural regularities rather than individual-level psychological capacities as evidence of norms. This avoids the pitfalls of a human-biased definition and enables a more unbiased investigation of normative capabilities in AI agents. Previous work in comparative norms has distinguished between behavioural regularities and genuinely normative ones by whether they are socially maintained (i.e. through group pressure). We operationalise social maintenance as comprising three elements: minimal social interaction (measured by information flow), stake, and mechanisms to affect stake. We further add and justify a multiple equilibria criterion, capturing the fact that norms do not converge on a single equilibrium but represent one among many possible stable states; a defining feature of normative behaviour. We then highlight flaws in previous LLM norm research, including attempts to identify human-like psychology as evidence of norms or to externally load motivations into agents. Finally, we propose an experimental design to distinguish between normative and mere behavioural regularities in LLMs and discuss how it allows us to probe the "norm psychology" of LLMs.

Valeria Widler, Zuse Institute Berlin & Freie Universität Berlin

A multi-level causal mechanism of political opinion dynamics: Polarization as incomplete collective inference?

How do decentralized social interactions allow societies to aggregate distributed information into coherent collective beliefs, and under what conditions does this process lead to persistent polarization? I present a framework that treats political opinion dynamics as a process of collective inference in which decentralized agents collectively construct shared representations of their social environment through locally mediated information exchange. In this framework, consensus arises when distributed information is effectively aggregated across the population, whereas polarization reflects a breakdown in this collective inference process. I investigate what sequence of dynamical processes leads from local interactions to global consensus to gain a complete picture of the causal mechanism. I introduce an agent-based model in which agents update beliefs through evidence accumulation while selec-

tively weighting information from different social groups. Identity-based social influence restricts information integration to local communities, while migration of agents allows information to propagate across community boundaries. The model therefore captures how the informational connectivity of social networks emerges from identity-based social learning at the individual level and its effect on collective belief formation. The analysis identifies distinct regimes and phases of collective belief formation. Consensus emerges through a multilevel causal mechanism linking individual evidence accumulation and the emergence of downward causation. By decomposing this process into its constituent phases, the study characterizes the conditions under which collective information integration succeeds. I will discuss how the mechanism might determine whether a system fractures into polarized camps.

Yick Chung, University of Milan

The functional specificity hypothesis: the role of prosocial emotions in the evolution of cooperation

Why did evolution favor multiple prosocial emotions instead of a single mechanism? Using an evolutionary agent-based model, the authors test the functional specificity hypothesis, examining three mechanisms: guilt (partner-specific expectation violation), kindness (cooperation reward active regardless of norms), and identity (self-discrepancy from internalized standards). Across 3,000 simulations spanning stranger- to partner-dominated environments, guilt acts as an environmental specialist (cooperation 5.9%-84.2%), kindness as a generalist (34.9-41.9% regardless of environment), and identity as cooperation-dependent and not self-sustaining (zero cooperation in identity-only populations under asymmetric weighting). Under selection over 500 generations, kindness alone shows universally positive selection, identity is universally selected against, and guilt declines 66% less in partner-rich environments. The results identify two dimensions of variation — specialists vs. generalists, self-sustaining vs. cooperation-dependent - explaining why evolution maintained mechanism diversity rather than converging on a single cooperation mechanism.

Yuxin Liu, Princeton University

Genetic Selectivity in Migration: Polygenic Insights into Heterogeneous Outcomes of Domestic Migration in the US

Migration studies often grapple with selection bias and unobserved confounders, such as education, socioeconomic status (SES), and health, which complicate the understanding of migration outcomes. This study leverages genetic data to address these challenges, utilizing polygenic scores (PGS) to examine the genetic selectivity of internal migration in the United States and its implications for mobility behaviors. Employing data from AddHealth, the research investigates the genetic profiles of U.S. migrants, focusing on traits such as migration distance, educational attainment, health, and risk-prone behaviors. Specifically, it explores whether long-distance migrants exhibit positive genetic selectivity (e.g., higher PGS for educational attainment and health) while short-distance, frequent movers display negative selectiv-

ity (e.g., higher PGS for risk-prone behaviors and lower SES). The study proceeds in three steps: (1) validating the robustness of migration distance PGS across datasets, (2) comparing genetic profiles across migration groups to assess the direction and magnitude of genetic selectivity, and (3) employing within-family designs to establish causal relationships between genetic traits and migration behaviors. By integrating genetic insights with sociological perspectives, this research aims to reconcile two competing theories of migration outcomes—one emphasizing the advantages of positively selected migrants and the other highlighting the challenges faced by disadvantaged movers. The findings are expected to provide a nuanced understanding of migration selectivity, demonstrating how genetic endowment shapes diverse migration patterns and outcomes. This study offers innovative methodological contributions and practical implications for addressing disparities among migrant populations in the U.S.

Participants

Ingela **ALGER**
Stefan **AMBEC**
Stavros **ANAGNOU**
Andrea **BARONCHELLI**
Jenna **BEDNAR**
Julian **BERGER**
Paul **BERGMANN**
Cristina **BICCHIERI**
Adrien **BLANCHET**
Adam **BODE**
Sylvie **BORAU**
Isabelle **BROCAS**
Charlotte **CAVAILLÉ**
Victor **CHAVEZ**
Yick **CHUNG**
Sara **CONSTANTINO**
Camélia **CROZAT**
Cameron **CURTIN**
Philippe **DE DONDER**
Kaleda **DENTON**
Maxime **DEREX**
Konrad **DIERKS**
Jonathan **DIEZ**
Ray **DUCH**
Charles **EFFERSON**
Jesse **FENNEMAN**
Juan **FERRER**
Vincenz **FREY**
Sergey **GAVRILETS**
Victor **GAY**
Alexandros **GELASTOPOULOS**
James **HAMMITT**
Marion **HOFFMAN**
Marijn **KEIJZER**
Petr **KRAUTWURM**
Tadija **LATINOVIĆ**
Jojo **LEE**
Shuai **LIU**
Yuxin **LIU**
Emiliano **LORINI**
Ludovic **MAISONNEUVE**
Catherine **MOLHO**
Thomas **MORGAN**
Maximilian **MÜLLER**
Aleix **NICOLÁS OLIVÉ**
Campbell **OLYMPIA**
Jorge **PEÑA**
Alfonso **PÉREZ ESCUDERO**
Felipe **PERILLA**
Pantelis **PIPERGIAS ANALYTIS**
Ilaria **PRETELLI**
Lorenzo **RAVIZZONI**
Arnaud **REYNAUD**
Juliette **SAETRE**
Paul **SEABRIGHT**
Alberto **SIMPSER**
Clément **SIRE**
Jonathan **STIEGLITZ**
Máté **SZILASSI**
Guy **THÉRAULAZ**
Marcus **TORRES**
Nicolas **TREICH**
Dari **TRENDAFILOV**
Pablo **VARAS ENRIQUEZ**
Valeria **WIDLER**
Chiara **ZANARDELLO**
Filippo **ZIMMARO**