The rise and fall of technological development in virtual communities

Natalia Vélez[1*], Charley M. Wu[2,3,4†], Samuel J. Gershman[5], Eric Schulz[6†]

[1*]Department of Psychology, Princeton University, Princeton, NJ, USA.
[2]University of Tübingen, Tübingen, Germany.
[3]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany.
[4]Department of Computational Neuroscience, Max Planck Institute for Biological Cybernetics, Tübingen, Germany.
[5]Department of Psychology and Center for Brain Science, Harvard University, Cambridge, MA, USA.
[6]Helmholtz Institute for Human-Centered AI, Munich, Germany.

*Corresponding author. E-mail: nvelez@princeton.edu;
Contributing authors: charley.wu@uni-tuebingen.de; gershman@fas.harvard.edu; eric.schulz@helmholtz-munich.de;
†These authors contributed equally to this work.

Abstract

Humans have developed technologies to adapt to virtually every habitat on Earth. But why do some communities develop thriving technological repertoires while others stagnate? We address this question by analyzing player behavior in One Hour One Life (OHOL), a multiplayer online game where players can build technologically advanced communities from scratch (N = 22,011 players, 2,700 communities, 428,255 playthroughs). Players are randomly assigned to a community in each playthrough and can contribute to it for up to one hour. Over time, through many players' contributions, communities can survive for weeks and amass rich technological repertoires. Thus, this dataset provides a unique quasi-experiment into how the composition of communities affects their growth and decline. Using this approach, we find that technological developments are the product of interactions between individuals and the communities where they are placed. Individuals take on jobs that align with those of their closest peers, and they selectively contribute new technologies in areas of their expertise that diverge from the rest of the community. In the aggregate, these two processes—aligning to form specialized communities, or diverging to form diverse ones—have opposing effects on the size and stability of a community's technological repertoire, and imbalances in specialization serve as an early indicator of population collapse. Our results suggest that, to survive, communities must balance between diversifying to develop new technologies, while specializing to maintain the ones they already have. Our approach provides a testbed for theories of large-scale social phenomena that would otherwise be difficult to test against real-world data or traditional laboratory experiments.

1 Introduction

Humans have adapted to virtually every habitat on Earth by developing rich technological repertoires [1, 2]. Folk tales about the origins of these repertoires often tell simple stories that highlight the contributions of a single, brilliant innovator: Odin discovered the runes, Jamshid invented medicine, Spider Woman passed down the secrets of pottery and agriculture, and so on. However, the weight of existing archaeological, ethnographic, and experimental evidence suggests that creating new technologies is anything *but* a solitary activity [3–5]. Contrary to the myth of the lone innovator, technological progress relies on specialized knowledge and labor and flourishes in large communities that place diverse industries in close proximity to one another [6–8]. It is typically assumed that these conditions create opportunities for individuals to develop new technologies by sharing expertise and labor [9, 10]. However, it is surprisingly challenging to test this assumption directly using existing methods. How do individuals unlock the opportunities for technological development that communities provide? Why do some communities grow and flourish, while others decline?

To address these questions, it is important to test technological development within an environment where we can observe both small-scale, individual behaviors and large-scale, community-wide dynamics. This is not possible with existing methods, which instead tend to prioritize a single level of social organization [5, 11]. For example, real-world data from archaeological, historical, and economic records offer snapshots of technological repertoires developed by large communities over long timescales (e.g., [10, 12, 13]. Yet these snapshots are often incomplete; outside of isolated case studies of famous figures, it is difficult to trace individual contributions to a community's development (e.g., [13]; but see [14]). By contrast, laboratory experiments allow researchers to study individual behavior with near-complete information (e.g., [15, 16]). However, it is difficult to emulate slow, large-scale processes of real-world technological development in short, simplified tasks. Here, we present a new and complementary approach that provides unprecedented access into both individual and community-wide contributions to technological development.

We analyzed player behavior in One Hour One Life[1] (OHOL), a multiplayer online game where players can create technologically advanced communities over a massive, multi-generational collaboration (22,011 players, 2,700 communities, 428,255 lives lived; Fig. 1). Each time a player starts a new session, they are randomly placed into a community. This random assignment provides a unique quasi-experiment into how the composition of communities affects their growth and decline. At the smallest scales, we can trace individual behavior with finer detail than is typically possible in real-world datasets, including the expertise that players have developed over prior lifetimes, the communities in which they have lived, and the players with whom they have interacted. Yet we can also observe slow, large-scale social dynamics at a scale beyond what is typically observed in laboratory studies: Communities can persist for several weeks in real time, spanning thousands of generations of players. With these data in hand, we can now examine how individual and community-wide processes of technological development support one another. Below, we examine three fundamental processes in technological development that cut across these scales: how individuals adapt their activities to the rest of the community, how community-wide technological repertoires are built up from individual contributions, and how division of labor across communities affects their growth and decline. Together, our work provides an integrated view of technological development across levels of social organization.

2 Results

2.1 Studying technological development in a microcosm

When a player starts a new session of OHOL, they are randomly assigned a new character. Each character starts out as the infant of another player whose character has already reached adulthood. In infancy, players are entirely dependent on the care of their "mother" and of other, nearby players. Characters age by one year in-game for every minute that passes in real time—gaining increasing autonomy through childhood, adolescence, and adulthood—and can live for up to one hour (Fig. 1b). Over time, through many links between parents and offspring, players can create enduring family lineages (Fig. 1c). Families are at the core of social interactions in OHOL. Each character is uniquely identified by their familial

[1] One Hour One Life was developed, published, and maintained by Jason Rohrer. This work is not affiliated with One Hour One Life; we are grateful to Mr. Rohrer for making the game's logs and source code publicly available.

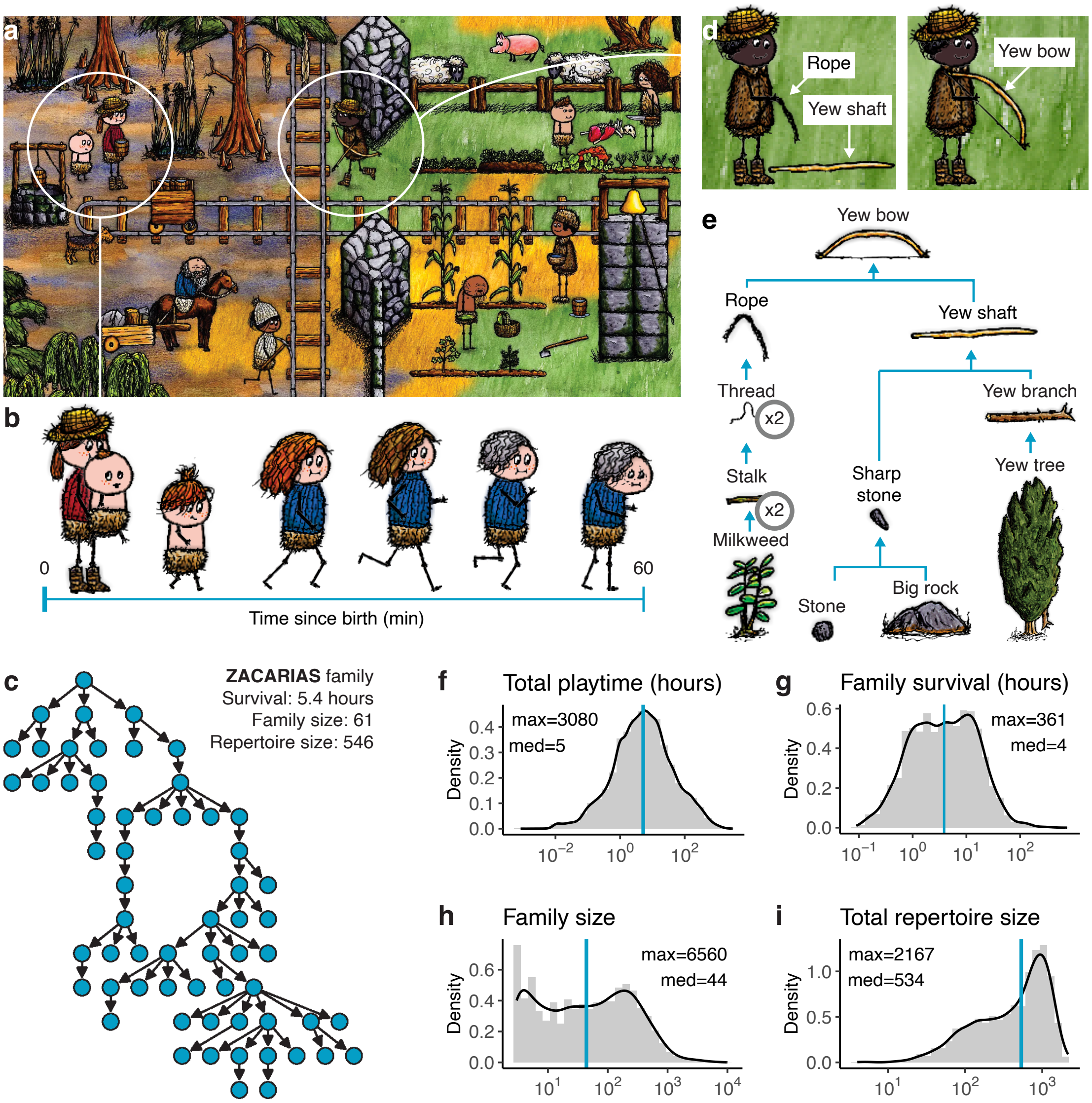


Fig. 1 **Overview of game mechanics.** **a,** Bird's-eye view of a representative community. **b,c** Family mechanics: **b,** New characters are born as the infant of another player whose character has reached adulthood. Characters age by one year for every minute that passes in real time, for up to 60 minutes. **c,** Family tree for a representative community: The top node shows the founding member, and each edge connects a character to their offspring. **d,e,** Crafting mechanics: **d,** New technologies can be created by combining simpler parts. **e,** Full procedure to make a yew bow, starting from naturally occurring primitives. **f–i,** Distributions of: **f,** playtime (i.e., hours that individual players have spent playing the game), **g,** family survival (i.e., hours between the birth of the founding member and the death of the last descendant), **h,** family sizes (i.e., total characters born to a community), and **i,** total repertoire sizes (i.e., total number of unique technologies used by a family over its whole history). Blue vertical lines show median values for each measure; insets show maximum (max.) and median (med.) values.

relationships. The game obscures individual players' identities, such that when players hover their mouse over a character, they do not see the player's username or other identifiers that are stable across sessions, but rather a temporary name and surname assigned by that character's mother and a small explanation of the character's familial relationship to them (e.g., "Elizabeth Bennet - You", "Jane Bennet - Your Sister"). Of the 9,384,198 social interactions we observed between characters, 71.09% of occurred between kin, and characters lived on average closer to kin than to non-kin (Fig. S1). Thus, throughout the paper, we will operationalize *communities* as family lineages. After applying exclusion criteria (see Supplementary Methods), we obtained a sample of player behavior between November 2019 to June 2020, spanning N = 2,700 communities, comprising 428,255 characters, operated by 22,011 unique players.

OHOL gives players few explicit goals, other than surviving to old age. Yet surviving is not trivial: Only 34% of characters in our sample survived to old age, while the rest starved (62%) or were murdered by other players (4%). To survive, players need to develop reliable food sources, clothing, tools, and shelter. The game map is seeded with natural resources including wild plants, animals, minerals, and bodies of water. From these resources, players can create new technologies by combining simpler parts. For example, players can pick milkweed to obtain fibers, combine the fibers to make string, combine the strings to make a rope, and tie the rope around a shaft carved from a yew branch to make a bow (Fig. 1d–e). Objects that players create in the game persist until they are spent or transformed; thus, players can build upon tools and structures that were first developed by their forebears generations ago.

We propose that the growth of communities in OHOL is analogous to real-world processes of technological development for two reasons. First, technological development in OHOL is best understood not as the work of lone geniuses, but rather as a multi-generational, community-wide collaboration. Because individual characters only live for up to an hour, the most advanced technologies cannot be developed by individual characters entirely from scratch. Advanced technologies, such as woven fabrics, cameras, cars, and radio transmitters, can require 353–502 unique ingredients to make and were developed, on average, 6–17 hours after a community's inception, many generations after the founding member had passed away (Fig. S2). Much as in real-world technological development, the primary problem faced by communities is not discovering technologies anew, but rather building capacity to produce technologies locally [17–19]. Second, everything that players create is prespecified by the game's technological tree, and players may have insights into the tree's structure from real-world intuitions and from discussions with other players. Thus, when a community first builds a car, we cannot determine whether they arrived at this technology *de novo* through serendipitous recombination, as is typically assumed in cultural evolutionary processes [3, 20, 21]. Instead, this development is akin to establishing an automotive industry in a new region [22, 23].

Our approach offers several advantages over existing methods for studying technological development [24, 25]. First, players invest considerable time and effort into the game. The most experienced players have logged thousands of hours of playtime (Fig. 1f); thus, it is possible for players to develop deep expertise in this game over longer timescales than what is typically feasible in laboratory studies. Second, players form communities that are larger, more long-lived, and more technologically complex than what can feasibly be created in laboratory studies. The longest-running community in our sample survived continually for 15 days (361 hours; median: 4 hours; Fig. 1g), and the largest had 6,560 members over its entire history (median: 44; Fig. 1h). With 3,798 craftable technologies, the game's technological tree is sufficiently complex that communities can also amass impressive technological repertoires (median: 534 unique technologies used during the family's entire history; max: 2,167 technologies; Fig. 1i). Finally, compared to real-world datasets, OHOL affords the unique opportunity to study large-scale community dynamics with nearly-complete information about what expertise individuals bring into the community and what they contribute.

With these data in hand, we can now ask how community-wide patterns of growth and decline are affected by the expertise, activities, and social interactions of individual community members. When players are born into a new community, they carry with them experiences from past lifetimes. Thus, each new playthrough of OHOL is akin to a migration. As in real-world technological development, these migrations have the potential to transform communities by bringing an infusion of ideas and skills [13, 26]. In the next two sections, we examine reciprocal interactions between individuals and communities—namely, how individuals adapt to new communities into which they are placed, and how communities are shaped by individual contributions.

2.2 Individuals adapt to match close peers

We characterized what jobs players performed in each lifetime by projecting their activities onto a lower-dimensional "jobspace" using non-negative matrix factorization (NMF; [27]). In brief, we measured a contingency matrix $\mathbf{A}$ of m characters $\times$ n objects, where each cell denotes the number of times each character interacted with each object. This matrix was factored into an $m \times k$ matrix of job embeddings ($\mathbf{J}$) that describes each character's activities, and an $k \times n$ matrix of object embeddings ($\mathbf{O}$) that captures

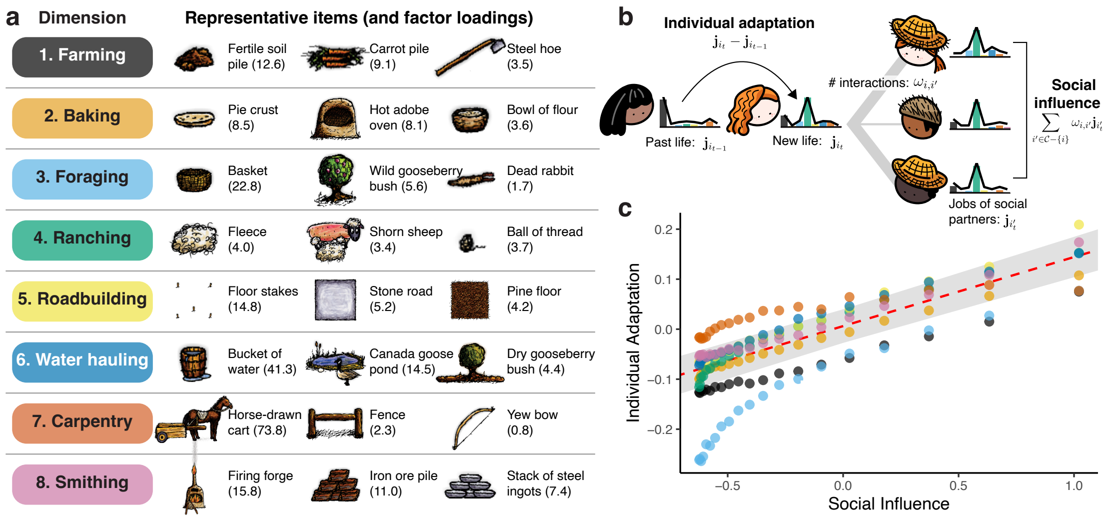


Fig. 2 Predicting individual adaptation. a, Each character's activities were projected onto a lower-dimensional "job-space" using non-negative matrix factorization. Each row denotes a dimension or "job" in the NMF solution, and entries show representative objects with high factor loadings on each job (see also Table S1 for all top loadings and Fig. S3 for more details on rank selection). **b,** Schematic of social adaptation: We predicted how each player i's current job embedding ($\mathbf{j}_{i_t}$) changed from the most recent lifetime that they spent in a different community (individual adaptation: $\mathbf{j}_{i_t} - \mathbf{j}_{i_{t-1}}$) based on a weighted average of the job embeddings of that player's new peers (social influence: $\sum_{i'} \omega_{i,i'} \cdot \mathbf{j}_{i'_{t-1}}$, where $\omega_{i,i'}$ denotes the number of social interactions between i and i' in the current lifetime t). **c,** Social adaptation results: Each point shows average adaptation scores for each job dimension (marked by color) computed along 20 quantiles of social influence values. The dotted red line shows the fixed effect of the model predictions, illustrating the relationship between social influence and individual adaptation across all jobs (shaded area denotes 95% CI).

relationships between objects ($\mathbf{A} \approx \mathbf{JO}$). Throughout the paper, we use an NMF solution with $k = 8$ job dimensions, which balances predictive accuracy and interpretability (see Fig. S3a for validation and rank selection). Inspecting object embeddings reveals that the dimensions recovered using this method align with intuitive "jobs" such as farming, foraging, and baking (Fig. 2a). Further, players tend to take on similar jobs in different playthroughs (Fig. S3b), which suggests that job embeddings capture meaningful patterns of expertise that players contribute to different communities in which they are placed.

Using this representation of players' jobs, we examined how individuals adapt to the communities in which they are placed. Specifically, individuals may *conform* to the community—taking on jobs that are more similar to the people with whom they interact—or *differentiate* themselves by taking on different jobs. Our dataset also offers a unique opportunity to determine whether individuals adapt to the activities of the community at large, or focus on adapting to their closest peers. We tested these hypotheses by predicting how players' jobs change from one lifetime to the next, based on the jobs performed by members of their new community (Fig. 2b). Specifically, we constructed a linear mixed-effects model that predicts how much each player i's job embedding changed in lifetime t (individual adaptation: $\mathbf{j}_{i_t} - \mathbf{j}_{i_{t-1}}$) based on a weighted average of the jobs performed by each peer i' in their new community (social influence: $\sum_{i'} \omega_{i,i'} \cdot \mathbf{j}_{i'_t}$, where $\omega_{i,i'}$ is the number of social interactions between i and i' in the current lifetime). We modeled adaptation along each job dimension as a separate observation with separate random slopes, and treated players and communities as random effects. Each player's total playtime and total number of social interactions were included as additional fixed effects. Figure 2c shows a clear positive trend, where individuals adapt to become more similar to their social partners across all job dimensions ($\beta_{\text{influence}} = 0.14, SE = 0.006, t = 24.0, p < .001$); this result is consistent with a conformity account. Individuals also adapted more when they had more interactions with other community members ($\beta_{\text{interactions}} = 0.01, SE = 0.0005, t = 29.0, p < .001$) and less experience playing the game ($\beta_{\text{gametime}} = -0.03, SE = 0.0005, z = -58.0, p < .001$). Our regression model outperforms control models (Fig. S3e–f) where we swapped the job embeddings of social partners with randomly sampled jobs from the entire population (random jobs model) or shuffled the links of the community's social network (random network

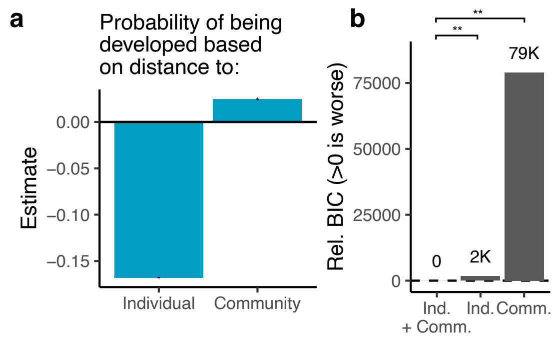


Fig. 3 **Tracing the impact of individual and community expertise. a,** Regression results. Each coefficient reflects the probability that a technology was newly developed by a community in a given epoch, based on the cosine distance between the technology's object embedding and the prior expertise of the individual who added the technology to the community's repertoire (Individual) and of the rest of the community (Community). Negative values that indicate that technologies that are farther from this expertise are less likely. Error bars, denoting standard errors, are not visible at this resolution. **b,** Model comparisons. BIC scores for mixed-effects logistic regression models that predict which technologies are added based on both individual and community expertise (Ind. + Comm.), or based on individual (Ind.) and community (Comm.) expertise separately. Y-axis values show relative BIC, or the difference of BICs between each model and the winning model (e.g., $\text{BIC}_I - \text{BIC}_{\text{I+C}}$); brackets denote pairwise comparisons against the Ind. + Comm. model performed using ANOVA; asterisks denote $p < .001$.

model). These controls suggest that individuals do not gravitate towards jobs that are more frequent across the entire population, and they also do not conform indiscriminately to the whole community.

Each community provides opportunities for individuals to acquire new expertise. Our results provide a more detailed view of how knowledge is transferred in communities: Rather than adapting to the community as a whole, individuals selectively adapt their activities based on the niche that they occupy within a community's social network. Next, we examine how communities are influenced by the expertise that individuals bring in.

2.3 Individual contributions to community growth

Next, we examined whether it is possible to predict what impact an individual will have on a community, based on fit between their expertise and that of the rest of the community. One deflationary hypothesis, consistent with folk tales of lone innovators, is that technological developments are solely driven by individual expertise, with no effect of community expertise. Alternatively, individuals may fall in step with the rest of the community, selectively developing technologies that are aligned with both individual and community expertise. Finally, individuals may contribute technologies that are within their unique domain of expertise; in this case, we would expect individuals to develop technologies that are closer to their expertise but *farther* from the community's expertise. This last hypothesis is consistent with classic economic and sociological proposals of the value of "structural holes" in the generation of new ideas [28, 29]; gaps in a community's knowledge create opportunities for individuals to contribute.

To test these hypotheses, we split each community's activities into 20-minute epochs, beginning at the birth of its founding member and ending with the death of its last descendant (rounded up to the closest 20-minute increment). In each epoch τ, we defined each community's *effective repertoire* $\mathcal{R}_\tau$ as the set of unique objects that community members used within that window. The key outcome measure for this analysis, *technological gains* $\mathcal{G}$, was defined as the set of objects added to a community's effective repertoire from the prior epoch ($\mathcal{G} = \mathcal{R}_\tau - \mathcal{R}_{\tau-1}$). Each object in $\mathcal{G}$ was paired with an undiscovered object that was selected to minimize differences in popularity and complexity within pairs (see Supplementary Methods for details on the matching procedure and Fig. S4 for validation). Our analysis included 6.7M object pairs. We then used a mixed-effects logistic regression to classify whether an object was discovered or not based on the cosine distance between its object embedding and (i) the prior expertise of the individual developer i who first interacted with the object (individual expertise: $\mathbf{j}_{i_{t-1}}$), (ii) the average expertise of all other community members i' in the set $\mathcal{C} - \{i\}$ (community expertise: $\sum_{i' \in \mathcal{C}-\{i\}} \mathbf{j}_{i'_{t-1}} / |\mathcal{C} - \{i'\}|$), adapting the modeling procedure in [30, 31].

Results are summarized in Figure 3. Consistent with the structural holes hypothesis, we find that communities are more likely to develop technologies that are close to the individual's expertise ($\beta_\text{I} =$

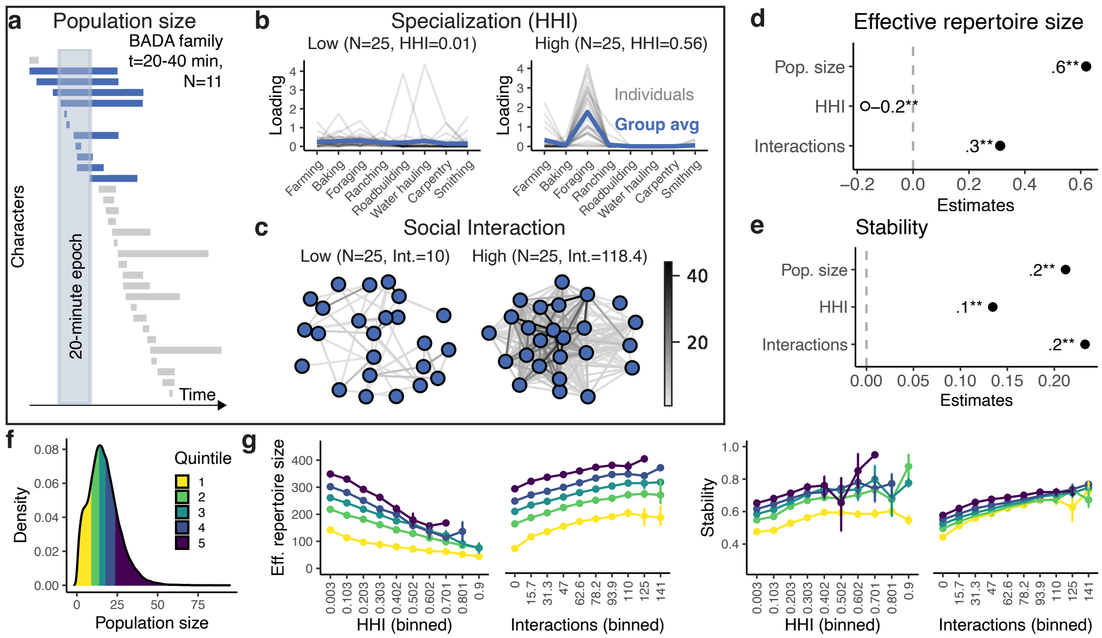


Fig. 4 Predicting community-wide patterns of growth. a–c, Predictors: Each community's history was split into 20-minute epochs. We identified the set of all community members alive in each epoch τ, $\mathcal{C}_\tau$, and derived several metrics: **a,** population size (i.e., the number of living community members, $|\mathcal{C}_\tau|$), **b,** specialization (i.e., the Herfindahl-Hirschman Index of the community's average job embedding, $\sum_{i' \in \mathcal{C}_\tau} \mathbf{j}_{i'}/|\mathcal{C}_\tau|$), and **c,** social interaction (i.e., the average number of social interactions per character, $\sum_{(i,i') \in [\mathcal{C}_\tau]^2} \omega_{i,i'}/|\mathcal{C}_\tau|$). **d,e** Regression results: Fixed effects for models predicting the **d,** size and **e,** stability of effective repertoires. Filled circles denote positive effects, and empty circles negative events; labels indicate coefficient values; asterisks denote significant predictors ($p < .001$). Error bars (denoting standard errors) are not visible at this resolution. **f,** Distribution of population sizes; fill colors denote quintiles. **g,** Effects of specialization and social interaction. Each point denotes the average size (left) and stability (right) of effective technological repertoires, based on binned values of population size (colors) and of specialization and social interaction (x-axes). Error bars denote bootstrapped 95% CI.

-0.16, SE $= 0.0006$, $z = -280$, $p < .001$) and far from the rest of the community's expertise ($\beta_{\mathrm{C}} = 0.02$, SE $= .0006$, $z = 41.5$, $p < .001$). The technologies that communities develop are best captured by a model that considers both individual and community expertise (Ind. + Comm. model BIC: 18474933), compared to a model that considers only individual expertise (Ind. model BIC: 18476638, ANOVA vs. Ind. + Comm.: $\chi^2(1) = 1720.9$, $p < .001$) or only community expertise (Comm. model BIC: 18553852, ANOVA vs. Ind. + Comm.: $\chi^2(1) = 78978$, $p < .001$). Thus, we find that both individual and community expertise make distinct contributions to technological development—individuals bring in new expertise into the communities where they are placed, and communities create opportunities where that expertise can be applied.

2.4 Effects of community composition on technological growth

Our results suggest that individuals face competing pressures to both align with and diverge from their community. They tend to match their activities with peers but contribute most in the areas where they differ. We now explore whether these two dynamics—aligning to form specialized communities and diverging to form diverse ones—have distinct impacts on technological development at the community level. Prior theoretical proposals suggest that specialized and diverse communities have unique strengths [32, 33]: More diverse communities are better able to develop new technologies quickly, but more specialized communities maintain existing technologies more efficiently. Thus, we hypothesize that more diverse communities may have larger, yet less stable technological repertoires.

To empirically test the relationship between technological development and community diversity, we used mixed-effects regressions to model the *size* ($|\mathcal{R}_\tau|$) and *stability* ($|\mathcal{R}_\tau \cap \mathcal{R}_{\tau-1}|/|\mathcal{R}_{\tau-1}|$) of effective repertoires over time as a function of three community characteristics: population size (the number of

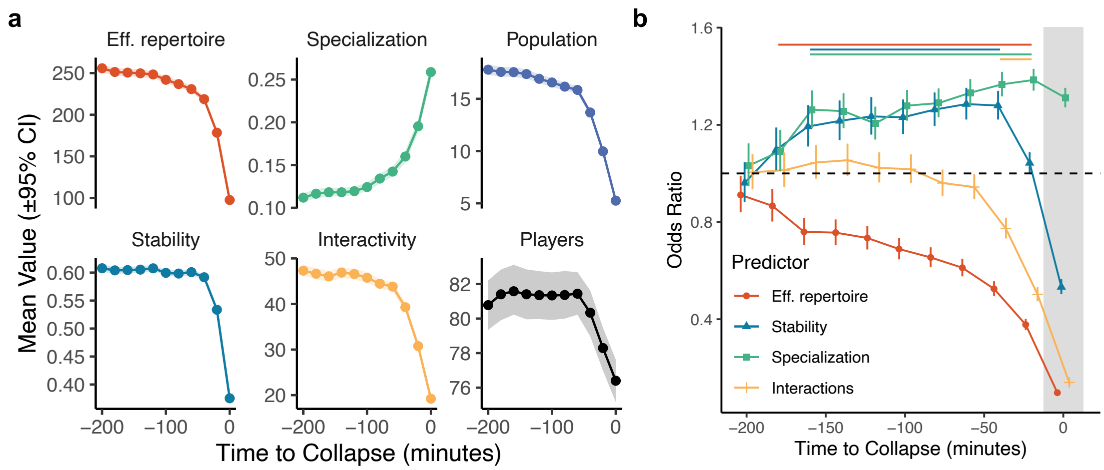


Fig. 5 Technological development as an early indicator of community collapse. We predicted the probability of community collapse based on community-level indicators of technological development (M1: Eff. repertoire, Stability) and community characteristics (M2: Specialization, Interactions), at offsets spanning 0–200 minutes prior to collapse. **a,** Average timecourse of community-level indicators, relative to the onset of collapse. Ribbons indicate 95% confidence intervals. **b,** Regression results. Each point shows model coefficients during a given epoch after statistically adjusting for the total number of players online throughout the server during that epoch; the gray shaded region highlights model estimates at the time of collapse. Error bars denote 95% Wald confidence intervals; horizontal lines indicate clusters where each predictor (marked by color) significantly predicted collapse after Bonferroni correction ($\alpha = 0.005$).

community members alive during epoch τ), average social interactions per capita, and specialization (measured as the Herfindahl-Hirschman Index, HHI of the community's average job embedding; Fig. S4a–c). Specialization is inversely related to diversity: Communities whose activities are concentrated in a single job have the highest specialization scores (high HHI), while diverse communities that perform many jobs have lower specialization scores (low HHI). Figure S5 shows the distribution of these community characteristics.

Regression results are shown in Figure 4d–e, and average repertoire sizes and stability split by community characteristics are shown in Figure 4f–g. Consistent with prior work showing that innovations and technological developments are concentrated in larger and more densely connected communities [10, 34], here we find that both population size and social interactions had positive effects on the size ($\beta_{\text{pop}} = 0.62, SE = 0.003, z = 228.16, p < .001$; $\beta_{\text{int}} = 0.31, SE = 0.003, z = 116.42, p < .001$) and stability ($\beta_{\text{pop}} = 0.21, SE = 0.002, z = 117, p < .001$; $\beta_{\text{int}} = 0.23, SE = 0.002, z = 129, p < .001$) of technological repertoires.

Of these community-level characteristics, only specialization has opposing effects on repertoire size and stability. More specialized communities had smaller ($\beta_{\text{spec}} = -0.17, SE = 0.003, z = -68.01, p < .001$) yet also more stable ($\beta_{\text{spec}} = 0.13, SE = 0.002, z = 70.9, p < .001$) technological repertoires. These results suggest that the best approach to building complex, lasting repertoires is not to specialize narrowly or to diversify to cover every activity. Instead, communities may need to divide labor to balance the demands of developing and maintaining technologies. To test this idea, we now examine whether imbalances in specialization serve as an early sign of collapse.

2.5 Specialization serves as an early warning sign of collapse

If communities are continually balancing two opposing processes—diversifying to develop new technologies, or specializing to maintain them—then we hypothesize that fluctuations in specialization may serve as an early warning sign of collapse. We define collapse as the timepoint when a community's population reaches 0. We then regressed community-level predictors at different temporal offsets to predict the time of collapse (see Supplementary Methods). Figure 5a shows the time course of these predictors, and Figure 5b shows the log-odds of collapse based on the value of these community-level predictors at different temporal offsets prior to collapse, after statistically adjusting for the total number of players online. Effective repertoire sizes consistently decreased prior to collapse and predicted the onset of collapse as early as 3 hrs (9 epochs) prior. These shrinking technological repertoires are accompanied by changes in social organization: Specialization consistently increases prior to collapse and predicts the onset of collapse as early as 2hrs 40 min (8 epochs) prior. Because each character can survive for up to one hour, these indicators signal collapse several generations before populations die out. The precise timing of these collapses does not coincide with periods of the day when players are less active, which suggests that collapses do not merely occur because there are no other players available to continue the lineage (Fig. S6). Our results suggest that technological development and diverse expertise are not merely symptomatic of a thriving community, but are essential to and predictive of its continued existence.

3 Conclusions

Influential theories from cultural evolution, economic geography, and urban science [1, 3, 8–10, 12, 35, 36] propose that placing many people with diverse expertise in close proximity creates opportunities for people to develop new technologies by learning from one another and combining their expertise. However, it is surprisingly challenging to test this proposal directly using existing methods, which afford either a coarse, bird's-eye view of real-world communities or a fine-grained, yet limited view of individual behavior within simplified experiments. Here, we leveraged two unique strengths of our dataset to examine why some communities build thriving technological repertoires while others stagnate. First, we provide a detailed view of how communities create opportunities for individuals to acquire expertise and contribute new technologies. Rather than falling in step with the community as a whole, we find that individuals selectively adapt their activities to match their closest peers, and they contribute technologies in areas where their expertise diverges from the rest of the community. Second, we took a broader view and examined how aligning and diverging from the community affect technological growth and survival at the community level. We find that specialization has opposing effects on the size and stability of technological repertoires. These results suggest that—rather than homing in on a single activity, or diversifying as much as possible—communities instead have to navigate a delicate balance between developing and maintaining technologies. Consistent with this idea, imbalances in specialization serve as an early warning sign of population collapse.

Together, these results provide an integrated view of technological development across levels of social organization. Much as orchestras create music by combining the distinct sounds of many instruments, communities develop technologies through the collaboration of individuals with distinct jobs, expertise, and positions in the social network. More broadly, our approach provides a testbed for theories of large-scale social phenomena that would be difficult to measure directly in real-world communities or emulate in experimental studies (Supplementary Discussion). This work provides a foundation to better understand how communities create conditions where individuals can thrive, and how individuals leave a distinct mark on the communities to which they belong.

4 Methods

Ethical approval declaration: The game's developer, Jason Rohrer, has made all data publicly available at http://publicdata.onehouronelife.com. These data were not collected for the purpose of research, and were provided without identifiable information by someone without any role in this research study except providing. Thus, the IRB at Harvard University and Princeton University determined that this work does not constitute human subjects research.

Sample selection: We analyzed game data from November 2019 to June 2020. Our raw data contain N = 755,029 characters, comprising N = 27,358 communities. Attrition presents a major challenge in analyzing this dataset. After excluding characters who disconnected early or who did not interact with the map, or families with < 3 eligible members, our final sample includes N = 2,700 communities composed of N = 428,255 avatars, operated by N = 22,011 players. Further details on sample selection can be found in Supplementary Methods.

Metrics: The raw data contains a demographic record of all births and deaths in the dataset ("life logs"), a log of every event where a player picked up, transformed, or dropped an object on the floor ("map logs"), and game data specifying each object's properties and how objects are transformed when they are combined or used (release v.341, retrieved from https://github.com/jasonrohrer/OneLifeData7 on June 17, 2020). Below, we describe how key metrics were derived from these data sources. Further details on these metrics—as well as alternative metrics used in control analyses—can be found in Supplementary Methods.

- *Technological tree*: We reconstructed the game's technological tree by searching recursively through transition files. We began the search with 80 naturally-occurring objects and searched for transition files that produce new objects when players combine two naturally-occurring objects or interact with a naturally-occurring object with their empty hands (e.g., plucking a branch off a tree). At each recursive step, we searched for transitions that produced new objects after interacting with the most recently-discovered objects, combining them, or combining one recently-discovered object with an object discovered in a previous recursive step, until no more technologies could be added to the technological tree.
- *Technological depth*: The number of recursive steps needed to search through the game's technological tree to develop a given object. Intuitively, naturally-occurring objects have a depth of 0, their immediate products have a depth of 1, and each subsequent object has a depth of $\max(d_I) + 1$, where d_I are the depths of the ingredients needed to produce the object.
- *Communities*: Family lineages. Families were reconstructed from life logs by creating a directed graph that connects each character to its parents until the graph terminates at the lineage's founding member.
- *Jobs*: We created a contingency table of m characters $\times$ n objects—where each cell denotes the number of times each character interacted with each object—and projected it onto a lower-dimensional space using non-negative matrix factorization [27]. This method picks out an intuitive space of "jobs" that capture players' typical activities within the game, including farming, foraging, baking, and water-hauling.
- *Expertise*: The job embedding of the character that a player operated in their most recent past lifetime; if players were born to the same lineage in multiple consecutive lifetimes, we instead took their most recent job in a different lineage.
- *Interaction weights*: We detected events in the map logs where one avatar picked up or modified an object that another avatar had placed on the ground. By detecting all such events in our dataset, we generated a weighted graph where each node corresponds to a character and the weight of the edge $\omega_{i,i'}$, corresponds to the number of interactions between characters i and i'.
- *Individual adaptation*: The difference between a player i's current job at lifetime t and their expertise ($\mathbf{j}_{i_t} - \mathbf{j}_{i_{t-1}}$).
- *Social influence*: A weighted average of the jobs performed by each of player i's peers in community $\mathcal{C}$ at lifetime t, weighted by $\omega_{i,i'}$, the number of interactions between them ($\sum_{i' \in \mathcal{C}-\{i\}} \omega_{i,i'} \mathbf{j}_{i'_t}$).
- *Epochs*: 20-minute windows in a community's history, beginning with the birth of the community's founding member and ending with the death of its last descendant (rounded up to the nearest 20-minute increment). The community characteristics described below—population size, specialization, interactions, repertoire size, and stability—were defined based on the activities of family members during each epoch.
- *Contemporaries*: The set of family members alive during an epoch τ, $\mathcal{C}_\tau$. *Population size* was defined as the size of this set, $|\mathcal{C}_\tau|$.
- *Specialization*: The Herfindahl-Hirschman Index (HHI) of the average job embeddings of all contemporaries,
- *Interactions*: The average number of social interactions between contemporaries, i.e., $\sum_{(i,i') \in \mathcal{C}_\tau^2} \omega_{i,i'} / |\mathcal{C}_\tau|$.
- *Effective technological repertoires*: The set of unique items that community members interacted with in epoch τ, $\mathcal{R}_\tau$. *Effective repertoire size* is the size of this set, $|\mathcal{R}_\tau|$.
- *Stability*: The proportion of objects present in a community's effective technological repertoire at epoch τ that were retained from the previous epoch: $(\mathcal{R}_\tau \cap \mathcal{R}_{\tau-1})/|\mathcal{R}_{\tau-1}|$. We excluded each community's first epoch from analysis, as stability is undefined where $|\mathcal{R}_{\tau-1}| = 0$.
- *Newly-developed technologies*: Items present in a community's effective repertoire at epoch τ that were not present in the previous epoch ($\mathcal{D} = \mathcal{R}_\tau - \mathcal{R}_{\tau-1}$).
- *Undeveloped matches*: Each new development in the set $\mathcal{D}$ was paired with an *undeveloped* match. We first sampled $|\mathcal{D}|$ objects that were not in the community's repertoire, where the probability p_o of sampling each object o was proportional to the total number of community repertoires that contain that object. This method prioritizes technologies that are popular across all communities, but that were not developed by the specific community being analyzed. We then assigned each developed object

to an undeveloped match, minimizing the differences in technological depths between each pair using a standard algorithm for linear sum assignment [37].

- *Distance to expertise*: Cosine distance between the job embedding for a newly-developed technology x (or its undiscovered match) and the expertise of the character i who first introduced it to the community repertoire $1 - S_C(\mathbf{o}_x, \mathbf{j}_{i_{t-1}})$.

Statistical analyses: Further details on the analyses described below can be found in Supplementary Methods.

- *Measuring how individuals adapt to their communities:* We used a mixed-effects linear regression to predict individual adaptation—i.e., how players' job embeddings changed from one lifetime to the next—as a function of social influence. We modeled adaptation along each job dimension as a separate observation. The model also included each player's total playtime and the total number of social interactions they had as additional fixed effects, as well as random effects of player, community, and job dimension.
- *Tracing individual contributions to technological development:* Each technology that was added to a community repertoire at epoch τ was paired with an undeveloped match using the method described in *Metrics*, above. We used a mixed-effects logistic regression to predict whether a given object was developed or not based on the proximity between that object's job embedding and the expertise of the character who added the object to the community repertoire (individual expertise), as well as to the average expertise of all other contemporaries (community expertise).
- *Predicting community-wide growth:* We predicted each community's effective repertoire size and stability at a given epoch using mixed-effects regressions with fixed effects of population size, interactions, and specialization, and community as a random intercept. To account for temporal autocorrelations in the data, both models included an AR(1) error structure. Effective repertoire size was fitted to a linear regression, and stability was fitted to a binomial regression for proportional outcomes.
- *Predicting community collapse:* We defined collapse as the timepoint when a community population reaches 0. We then used a series of logistic mixed-effects regression models to predict the time of collapse, where one class of models used community-level predictors (i.e., interactivity, and specialization) and the other used community-level outcomes (i.e., effective repertoire and stability). Each model was computed at different temporal offsets, with predictors lagged from 0–10 epochs, where each included total number of online players as a control (see Fig. S6) and family as a random intercept. Figure 5 visualizes the coefficients and the solid horizontal lines correspond to significant clusters that survived multiple test correction ($\alpha = .005$).

Supplementary information. Supplementary Methods, Supplementary Analyses, Supplementary Discussion, Supplementary Figures 1–6, Supplementary Table 1, and Supplementary References.

Acknowledgements. We are grateful to the developer of One Hour One Life, Jason Rohrer, and to the community of players who have spent years exploring the game's data, including the developers of ohol-family-trees (github.com/JustinLove/ohol-family-trees) and onetech (github.com/Kazetsukai/onetech). We thank Grace W. Deng for assistance in reconstructing the technological tree and family lineages. The computations in this paper were run on the FASRC Cannon cluster (supported by the FAS Division of Science Research Computing Group at Harvard University), the MPIB cluster, the MPI BioCyb cluster, the Uni Tübingen ML cloud cluster, and the PNI compute cluster at Princeton University. This research was supported by the National Institute of Mental Health (award number K00MH125856 to N.V). C.M.W. is supported by the German Federal Ministry of Education and Research (BMBF): Tübingen AI Center, FKZ: 01IS18039A and funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy–EXC2064/1–390727645.

References

[1] Henrich, J. *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter* (Princeton University Press, 2016).

[2] Richerson, P. J. & Boyd, R. *Not by genes alone: How culture transformed human evolution* (University of Chicago Press, 2008).

[3] Muthukrishna, M. & Henrich, J. Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150192 (2016).

[4] Momennejad, I. Collective minds: social network topology shapes collective cognition. *Philosophical Transactions of the Royal Society B* **377**, 20200315 (2022).

[5] Wu, C. M., Dale, R. & Hawkins, R. D. Group coordination catalyzes individual and cultural intelligence. *Open Mind* (in press).

[6] Glaeser, E. L., Kallal, H. D., Scheinkman, J. A. & Shleifer, A. Growth in cities. *Journal of Political Economy* **100**, 1126–1152 (1992).

[7] Feldman, M. P. & Audretsch, D. B. Innovation in cities: Science-based diversity, specialization and localized competition. *European Economic Review* **43**, 409–429 (1999).

[8] Jacobs, J. *The Economy of Cities* (New York: Vintage Books, 1970).

[9] Bettencourt, L. M., Lobo, J., Helbing, D., Kühnert, C. & West, G. B. Growth, innovation, scaling, and the pace of life in cities. *Proceedings of the National Academy of Sciences* **104**, 7301–7306 (2007).

[10] Balland, P.-A. *et al.* Complex economic activities concentrate in large cities. *Nature human behaviour* **4**, 248–254 (2020).

[11] Vélez, N., Christian, B., Hardy, M., Thompson, B. D. & Griffiths, T. L. How do humans overcome individual computational limitations by working together? *Cognitive Science* **47**, e13232 (2023).

[12] Powell, A., Shennan, S. & Thomas, M. G. Late pleistocene demography and the appearance of modern human behavior. *Science* **324**, 1298–1301 (2009).

[13] Koch, P., Stojkoski, V. & Hidalgo, C. A. The role of immigrants, emigrants and locals in the historical formation of european knowledge agglomerations. *Regional Studies* 1–15 (2023).

[14] Jara-Figueroa, C., Jun, B., Glaeser, E. L. & Hidalgo, C. A. The role of industry-specific, occupation-specific, and location-specific knowledge in the growth and survival of new firms. *Proceedings of the National Academy of Sciences* **115**, 12646–12653 (2018).

[15] Derex, M. & Boyd, R. Partial connectivity increases cultural accumulation within groups. *Proceedings of the National Academy of Sciences* **113**, 2982–2987 (2016).

[16] Creanza, N., Kolodny, O. & Feldman, M. W. Greater than the sum of its parts? modelling population contact and interaction of cultural repertoires. *Journal of The Royal Society Interface* **14**, 20170171 (2017).

[17] Krugman, P. Increasing returns and economic geography. *Journal of Political Economy* **99**, 483–499 (1991).

[18] Hidalgo, C. A., Klinger, B., Barabási, A.-L. & Hausmann, R. The product space conditions the development of nations. *Science* **317**, 482–487 (2007).

[19] Neffke, F., Henning, M. & Boschma, R. How do regions diversify over time? industry relatedness and the development of new growth paths in regions. *Economic Geography* **87**, 237–265 (2011).

[20] Mesoudi, A. Cultural evolution: a review of theory, findings and controversies. *Evolutionary Biology* **43**, 481–497 (2016).

[21] Derex, M. & Mesoudi, A. Cumulative cultural evolution within evolving population structures. *Trends in Cognitive Sciences* **24**, 654–667 (2020).

[22] Argote, L. *Organizational learning: Creating, retaining and transferring knowledge* (Springer Science & Business Media, 2012).

[23] Kogut, B. & Zander, U. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science* **3**, 383–397 (1992).

[24] Allen, K. *et al.* Using games to understand the mind. *Nature Human Behaviour* 1–9 (2024).

[25] Long, B., Simson, J., Buxó-Lugo, A., Watson, D. G. & Mehr, S. A. How games can make behavioural science better. *Nature* **613**, 433–436 (2023).

[26] Trippl, M. & Maier, G. *Knowledge spillover agents and regional development* (Springer, 2011).

[27] Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791 (1999).

[28] Burt, R. S. Structural holes and good ideas. *American Journal of Sociology* **110**, 349–399 (2004).

[29] Granovetter, M. S. The strength of weak ties. *American Journal of Sociology* **78**, 1360–1380 (1973).

[30] Schulz, E. *et al.* Structured, uncertainty-driven exploration in real-world consumer choice. *Proceedings of the National Academy of Sciences* **116**, 13903–13908 (2019).

[31] Brändle, F., Stocks, L. J., Tenenbaum, J. B., Gershman, S. J. & Schulz, E. Empowerment contributes to exploration behaviour in a creative video game. *Nature Human Behaviour* **7**, 1481–1489 (2023).

[32] Ben-Oren, Y., Kolodny, O. & Creanza, N. Cultural specialization as a double-edged sword: division into specialized guilds might promote cultural complexity at the cost of higher susceptibility to cultural loss. *Philosophical Transactions of the Royal Society B* **378**, 20210418 (2023).

[33] March, J. G. Exploration and exploitation in organizational learning. *Organization Science* **2**, 71–87 (1991).

[34] Henderson, J. V. The effects of urban concentration on economic growth (2000).

[35] Henrich, J. Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses—the tasmanian case. *American Antiquity* **69**, 197–214 (2004).

[36] Audretsch, D. B. & Feldman, M. P. Knowledge spillovers and the geography of innovation. *Handbook of Regional and Urban Economics* **4**, 2713–2739 (2004).

[37] Crouse, D. F. On implementing 2d rectangular assignment algorithms. *IEEE Transactions on Aerospace and Electronic Systems* **52**, 1679–1696 (2016).

Supplementary materials: The rise and fall of technological development in virtual communities

Natalia Vélez[1], Charley M. Wu[2†], Samuel J. Gershman[3], Eric Schulz[4†]

[1*]Department of Psychology, Princeton University, Princeton, NJ, USA.
[2]University of Tübingen, Tübingen, Germany.
[3]Department of Psychology, Harvard University, Cambridge, MA, USA.
[4]Helmholtz Institute for Human-Centered AI, Munich, Germany.

[†]These authors contributed equally to this work.

Supplementary Information Guide

1 Supplementary Methods

1.1 Source data

We analyzed player behavior in *One Hour One Life*, a multiplayer online game where players can build complex societies from scratch. The game's developer, Jason Rohrer, has made a slice of game data publicly available at http://publicdata.onehouronelife.com. The game's public logs have been continually updated in two-week slices from March 2017—when the game was first launched—to the present day. We used two primary sources for our analyses—life logs and map logs—and augmented our analyses using the game's data files (available at https://github.com/jasonrohrer/OneLifeData7).

Life logs contain a record of every birth and death that occurred during a game. Each row in the map logs corresponds to a single birth or death. Each birth entry contains a record of when and where the character was born, the character's gender, which player operated that character, and who that character's parent was. Each death entry contains a record of when and where the character died, the character's age at the time of death, the cause of death (e.g., "hunger" or "murder") and, if murdered, the identity of the character's murderer. Players were tagged using an anonymized hash; this enables us to track players' activities across multiple playthroughs without collecting identifying information.

Map logs contain a record of every state change in the game's map. The map state changes when (1) an object naturally changes or decays (e.g., a plant blooms), or when a character causes a change by (2) placing an object on an empty spot on the map, (3) picking up an object, or (4) using an object in their hands to transform an object on the floor (e.g., when a player uses an axe to cut wood into kindling). Each entry in the map log records the time and location of the change and the new state of that map location. Starting in November 2019, the map logs also record which character caused the change.

Game data files dictate the structure of the game. These files specify (1) the properties of all objects available within the game (e.g., spawn probabilities of naturally-occurring resources, caloric values of edible objects, heat output of fires), (2) transition rules that dictate how objects can be transformed and used, and (3) the graphics, sounds, and textures players see. These files were used to reconstruct the game's technological tree.

1.2 Sample selection

We curated a dataset of player behavior between November 2019 and June 2020. We began data collection after a software update that caused the map logs to start tracking which character caused each change in the map's state (v.280); this change to the map log format enables us to trace a single character's interactions with the game map over an entire lifetime. Our raw data contain the records of N = 755,029 unique characters in the life logs, comprising N = 27,358 family lineages.

Attrition presents a major challenge in analyzing this dataset. For example, 19,160 (70.03%) of family lineages in our raw data die out within the first 5 minutes, as the founding member immediately disconnects or restarts the game. Thus, we applied several inclusion criteria to focus our analyses on characters and communities that remained in the game long enough to develop new technologies.

First, we only included characters who were present in both the life logs and map logs. This criterion excluded 113,596 characters who appeared only in the life logs (likely because they did not cause any changes to the environment that would appear in the map logs) and 47,152 characters who were only present in the map logs (likely due to data loss in the game's original life logs), leaving N = 641,433 characters.

Next, we excluded characters whose players disconnected while the character was still in infancy. Characters age by one year in game time for every minute that passes in real time. As they age, characters also pass through discrete life stages that affect their appearance and their ability to interact with the environment: Characters are born as helpless infants, grow into children, mature into adults, then finally die of old age. During infancy (defined as the first three minutes of a playthrough), players cannot interact with objects on their own—their character instead has to be fed, clothed, and carried by other characters who have already reached adulthood. Players also have access to a special command during infancy that enables them to disconnect and restart the game in a new family. Thus, players who disconnected from game sessions or whose characters died in infancy left the game session before being able to contribute to the communities they were placed in. We filtered out characters who died before the age of four or whose player disconnected from the game session, leaving N = 453,366 characters.

Finally, we excluded communities who had too-few contributing members. We first defined communities by tracing family lineages (see "Defining communities" below), and then excluded communities that had fewer than 3 members remaining after applying the exclusions above. This criterion excluded N = 25,111 characters and N = 24,153 communities. Ultimately, our sample includes N = 2,700 communities composed of N = 428,255 characters, operated by N = 22,011 players.

1.3 Defining and validating communities

In order to examine what factors drive technological development within communities, we first need to define what we mean by a "community." The most expedient way to define communities is by family lineages. Each time a player starts a new game, their character is born as the infant of another player whose character has reached adulthood. (Note that characters reproduce through parthenogenesis; they have mothers, but no fathers.) We reconstructed family lineages from life log data by building a directed graph that connects each character to their parent, their parent to their grandparent, and so on. Using this method, each lineage can eventually be traced back to a single female founder (termed "Eves" in the game) who spawns in an unoccupied area of the map as a fully-formed adult. In the main text, when we mention "communities", we refer to these family lineages.

However, there are many aspects of community functioning that family lineages may lack. Intuitively, a community should be composed of individuals who live in close proximity and who share the fruits of their work with one another. Two characters who are descended from the same lineage are not guaranteed to belong to the same "community" in this intuitive sense. Family lineages might lack these characteristics if families tend to disperse over time, or if characters tend to live within multi-family settlements. Thus, we validated our operationalization of communities using two measures. First, we measured the distance between each character's place of birth and the place of birth of all other characters who were alive during that character's lifetime. We then identified each of these other characters as kin or non-kin—based on whether they belong to the same family lineage or not—and measured the median distance to members of each group (Fig. S1a–b). Our results suggest that characters live significantly closer to kin than to non-kin (paired t-test between median distances to kin and non-kin by character: $t(425602) = 1143, p < .001$). Characters are born at a median distance of 53.1 tiles away from kin and 1423.8 tiles away from non-kin. If we assume that characters move at the game's default walking speed (3.75 tiles/minute), then they would have to walk for 0.24 minutes to reach kin and 6.33 minutes—over a tenth of a lifetime—to reach non-kin. Second, we measured how often characters interact with members of the same family lineage (see *Defining "social interactions"*, below, for implementation details). Overall, of the 9,384,198 social interactions detected in our dataset, 71.09% occurred between kin. Put together, family lineages capture meaningful spatial and functional patterns in our data: On average, characters live closer to kin than to non-kin and interact more often with them.

1.4 Detecting social interactions

We detected social interactions by detecting instances where one character picked up or modified an object that another character had placed on the ground. A typical interaction is illustrated in Figure S1c: Alice first places a yew branch on the ground at coordinates $(x, y) = (1, 1)$. Barbara then uses a stone hatchet to chop the branch into kindling, Chris picks up the kindling to make a fire, leaving that spot empty, and Diana puts down a basket on the empty spot. We then built a graph representation of interactions between players by drawing edges between characters who interacted with the same, non-empty spot on the map in two consecutive events. Figure S1d shows the resulting graph for this toy example. (Note that there is no edge connecting Chris and Diana, because the square was empty when David placed a basket on it. Therefore, no objects changed hands between the two.) By iterating over the map log data, we obtained a weighted, directed graph of N = 9,384,198 social interactions between 428,255 characters. We used the full interactions graph to measure the total number of interactions that occurred between kin (see *Defining "communities"*, above); Figure S1e shows all interactions with kin (blue nodes) and non-kin (red nodes) for one representative character in this graph (black node). In analyses in the main text, we split the graph into subgraphs for each lineage to measure how often members of the same community interacted with one another.

1.5 Reconstructing the game's technology tree

Within the game, players can craft new objects by combining an object in their character's hand (the "actor" object) with an object on the floor (the "target" object). One Hour One Life's crafting system is defined by thousands of "transition files" within the game's data, which specify how actors and targets change following an interaction. Each transition describes an event of the form:

$$(\text{initial actor}, \text{initial target}) \rightarrow (\text{final actor}, \text{final target})$$

For example, the final transition needed to produce a yew bow (Figure 1e, main text) can thus be described as:

$$(\text{rope}, \text{yew shaft}) \rightarrow ([\text{empty hands}], \text{yew bow})$$

Simply drawing a graph that connects objects that are involved in the same transitions would not be enough to reconstruct the game's technological tree, as transition files also describe object changes that happen autonomously (e.g., objects growing, cooling, drying, and decaying) and contain infinite loops. For example, the following transition dictates what happens when places two baskets on top of one another to make a stack:

$$(\text{basket}, \text{basket}) \rightarrow ([\text{empty hands}], \text{stack of 2 baskets})$$

And the following transition dictates what happens when a character removes a basket from the stack:

$$([\text{empty hands}], \text{stack of 2 baskets}) \rightarrow (\text{basket}, \text{basket})$$

If we simply connected these steps to reconstruct a recipe to produce a stack of baskets, the recipe would have infinite steps, as one could cycle through these two transitions endlessly. To isolate just the transitions needed to produce *new* objects, we searched through transition files using a recursive procedure that was adapted from an open source, fan-made crafting guide (https://onetech.info/). Figure S2a shows a schematic of the procedure. First, a *workspace* was seeded with the 80 naturally-occurring objects in the game (e.g., Figure 1e shows four such objects in the bottom level of the graph: milkweed plants, stones, big rocks, and yew trees). In this first recursive step, we exhaustively searched for all transitions involving (a) pairs of objects in the workspace, (b) empty hands interacting with single objects in the workspace, and (c) single objects in the workspace being combined with themselves, to test if there existed transitions involving these objects that produce new items. For example, running this procedure on the naturally-occurring objects in Figure 1e would uncover the following new transitions:

$$\begin{aligned}
([\text{empty hands}], \text{milkweed}) &\rightarrow (\text{milkweed stalk}, \text{milkweed debris}) \\
([\text{empty hands}], \text{yew tree}) &\rightarrow (\text{yew branch}, \text{yew tree}) \\
(\text{stone}, \text{big rock}) &\rightarrow (\text{sharp stone}, \text{big rock})
\end{aligned}$$

All previously-discovered objects would then be moved from the workspace to a *cache*, and all newly-discovered objects (in this example, milkweed stalks, milkweed debris, yew branches, and sharp stones) would be added to the workspace. In subsequent recursive steps, we would additionally search through transitions involving single objects in the cache and single objects in the workspace. This process continued until no more craftable objects were found (3798 objects total, with a maximum depth of 111 recursive steps).

We can thus measure technological depth as the number of recursive steps needed to discover an object using this procedure, and the number of ingredients needed to produce a technology as the total number of unique objects involved in the transitions needed to produce the object, starting from naturally-occurring primitives.

1.6 Defining and validating jobs

In order to characterize what jobs characters perform in each playthrough, we first created an $m \times n$ contingency matrix $\mathbf{A}$ where the $m = 428,255$ rows correspond to each unique character and the $n = 2,804$ rows correspond to each unique item (after filtering out items that were never used). Each element $a_{nm} \in \mathbf{A}$ describes the number of times character m interacted with item n. We then used non-negative matrix factorization (NMF)[1] to learn latent "job" embeddings that capture each character's activities as a low-dimensional vector.

NMF describes a method of approximating a matrix $\mathbf{A}$ as the product of two factorized matrices $\mathbf{J}$ and $\mathbf{O}$, such that $\mathbf{A} \approx \mathbf{JO}$. Importantly, a non-negative constraint is placed on $\mathbf{J}$ and $\mathbf{O}$ as a method of extracting sparser and more interpretable latent features. Unlike other dimensionality techniques such as PCA, each additional latent feature only provides additive components, making their contributions more interpretable. Each row of the matrix $\mathbf{J}$ can be interpreted as a low-dimensional embedding of each character in $\mathbf{A}$, while each column of the matrix $\mathbf{O}$ can be interpreted as a low-dimensional embedding of each item in $\mathbf{A}$. Characters that interacted with similar items will have similar row vectors in $\mathbf{J}$, while items that are interacted with by similar characters will have similar column vectors in $\mathbf{O}$. NMF was performed using `scikit-learn` [2] with Coordinate Descent[3] and initialized with random matrices.

An important parameter in the model is the number of latent dimensions k, which define the length of each character's job embedding in $\mathbf{J}$ and of each object's embedding in $\mathbf{O}$. Intuitively, we searched for a lower-dimensional representation of players' activities that balances accuracy and sparsity—i.e., that captures as much of the variation in player's activities as possible with the fewest dimensions. We computed NMF models with $k \in 2, 4, \ldots, 50$ and compared them on the basis of an interpretability score, inspired by [4], that combines measures of accuracy and sparsity:

$$\text{interpretability}(k) = \frac{\text{accuracy}(k) + \text{sparsity}(k)}{2} \tag{1}$$

Accuracy is defined as the ratio of the variance explained for a given number of dimensions k relative to the variance explained by the maximum number of dimensions under consideration ($k_{\max} = 50$):

$$\text{accuracy}(k) = \frac{\text{varExplained}(k)}{\max_{k' \in 2,4,\ldots,50} \text{varExplained}(k')} \tag{2}$$

Sparsity is defined as a simple inverse function of the number of dimensions:

$$\text{sparsity}(k) = 1 - \frac{k}{k_{\max}} \tag{3}$$

Figure S3a shows the results of this analysis; an NMF solution of $k = 8$ job dimensions achieves the highest interpretability score by balancing accuracy and sparsity.

1.7 Control models: Random jobs and random networks

In the social adaptation analysis, we compare the results of our main model to two controls. First, the *random jobs* control was created by replacing each community member's job embedding with job embeddings that are uniformly sampled from the whole population. Intuitively, these random job embeddings favor combinations of jobs that are more prevalent in the whole population. Thus, this is a particularly important control to understand whether individuals are specifically adapting to their peers. Second, the *random networks* control was generated by shuffling connections in the whole community's interaction graph. This method preserves the total number of social interactions observed throughout a family's history, but it obscures information about an individual's closest peers.

2 Supplementary Discussion

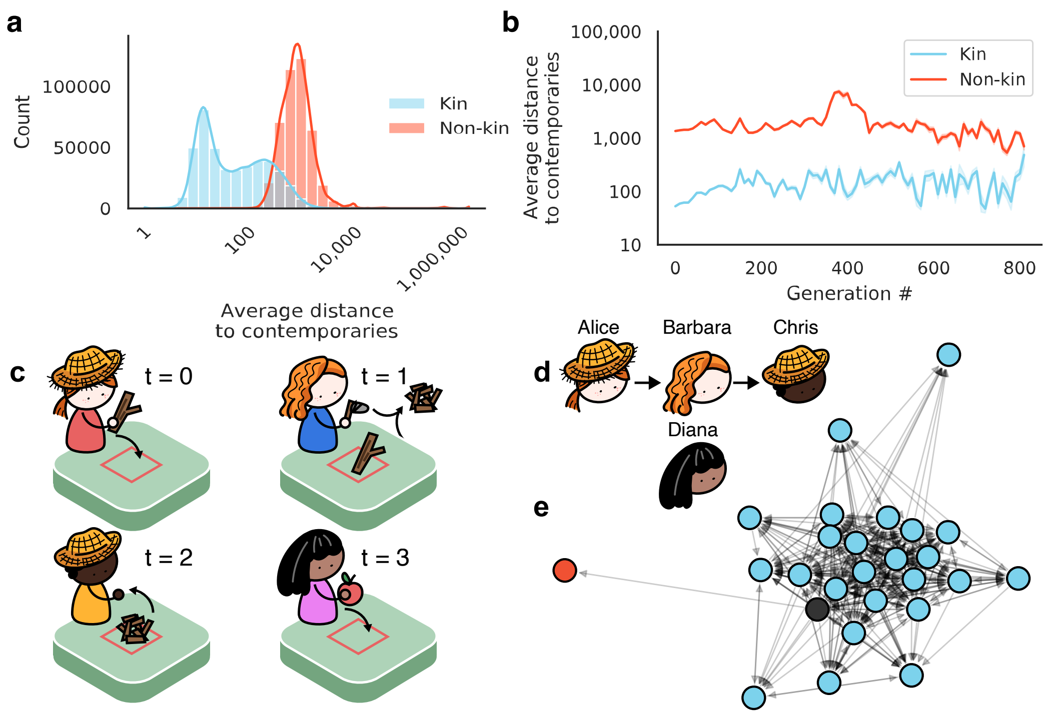
In the main text, we highlight the unique opportunities afforded by studying technological development in One Hour One Life. Here, we propose further opportunities to study large-scale social phenomena using this approach and highlight weaknesses that may be addressed by complementing this approach with experimental studies and real-world datasets.

The One Hour One Life dataset was not originally intended for research, and it therefore has many of the limitations and strengths common to any "found" digital dataset [5]. First, the dataset is inherently incomplete. Most notably, players can interact with one another by typing transient messages that appear in speech bubbles above the character's head or by leaving more durable notes on paper, signs, and gravestones. However, we cannot observe these social interactions in the map logs, and we have to infer social interactions indirectly based on how players interact with the environment. As a result, we do not know, for example, to what extent the patterns of specialization that we observe are tacit or explicitly coordinated [6], or to what extent players acquire expertise through instruction from a more experienced teacher [7–9]. These are fruitful avenues for experimental work.

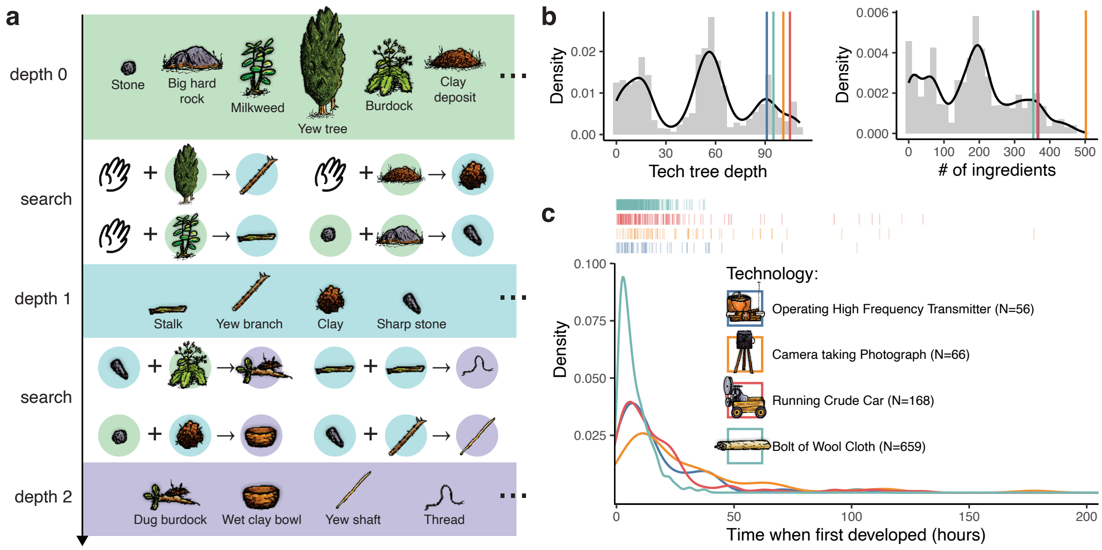
Second, while OHOL affords a view of technological development within a naturalistic environment, it is not quite natural; instead, the game was designed by a developer to encourage certain behaviors and discourage others. In the main text, we noted one way that this feature constrains our interpretation of the data. Namely, because everything that players can create is pre-specified, we cannot say whether a new technology constitutes an "innovation" [10]; instead, the primary problem faced by communities is gathering the resources, materials, and know-how needed to make it [11, 12]. This problem of naturalism is not unique to found datasets. Experimental studies also retain some naturalistic features while distorting or stripping away others. The key difference is that researchers have the agency to decide what features are relevant to include in an experimental study (for better or for worse [13]) and can assess the causal effects of experimental manipulations. Thus, it is important to ground the questions that we ask of this dataset with theories and empirical observations of real-world communities, and to augment this approach with targeted experimental studies.

By contrast, OHOL provides an important testbed for large-scale social phenomena that would be difficult to measure in real-world datasets or emulate in experimental environments. A key strength of this dataset is that it is large enough and collected over a long enough timescale for thousands of communities to blink in and out of existence. Like proverbial laboratories of democracy, these communities provide a quasi-experiment into what makes a community successful. In the present work, we focused on a particular metric of success: what conditions enable communities to develop technological repertoires. Yet this is not the only sense in which a community can be successful. Intuitively, we can also measure a community's success based on how resilient it is to external shocks [14], how efficiently it meets the basic needs of its members [15], how smoothly it manages internal conflict [16], or how faithfully it preserves shared natural resources for future generations [17]. Each of these metrics is a viable direction for future work using this approach.

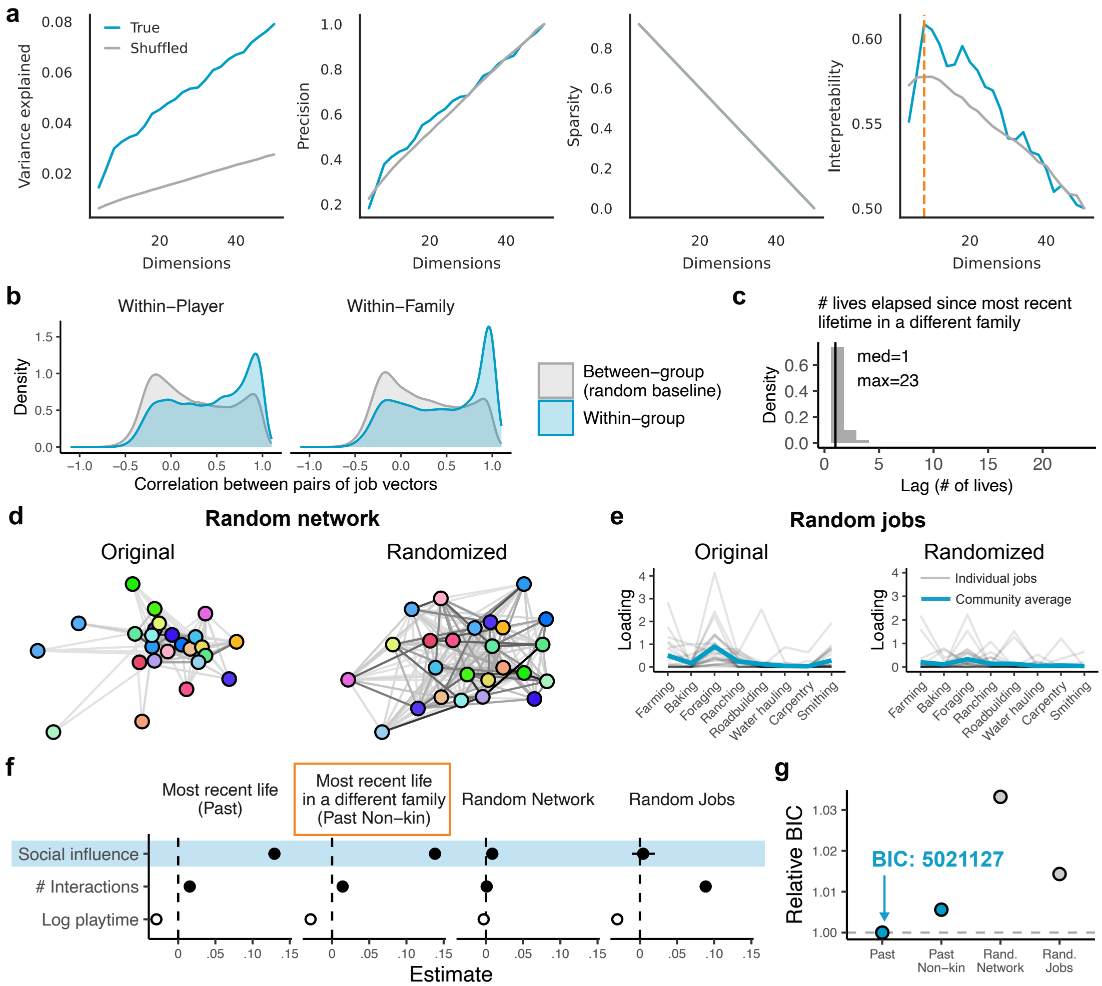
3 Supplementary Figures



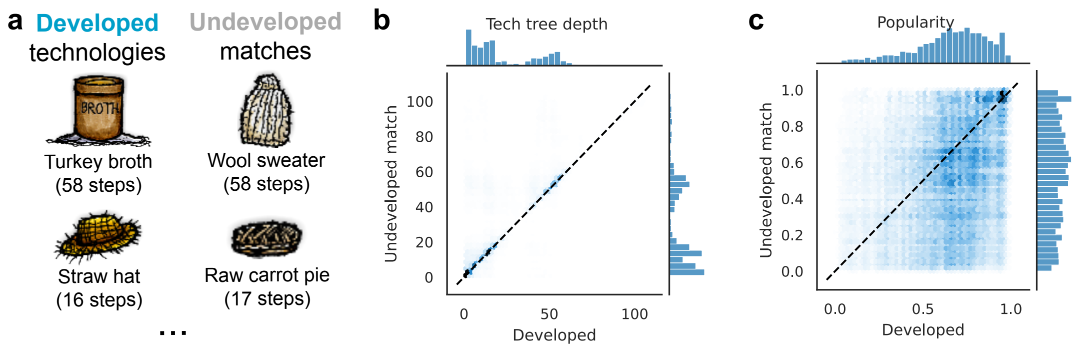
Supplementary Figure 1: **Defining communities and social interactions.** We defined communities as family lineages. We validated this operationalization by testing whether characters **a,b** are physically closer to and **c,d,e** share objects more often with kin, compared to non-kin. **a,** Distribution of average distances to kin and non-kin: We estimated how close characters live to each other by measuring the distance between each character's place of birth and the birthplace of all other living characters. Each datapoint in these distributions corresponds to one character. **b,** Average distance to kin and non-kin over time. We measured characters' average distance to kin and non-kin, based on the generation that there were born into. (By convention, the founding member of each family is born in generation 0, her immediate descendants are generation 1, and so on.) Ribbons denote bootstrapped 95% confidence intervals. **c,d** Schematic of social interactions measure: Two characters can be said to have interacted with one another if one character transforms or picks up an item that another character has placed on the game's map. **c,** Toy example, depicting four characters interacting with a patch of the map. **d,** Interaction graph of the toy example shown in panel c. Note that the last character is not connected to anyone, as they placed an apple on an empty patch of the map, rather than interacting with the objects that other characters had placed on that patch before. **e,** Full interaction graph for one representative character within our sample. The black note denotes the character highlighted in this analysis; blue nodes denote kin who interacted directly with this character, and red notes denote non-kin. Overall, the majority of this character's interaction partners are kin.



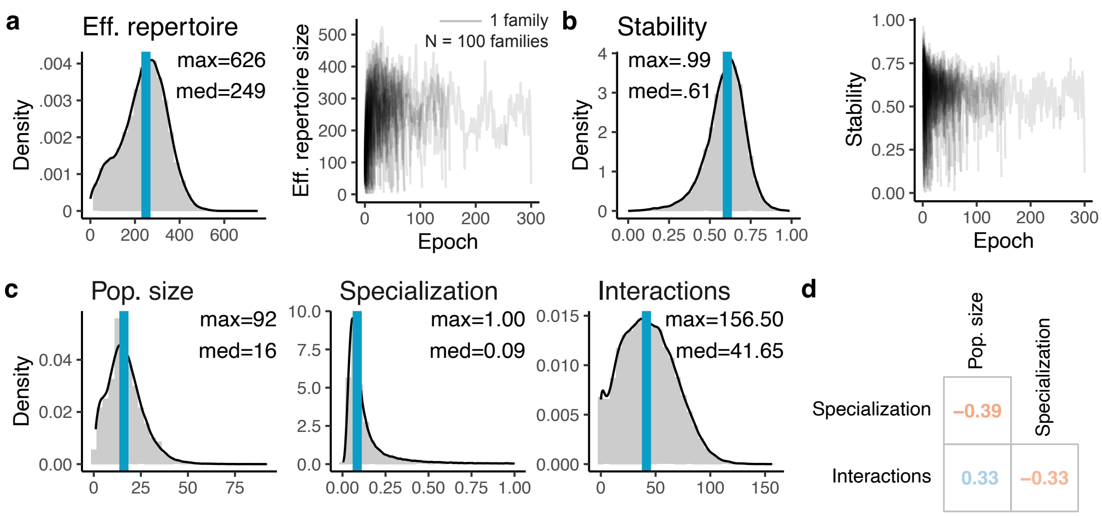
Supplementary Figure 2: **Reconstructing the game's technological tree. a,** Schematic of procedure used to recursively search through the game's technological tree. This procedure is described in more detail in Supplementary Methods. **b,** Overall distribution of tech tree depth (the number of recursive steps needed to discover an object) and number of ingredients (the number of unique objects needed to produce a technology from start to finish). Vertical lines highlight representative examples of advanced technologies: a high-frequency radio transmitter, a camera taking a photograph, a running car, and a bolt of wool cloth. **c,** Time elapsed when an advanced technology was first developed, relative to the community's founding. Each tick mark denotes one community that developed each item; sample sizes in the legend indicate the total number of communities who added each of these objects at least once to their technological repertoires.



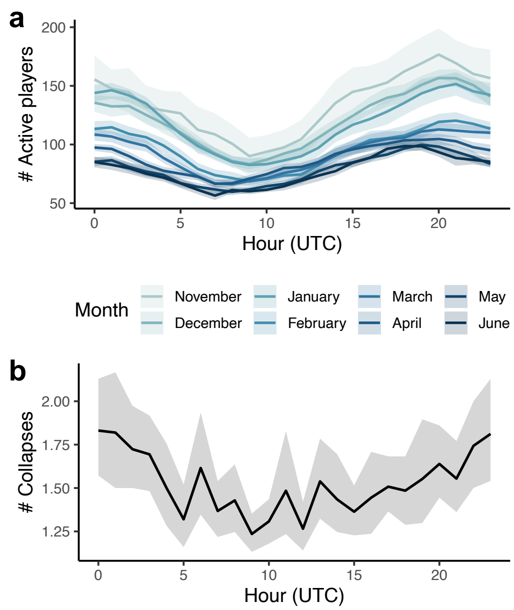
Supplementary Figure 3: **Validation of NMF analysis**. **a,** Rank selection. Variance explained, precision, sparsity, and interpretability scores for NMF solutions derived from the true activity matrix (blue line); by comparison, we have also included results for a matrix with shuffled rows (gray line). An NMF solution with $k = 8$ latent dimensions (orange dotted line) best balanced precision and sparsity. **b,** Consistency in jobs performed by each player (left) and community (right). The blue curves show the correlation between pairs of job embeddings sampled from the same player or from the same community, as compared to a random baseline where pairs of jobs are sampled uniformly from the whole dataset (gray). **c,** In the main text, we characterized players' expertise based on the jobs they performed in their most recent lifetime in a different community. Here, we show the distributions of lag between the player's current life and the lifetime used to compute their expertise, measured in number of lifetimes. In most cases, expertise was computed based on the most recent lifetime (lag 1). **d,e** Random network and random jobs control. **d,** Example interaction graph for a representative community, before (Original) and after (Randomized) shuffling the edges. **e,** Example job embeddings for a representative community, before (Original) and after (Randomized) replacing job embeddings by sampling uniformly from the whole dataset. **f,** Model coefficients for social adaptation analysis. The color of each dot denotes the sign; blue area highlights the effect of social influence estimated in different models; and the orange box highlights the model reported in the main text. **g,** BIC scores. Blue dots denote models that used the true job embeddings and network structure; gray dots denote random baselines.



Supplementary Figure 4: **Comparing developed and undeveloped technologies.** In the main text, we tested whether we can predict which specific technologies are added to a community's repertoire based on individual and community expertise. To address this question, we matched each technology that was developed to a plausible, yet undeveloped technology. **a,** Representative pairs of developed and undeveloped technologies. **b,c** Comparison between the **b,** technological depth and **c,** popularity of developed and undeveloped technologies. Technological depth was measured as the number of recursive steps needed to produce a technology (Fig. S2) and popularity was measured as the proportion of communities who added the technology to their repertoires at least once in the community's history. Note that our matching procedure only explicitly minimized differences in depth; nonetheless, we see that the undeveloped pairs are comparable in both depth and popularity.



Supplementary Figure 5: **Predicting the growth of communities. a,b** Effective repertoire size and stability. The plots on the left of each panel show the distribution of **a,** effective repertoire sizes and **b,** stability by community and epoch; the plots on the right show timecourses of effective repertoire size and stability for a representative subsample of N = 100 communities. **c,** Distribution of population sizes, specialization, and interactions. Vertical blue lines indicate median values. **d,** Correlations between population size, specialization, and frequency of interactions.



Supplementary Figure 6: **Predicting the onset of collapse**. In the main text, we predict when communities collapse based on indicators such as specialization and social interactions. One alternative, deflationary explanation for why communities collapse is that they may simply die out because there are not enough players online on the server to carry on the lineage. If this is the case, then we may expect to see collapses to peak at times of the day when players are least active. **a,** Average number of daily players per hour, split by each month in our dataset. **b,** Average number of collapses per hour. Shaded areas denote 95% CI. Contrary to this deflationary account, collapses do not peak at approximately 10 UTC, when the server is least active.

4 Supplementary Tables

Dimension	Object name	Loading
1) Farming	Fertile Soil Pile	12.6
	Deep Tilled Row	11.6
	Fertile Soil	10.1
	Shallow Tilled Row	9.7
	Carrot Row	9.2
	Carrot Pile	9.1
	Hardened Row	3.6
	Steel Hoe	3.5
	Milkweed Debris	3.4
	Dry Planted Carrots	3.3
2) Baking	Clay Bowl	26.6
	Clay Plate	16.3
	Stack of Clay Bowls	13.5
	Stack of Clay Plates	12.9
	Raw Pie Crust	8.5
	Hot Adobe Oven	8.1
	Bowl of Dough	5.8
	Bowl of Flour	3.6
	Sliced Bread	3.3
	Raw Mutton Pie	3.1
3) Foraging	Basket	22.8
	Wild Gooseberry Bush	5.6
	Sharp Stone	4.4
	Dead Rabbit	1.7
	Skinned Rabbit	1.5
	Milkweed Stalk	1.5
	Thread	1.4
	Clay Pit	1.3
	Snared Rabbit Family Hole	1.3
	Baited Snared Rabbit Family Hole	1.2
4) Ranching	Bowl of Gooseberries	64.1
	Domestic Gooseberry Bush	58.3
	Empty Domestic Gooseberry Bush	9.8
	Bowl of Gooseberries and Carrot	9.0
	Fleece	4.0
	Ball of Thread	3.7
	Shorn Domestic Sheep (with fleece)	3.4
	Shorn Domestic Sheep (no fleece)	3.3
	Gooseberry	3.2
	Fed Shorn Domestic Sheep	3.1
5) Roadbuilding	Stakes	19.4
	Floor Stakes	14.8
	Stone	13.8
	Hand Cart	6.2
	Stone Road	5.2
	White Pine Tree	4.2
	Pine Floor	4.2
	Auto-Orienting Wall Stakes	3.2
	Hand Cart with Tires	3.2
	Boards	2.2
6) Water hauling	Partial Bucket of Water	41.3
	Canada Goose Pond (with egg)	14.5
	Canada Goose Pond	6.0
	Full Bucket of Water	5.7
	Vigorous Domestic Gooseberry Bush	5.5
	Empty Bucket	5.5
	Dry Domestic Gooseberry Bush	4.4
	Cistern	4.1
	Pond	2.3
	Bowl of Water	1.9
7) Carpentry	Escaped Horse-Drawn Tire Cart	73.8
	Backpack	3.8
	Hitched Horse-Drawn Tire Cart	2.4
	Fence	2.3
	Written Object	1.4
	Yew Bow	0.8
	Chopped Softwood Tree	0.7
	Banana Plant	0.7
	Stone Pile	0.7
	Vulcanized Rubber Tire	0.6
8) Smithing	Firing Forge	15.8
	Flat Rock	12.1
	Iron Ore Pile	11.0
	Firebrand	9.9
	Wooden Tongs	9.4
	Kindling	8.9
	Firing Adobe Kiln	8.2
	Stack of Steel Ingots	7.4
	Stack of Wrought Iron	7.2
	Kindling Pile	6.6

Supplementary Table 1: **Top 10 objects that load most strongly onto each job.** Objects that are featured in Figure 2 (main text) are highlighted in yellow.

References

[1] Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401**, 788–791 (1999).

[2] Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011).

[3] Hsieh, C.-J. & Dhillon, I. S. *Fast coordinate descent methods with variable selection for non-negative matrix factorization*, 1064–1072 (2011).

[4] Yu, B. & Barter, R. *Veridical Data Science: The Practice of Responsible Data Analysis and Decision Making* (MIT Press, 2024).

[5] Salganik, M. J. *Bit by bit: Social research in the digital age* (Princeton University Press, 2019).

[6] Goldstone, R. L., Andrade-Lotero, E. J., Hawkins, R. D. & Roberts, M. E. The emergence of specialized roles within groups. *Topics in Cognitive Science* **16**, 257–281 (2024).

[7] Morgan, T. J. *et al.* Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nature Communications* **6**, 6029 (2015).

[8] Lucas, A. J. *et al.* The value of teaching increases with tool complexity in cumulative cultural evolution. *Proceedings of the Royal Society B* **287**, 20201885 (2020).

[9] Vélez, N., Chen, A. M., Burke, T., Cushman, F. A. & Gershman, S. J. Teachers recruit mentalizing regions to represent learners' beliefs. *Proceedings of the National Academy of Sciences* **120**, e2215015120 (2023).

[10] Muthukrishna, M. & Henrich, J. Innovation in the collective brain. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150192 (2016).

[11] Hidalgo, C. A., Klinger, B., Barabási, A.-L. & Hausmann, R. The product space conditions the development of nations. *Science* **317**, 482–487 (2007).

[12] Krugman, P. Increasing returns and economic geography. *Journal of Political Economy* **99**, 483–499 (1991).

[13] Jolly, E. & Chang, L. J. The flatland fallacy: Moving beyond low–dimensional thinking. *Topics in Cognitive Science* **11**, 433–454 (2019).

[14] Dakos, V. & Bascompte, J. Critical slowing down as early warning for the onset of collapse in mutualistic communities. *Proceedings of the National Academy of Sciences* **111**, 17546–17551 (2014).

[15] Turchin, P. Long-term population cycles in human societies. *Annals of the New York Academy of Sciences* **1162**, 1–17 (2009).

[16] Bandy, M. S. Fissioning, scalar stress, and social evolution in early village societies. *American Anthropologist* **106**, 322–333 (2004).

[17] Ostrom, E. *Governing the commons: The evolution of institutions for collective action* (Cambridge university press, 1990).