The spectrum generated with circular polarization contains both rotation- and vibration-generated lines. Consequently, there is a second coherent molecular excitation, implying the introduction of one additional random phase. This in itself does not destroy the phase coherence of either the pure vibrational or the pure rotational spectra, but to experimentally produce sub-fs pulses, these degrees of freedom need to be relatively controlled to correct for this additional random phase. Also, one could optimize the medium pressure to suppress completely the vibrational lines (6), leaving only the mutually coherent rotational lines.

The extension of using this HC-PCF in other Raman excitation regimes is straightforward. For example, its combination with the adiabatic preparation technique would enhance the conversion efficiency and further reduce the required pumping powers involved while ensuring better control over the spectral components’ phases. This would enable the generation and synthesis of attosecond pulses with much lower pumping powers. Furthermore, in addition to the intrinsic fundamental importance of the discovery of this “Von Neumann–Wigner”–related waveguidance to photonics, it will provide us with new tools to develop next-generation HC-PCFs with even broader bandwidth and lower transmission loss. For example, with lower loss figures, one could synthesize ultrashort pulses using continuous-wave pumps. This would permit the synthesis of arbitrary optical waveforms with a degree of control approaching that in electronics.

References and Notes
15. Supporting online material is available on Science Online.
24. We are grateful to T. Birks, D. Bird, and C. Wu for stimulating discussions. This work is funded by the UK Engineering and Physical Sciences Research Council (EPSRC). M.G.R. was supported by the U.S. NSF (grants ECS-0621723/PHYS-0456974/PHYS-0558482).

Supporting Online Material
www.sciencemag.org/cgi/content/full/318/5853/1118/DC1
Methods
SOM Text
Figs. S1 to S7
References
13 August 2007; accepted 4 October 2007
10.1126/science.1149091

Engineering Entropy-Driven Reactions and Networks Catalyzed by DNA

David Yu Zhang,† Andrew J. Turberfield,‡ Bernard Yurke,‡* Erik Winfree†

Artificial biochemical circuits are likely to play as large a role in biological engineering as electrical circuits have played in the engineering of electromechanical devices. Toward that end, nucleic acids provide a desirable substrate for the regulation of biochemical reactions. However, it has been difficult to incorporate signal amplification components. We introduce a design strategy that allows a specified input oligonucleotide to catalyze the release of a specified output oligonucleotide, which in turn can serve as a catalyst for other reactions. This reaction, which is driven forward by the configurational entropy of the released molecule, provides an amplifying circuit element that is simple, fast, modular, composable, and robust. We have constructed and characterized several circuits that amplify nucleic acid signals, including a feedforward cascade with quadratic kinetics and a positive feedback circuit with exponential growth kinetics.

The development of modular biochemical circuit elements poses several challenges.

First, distinct signals must be carried by distinct chemical species, motivating the use of information-carrying molecules whose sequences can be used to encode signal identity. Second, “wiring up” a gate to specified inputs and outputs involves the design and synthesis of new molecules; this calls for modular gate designs. Third, a fast and robust catalytic mechanism must be identified and coupled to a suitable energy source in order to create gates with signal gain. Fourth, it must be possible to construct circuits of arbitrary complexity that can produce an unlimited variety of dynamical behaviors. Finally, there should be no leak or crosstalk between distinct signals and gates. It is difficult to meet all these challenges simultaneously.

Nucleic acids are attractive for this purpose because the combinatorial sequence space allows for an enormous diversity of signal carriers, and the predictability and specificity of Watson-Crick base pairing facilitate the design of gate architectures. The “RNA world” hypothesis further suggests that sophisticated biochemical organization can be achieved with nucleic acids alone (1), and nucleic acids have indeed been shown to be a versatile construction material for engineering molecular structures and devices (2, 3), including catalytic (4–8) and logical (9–12) control elements and circuits (13–17). Engineering (deoxy)ribonucleic-based logic gates has been very effective, resulting in systems containing over 100 gates operating independently in parallel (10) as well as systems demonstrating cascading of a signal between two gates (13, 15, 16). Alternatively, hybridization-based systems, usually driven by the energy of base-pair formation, have proven especially suitable for cascading signals, as demonstrated by a circuit five layers deep (17). That work, relying primarily on noncatalytic logic gates, identified amplification and signal gain as essential for scaling up to large cascaded circuits. We provide a solution to this problem.
The entropy-driven catalytic gate presented here is substantially simpler than previous hybridization-based designs; moreover, it is faster, better understood, and more modular. The net reaction is shown in Fig. 1A: Fuel strand (F) reacts with the three-stranded substrate complex (S), displacing output and signal strands (OB and SB) from linker strand (LB) to form waste complex (W). The total number of base pairs in the reactants and products is unchanged; the reaction is driven forward thermodynamically by the entropic gain of the liberated molecules. Fuel, signal, catalyst, and output are all single-stranded DNA molecules that can be of similar lengths; thus, each molecule may play multiple roles within a network. For example, the output of one gate may serve as the input to another. Notably, catalyst C and output OB may be entirely independent in sequence (18); this modularity implies that a catalytic gate can be designed to act at any point within a preexisting circuit. Unlike previous hybridization-based catalyst systems, the reaction design does not require unusual secondary structures such as pseudoknots and kissing loops. Undesired interactions can be avoided by design (19–21), resulting in reliable and predictable circuit behavior.

Strands are conceptually subdivided into functional domains (number labels in Fig. 1) whose sequences determine the pattern of interactions between circuit components. [Domain sequences are given in Table 1; see supporting online material (SOM) text S1 for design details.] The domains can be grouped by purpose: domains 3 and 5 are termed toehold domains, whereas domains 1, 2, 4, and 6 are termed specificity domains. Toehold domains are short enough to bind only fleetingly in the absence of additional binding (and need not be distinct), but they greatly accelerate the initiation of strand displacement reactions (22). Specificity domains ensure specific interactions [even a single mismatch can slow down branch migration substantially (23)] and determine the identities of the catalyst and output molecules. The lengths of the toehold domains determine kinetics and need to be between roughly 4 and 10 nucleotides (nt), but the specificity domains may be of any length sufficient to ensure thermal stability. Domains 1 and 6 of OB and SB, respectively, are inert, whereas their respective toeholds are sequestered in S.

Catalytic activity has two characteristic behaviors: the speedup of the target reaction and the re-release of the catalyst to allow for multiple turnover. To achieve these behaviors, we introduce and apply a design principle that we call toehold exchange (Fig. 1B): C first binds to the single-stranded toehold domain S on S to form the four-stranded intermediate I4, which then rearranges (by branch migration) to form I5. The binding between toehold domains 3 and S is too weak to keep SB attached, so I2 spontaneously dissociates into SB and I3. Newly exposed S then facilitates the binding of F, resulting in A4, which then quickly rearranges to release OB and I5. Finally, I5 rearranges so that C is attached only by the binding of 5 and S, which spontaneously dissociates to leave W and regenerate C. The reaction mechanism presented here, based on branch migration and driven by entropy, differs from the traditional view of catalysis in biological organisms in that it requires no enzymes and alters no covalent bonds.

It is important to ensure that alternative interactions do not interfere with intended gate functions. Toward this end, a key design principle is that the complements of the specificity domains never appear in their single-stranded form. Domain X is the complement of (and will hybridize to) domain x. (B) The proposed catalytic pathway. Reverse reactions are also present and modeled (with the exception of S+OB→I4, which occurs at a negligible rate). (C) Analysis by PAGE (12% native gel) of the reaction mechanism. Unless otherwise noted, all experiments were performed at 25°C in Tris-acetate (1X) buffer supplemented with 12.5 mM MgCl2. Here, [S] = [F] = 200 nM, [C] = 200 nM, except where C denotes 20 nM. "ann." denotes that species were annealed; "30 m" denotes that the reaction occurred for 30 min. See fig. S5 for the full gel, including control lanes. (D) Fluorescent reporter strategy. ROX denotes the carboxy-X-rhodamine fluorophore, and RQ denotes the Iowa Black Red Quencher. Domain 2 is subdivided into 2a, 2b, and 2c; 2ab consists of 2a and 2b (Table 1). (E) Demonstration of catalysis. Different amounts of C were introduced into the system at t = 0. Here, [S] = 10 nM = 1x, [F] = 13 nM, and [OR] = 30 nM. Fluorescence (in all figures) is reported in units such that 0.0 is the background fluorescence of the quenched reporter and 1.0 is the fluorescence of ~10 nM of triggered reporter. The control trace (black) shows the reaction with no substrate S and no catalyst C. Dotted lines show curves calculated with the reduced reaction model. sim, simulated.
I prevent the reaction from progressing past intermediate \( I \). The amount of \( I \) produced after 30 min (lane 2) is almost identical to that present at equilibrium, as assessed by annealing the reaction components (lane 3). This suggests that all reactions up to \( I \) are fast on this time scale. Similarly, the subsequent reaction between \( I \) and \( F \) is also fast (lanes 3 to 5). The complete system behaves as expected: The uncatalyzed reaction is slow (lanes 7 and 8), and a substoichiometric quantity (0.1×) of \( C \) enables the reaction to proceed rapidly to near-completion (lanes 9 and 10).

In order to measure the time course of the catalyzed reaction by means of a fluorescent reporter without interference from fluorophore-quencher interactions (25) (SOM text S3), we use an indirect reporter complex \( OR \). \( OR \) reacts stochiometrically with output \( OB \) to separate a fluorophore-labeled strand from a quencher-labeled strand, thereby increasing fluorescence (Fig. 1D). The rate constant for the reporter system was measured to be \( k_{\text{ROX}} = 4 \cdot 10^5 \text{M}^{-1} \text{s}^{-1} \) (SOM text S4). Because initial \( [OR] = 30 \text{nM} \) is in excess to \([S] = 10 \text{nM}\), the reporter complex remains substantially in excess, and the reporting delay should remain less than 100 s, which is short as compared to the half time of the catalyzed reactions. \( OR \) does not react substantially with \( S \), because there are no single-stranded toeholds to initiate interaction. Measurements of the kinetics of the catalyzed reaction over a 500-fold range of catalyst concentration are shown in Fig. 1E.

We modeled this system using the reduced reaction set shown below.

\[
S + F \xrightarrow{k_1} SB + OB + W
\]

\[
S + C \xrightarrow{k_{SB}} I3 + SB
\]

\[
I3 + F \xrightarrow{k_2} I5 + OB
\]

\[
I5 \xrightarrow{k_{OB}} C + W
\]

\[
OB + OR \xrightarrow{k_{\text{ROX}}} \text{Fluorescence}
\]

where \( k_0 = 2.3 \cdot 10^4 \text{M}^{-1} \text{s}^{-1} \),

\( k_1 = 6.5 \cdot 10^2 \text{M}^{-1} \text{s}^{-1} \),

\( k_2 = 4.2 \cdot 10^2 \text{M}^{-1} \text{s}^{-1} \),

\( k_3 = 4 \cdot 10^{-2} \text{s}^{-1} \) (fitted), and

\( k_{\text{ROX}} = 4 \cdot 10^5 \text{M}^{-1} \text{s}^{-1} \)

The first reaction shown models the uncatalyzed (leak) reaction. Intermediate steps in branch-migration reactions are omitted, because they are relatively fast at experimental concentrations (SOM text S5) (26) and because intermediates \( I1, I2, \) and \( I4 \) are not observed in PAGE analysis of reactants and products (Fig. 1C). Noting the approximate symmetry between the corresponding reactions, we assume that \( k_3 = k_1 \) and \( k_{-1} = k_2 \). The rate constants \( k_0, k_1, \) and \( k_2 \) were measured individually (fig. S4); \( k_3 \) was fit to the data of Fig. 1E. The time course of the catalyzed reaction over a wide range of catalyst concentrations is accurately reproduced by this reduced system of rate equations (Fig. 1E).

According to this model, the addition of catalyst can accelerate the reaction by over four orders of magnitude (\( k_3/k_0 = 1.8 \cdot 10^4 \)).

In the net reaction, each base pair that is broken is replaced by another of the same type, so the net free energy change from base-pairing interactions should be small. The reaction is driven by the gain in configurational entropy corresponding to the liberation of \( OB \) and \( SB \) at the cost of localizing \( F \). To confirm the dominance of this entropic driving force, we truncate \( F \) by removing up to 8 nt from its 5′ end, making the products more and more thermodynamically disfavored. Nonetheless, in all cases

---

**Table 1.** Domain sequences of basic catalytic reaction.

<table>
<thead>
<tr>
<th>Domain</th>
<th>Sequence</th>
<th>Length (nt)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5′-CTTCTCTACA-3′</td>
<td>10</td>
</tr>
<tr>
<td>2a</td>
<td>5′-CCATCG-3′</td>
<td>6</td>
</tr>
<tr>
<td>2b</td>
<td>5′-TCTCA-3′</td>
<td>6</td>
</tr>
<tr>
<td>2c</td>
<td>5′-ACTAACTTACGG-3′</td>
<td>12</td>
</tr>
<tr>
<td>3</td>
<td>5′-CCCT-3′</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>5′-CATTCAATACCTACGG-3′</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>5′-TCTCA-3′</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>5′-CCACATACATATT-3′</td>
<td>16</td>
</tr>
</tbody>
</table>

---

**Fig. 2.** Verification of entropic driving force. Analysis by PAGE (12% native gel) of reactions with truncated fuel strands. \([S] = [F] = 200 \text{nM} \); \([C] = 20 \text{nM}\), as denoted by the asterisk. All reactions were run at 25°C for 3 hours. “Ft2” denotes that two bases were truncated from the 5′ end of fuel strand \( F \).

**Fig. 3.** A two-layer cascaded network. (A) Schematic. See table S2 for sequences of new domains. (B) Kinetics. Indicated amounts of initial catalyst \( C0 \) were added at \( t = 0 \). Fluorescence derives from reporter complex \( OR \) (Fig. 1D) at 30 nM. Dotted lines show simulated traces; see SOM text S8 for details on reaction rates and modeling. a.u., arbitrary units. (Inset) Response to 0.0010x, 0.0003x, and 0.0001x catalyst. The asterisk indicates that three independent reaction traces are shown. 1.0 fluorescence units correspond to \( \approx 10 \text{nM} \) of triggered reporter.
the waste product is favored at equilibrium (Fig. 2; see SOM text S6 for discussion on entropy and free energies). The thermodynamic driving force, being dominated by center-of-mass configurational entropy of released molecules, is somewhat robust to environmental conditions such as temperature and salt concentrations that alter the strength of DNA hybridization (SOM text S7 and fig. S6).

To demonstrate cascaded circuit construction, we designed a two-layer feedforward network by introducing an upstream catalyst system whose output acts as the catalyst for the original system (Fig. 3A). For clarity, F, OB, and the other reagents and products from Fig. 1 are relabeled F1, OB1, and so forth. Catalyst C0 catalyzes the production of OB0 (which contains a subsequence identical to C from Fig. 1), which in turn catalyzes the production of OB1. The concentration of upstream catalyst C0 is constant, so initially [OB0] increases linearly with time, which causes [OB1] to increase quadratically with time (Fig. 3B). Eventually, the substrates and fuels are depleted, and the reaction halts, giving rise to an overall sigmoidal shape to the fluorescence traces (Fig. 3B). The model previously used can be extended to predict the behavior of this feedforward circuit data (SOM text S8).

This cascaded system can be used as an amplifier to detect small quantities of C0. Repeated fluorescence experiments show that we are able to distinguish reliably between 1 pM (0.0001×) catalyst C0 and 0× catalyst within 12 hours (Fig. 3B, inset). This corresponds to a roughly 900-fold amplification of the input signal. (1 pM of catalyst triggered ≈ 900 pM of reporter above the baseline set by the 0× reaction.) For comparison, 1 pM corresponds to about one molecule per eukaryotic cell volume. Repeated measurements of independent samples show less than 3% variability across all timepoints (SOM text S9).

Feedback is another important feature of both biological regulatory networks and artificial control circuits. Exponential growth kinetics can be achieved by redesigning the reaction presented in Fig. 1 such that output OB contains catalyst C as a subsequence (Fig. 4A). The reaction is then autocatalytic. Figure 4C shows the time course of this reaction for a wide range of catalyst concentrations. In a process dominated by initial exponential growth \( (c = ce^{ct}) \), the time to reach a threshold degree of completion depends logarithmically on the initial concentration \( c_0 \) (where \( c \) is the concentration of the exponentially growing species, \( λ \) is the characteristic time constant, and \( t \) is time). Thus, a linear trend in a log-linear plot of initial concentration to time to half completion \( (t_{1/2}) \) is indicative of exponential growth. [Such plots are used as calibration standards for quantitation methods such as real-time polymerase chain reaction (PCR) (27).]

Fig. 4. The autocatalyst system. (A) Proposed reaction pathway. See table S2 for sequences of new domains. (B) Reporter complex SR, used for monitoring autocatalytic and cross-catalytic reaction networks. TET denotes the tetrachlorofluorescin fluorophore. FQ denotes the Iowa Black Fluorescent Quencher. (C) Kinetics of autocatalysis. Indicated amounts of autocatalyst were added at \( t = 0 \). At 30 min, 1% (100 pM) was amplified 25-fold over the untriggered reaction. Reporter SR was present at 20 nM = 2×. Control sample (black) contained no fuel F. (D) Semi-log plot of \( t_{1/2} \) (in minutes) as a function of logarithm of molar trigger concentration. The orange line shows \( t_{1/2} \) of the untriggered reaction. (E) Performance in the simultaneous presence of total RNA and unfractionated cell lysate. Total RNA present in solution was 10× (by mass) that of the sum of all relevant catalyst DNA. Active cell lysate, as would be used in an in vitro translation system, was added to be 1.1% by volume of total reaction (17 μL in 1500 μL). The control reaction did not contain any substrate S. Experiments involving total RNA but not cell lysates did not show the observed drift.
the point that our autocatalyst could be used as an enzyme-free constant-temperature alternative to PCR for detecting known sequences.

For many applications in biotechnology, nucleic acid devices must remain functional in the presence of naturally occurring macromolecules. We therefore tested the autocatalyst system in the presence of an excess of mouse liver total RNA with rabbit reticulocyte lysate (Fig. 4E). Reactions proceeded to apparent completion with no more than a twofold slowdown, and the presence of a 3% trigger can still be detected.

The ability to construct larger circuits will enable the wide range of chemical circuit functions needed for sophisticated applications. Our aptamer-driven catalytic reaction networks are suited for scaling up to larger circuits. The modular molecular design makes synthesis of more complex components and networks with arbitrary topology straightforward. To demonstrate this, we constructed an aptamer-driven catalytic analog AND gate in which both of two catalysts are required to release output (SOM text S12 and fig. S11). For scaling up to large circuits, independent catalyst systems must have negligible crosstalk. The success of quantitative models that assume no crosstalk, as presented above, is encouraging; further evidence comes from a test of two independent catalyst systems operating in the same solution (fig. S12). Finally, catalytic systems have the potential to avoid the slowdown that plagued previous attempts to construct large nucleic acid circuits (17).

Future nucleic acid control circuits must be interfaced to molecular sensors and actuators. This may be achieved directly when the inputs and outputs are themselves nucleic acids, such as for the detection, analysis, and response to complex nucleic acid samples (9, 30) or for the control of nucleic acid nanomachines (2, 31). Nucleic acid circuits can also respond to and control more general chemical events: In principle, the release of an oligonucleotide could regulate covalent chemical events in isolation, there are no sequence constraints; we demonstrate a system with completely independent catalyst and output in SOM text S10; it has very similar kinetics to that of the reaction shown in Fig. 1.


In the system presented in Fig. 1, there is some sequence redundancy in the domain sequences chosen (for example, 2b and 5 are identical). This is because all four systems presented in this paper were designed together, with the goal of minimizing the number of differences between systems. In the design of a catalytic reaction in isolation, there are no sequence constraints; we demonstrate a system with completely independent catalyst and output in SOM text S10; it has very similar kinetics to that of the reaction shown in Fig. 1.


24. Materials and methods (SOM text S2) are available as supporting material on Science Online.

---

Radar Sounding of the Medusae Fossae Formation Mars: Equatorial Ice or Dry, Low-Density Deposits?


The equatorial Medusae Fossae Formation (MFF) is enigmatic and perhaps among the youngest geologic deposits on Mars. They are thought to be composed of volcanic ash, eolian sediments, or an ice-rich material analogous to polar layered deposits. The Mars Advanced Radar for Subsurface and Ionospheric Sounding (MARSIS) instrument aboard the Mars Express spacecraft has detected nadir echoes offset in time-delay from the surface return in orbits over MFF material. These echoes are interpreted to be from the subsurface interface between the MFF material and the underlying terrain. The delay time between the MFF surface and subsurface echoes is consistent with massive deposits emplaced on generally planar lowlands materials with a real dielectric constant of $\approx 2.9 \pm 0.4$. The real dielectric constant and the estimated dielectric losses are consistent with a substantial component of water ice. However, an anomalously low-density, ice-poor material cannot be ruled out. If ice-rich, the MFF must have a higher percentage of dust and sand than polar layered deposits. The volume of water in an ice-rich MFF deposit would be comparable to that of the south polar layered deposits.

Units of the Medusae Fossae Formation (MFF) occur discontinuously at equatorial latitudes along the boundary of the hemispheric dichotomy from Amazonis to Elysium Planitiae (~130°E to 240°E) (1, 2). The MFF may be among the youngest surficial deposits on Mars, unconformably overlying ancient Noachian heavily cratered highlands and young Amazonian lowlands (1–8). However, pedal cratering on the outer edge of the MFF...
Reports: “Engineering entropy-driven reactions and networks catalyzed by DNA” by D. Y. Zhang et al. (16 November 2007, p. 1121). In Fig. 4B, domain 4a should have been domain 4t, with a length of 7 nucleotides. The corrected figure appears below. The following text should also be added to the Fig. 4B caption: “Domain 4t has identity 5'-TTGAATG-3' and is a subsequence of domain 4a.”
Disclosure of NDIS profiles would not violate any meaningful privacy interests (12). (There are easier ways to determine whether an individual has a criminal record than searching such a database, and the profiles would not be useful for medical diagnoses.) The profiles in the Victoria, Australia, database have been widely circulated for years with no known harm occurring. The U.S. government regularly argues to courts that broad mandatory DNA collection statutes are not unconstitutional precisely because the 13 genetic loci are noncoding and thus have no power to reveal any sensitive information. Moreover, as most research scientists know well, the government frequently releases sensitive information under controlled conditions to verified researchers. Even within the criminal justice context, law enforcement officials have made available data about the age, race, gender, geographic residence, and a wide range of other information about criminal offenders so that researchers can conduct studies aimed at improving and enhancing effective law enforcement.

Disclosure of NDIS profiles would allow independent scientists to evaluate some of the population genetic assumptions underlying DNA testing using a database large enough to allow more sensitive evaluation of population structure. The publicly available population databases used to date for statistical estimation of the frequency of DNA profiles are relatively small (N ≈ 1000), consisting of convenience samples analyzed over a decade ago (2, 3). In contrast, NDIS has grown to over 7 million complete 13-locus short tandem repeat (STR) genotypes (4). Analysis of these data would allow more powerful tests of independence within and between loci, as well as assessment of the efficacy of the theta factors used to compensate for population substructure. (To the extent the data are identified by state, analysis of NDIS data could also yield important information about the most appropriate geographic scaling for allele frequency estimates.)

The large sample size also allows real-world tests of propositions that previously have been addressed only by simulation. For example, it would allow tests of the frequency with which three-person mixtures could produce profiles consistent with two contributors (5); kinship analysis could allow assessment of how match probabilities are affected by the number of close relatives in the database (6, 7); and multivariate analysis could be used to evaluate the extent to which DNA profiles cluster due to identity by descent. As studies of smaller databases have shown, researchers need not know a priori the precise number of relatives in the database, nor their ethnic/racial background, to perform these assessments (6, 8). Indeed, scholars who have examined smaller databases have called for examination of national databases (6, 8, 9). Access to the anonymized 13-locus genotypes would allow more powerful analyses of these important issues than was previously possible.

Analysis of NDIS can also yield valuable insights into the frequency and circumstances under which certain typing errors may occur. A review of a government database from Victoria, Australia, containing 15,021 9-locus STR profiles shows how important such a review can be for “quality control purposes” (10, 11). The study found an error rate of about 1 in 300 for the typing of reference samples, which raises concerns about missed opportunities to develop investigative leads.

Disclosure of NDIS profiles would not violate any meaningful privacy interests (12). (There are easier ways to determine whether an individual has a criminal record than searching such a database, and the profiles would not be useful for medical diagnoses.) The profiles in the Victoria, Australia, database have been widely circulated for years with no known harm occurring. The U.S. government regularly argues to courts that broad mandatory DNA collection statutes are not unconstitutional precisely because the 13 genetic loci are noncoding and thus have no power to reveal any sensitive information. Moreover, as most research scientists know well, the government frequently releases sensitive information under controlled conditions to verified researchers. Even within the criminal justice context, law enforcement officials have made available data about the age, race, gender, geographic residence, and a wide range of other information about criminal offenders so that researchers can conduct studies aimed at improving and enhancing effective law enforcement.
Weighing Reward and Punishment

IN THEIR REPORT “POSITIVE INTERACTIONS promote public cooperation” (4 September, p. 1272), D. G. Rand et al. find that targeted reward is at least as effective as targeted punishment in maintaining cooperation. In their experiment, infrequent reward may be sufficient because the group is small and interacts repeatedly. However, in real-world situations, punishment may be the more effective and cost-efficient option.

In many real-world cases, unlike Rand et al.’s example, the cost to Player A of giving Player B a material reward is roughly the same as the benefit Player B receives from the reward. (The benefit of nonpecuniary rewards, such as praise, may exceed their cost considerably. Rand et al. suggest this, but their experiment is not set up to provide evidence.) Thus, the cost of cooperation is simply shifted to those who provide the reward. However, the threat of punishment provides a less costly lever to force cooperation, even when the threat must be carried out. The cost of a match and a gallon of gasoline is much less than cost to repair the damage they could cause. Likewise, nasty words can hurt much more than the effort it takes to say them.

In real-world situations, when people are not interacting in a small group and when they are motivated by money, the threat of punishment is effective. Laws are based largely on this insight.

JONATHAN BARON

Response

BARON ARGUES THAT INFREQUENT PUNISHMENTS are more cost efficient than infrequent rewards. But our experiment does not represent a situation of intermittent rewarding. Instead, we have shown that contributions to the public good can be maintained by linking the public goods game to cooperative, wealth-producing pairwise interactions. Low contributors are denied cooperation in pairwise interactions, while high contributors are rewarded. Due to the ubiquity of such opportunities for targeted interaction, there is no need for costly peer punishment to enforce cooperation. Full cooperation in both the public and pairwise interactions leads to the best possible payoff. Thus, adding punishment cannot result in better outcomes.

Baron challenges the real-world applicability of the non-zero-sum rewards in our study. However, the availability of wealth-generating, non-zero-sum interactions is the essence of all social dilemmas—including the Prisoners’ Dilemma (1–5), of which our reward interaction is an example, as well as the Public Goods Game (6–9) itself. These games represent the multitude of different

References

cooperative interactions in which two or more people working together can achieve more than each person could alone. For example, consider mutually beneficial trade: Both parties pay the cost of abandoning something worth less to them than to the other, in order to gain something they find relatively more valuable. To enforce public cooperation, one can refuse to trade with those who do not contribute to the public good. Baron’s claim that life offers few opportunities to create material benefits for others through cooperation questions the relevance of all work on social dilemmas, including his own (10).

Baron concludes by mentioning the role of punishment in law. However, our paper and most others on costly punishment (4, 5, 7–9) investigate peer punishment, not institutionalized punishment. The latter deserves further empirical and theoretical exploration.

DAVID G. RAND,1,2 ANNA DREBER,1,3 TORE ELLINGSEN,3 DREW FUDENBERG,4 MARTIN A. NOWAK1,5,6

1Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA. 2Berkman Center for Internet and Society, Harvard University, Cambridge, MA 02138, USA. 3Department of Economics, Stockholm School of Economics, 11358 Stockholm, Sweden. 4Department of Economics, Harvard University, Cambridge, MA 02138, USA. 5Department of Mathematics, Harvard University, Cambridge, MA 02138, USA. 6Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.

*To whom correspondence should be addressed. E-mail: martin_nowak@harvard.edu

References