Journal of Big History

Volume II Number 2 Summer 2018

International
Big History
Association

The Journal of Big History (JBH)
ISSN 2475-3610

JBH, the journal of the International Big History Association (IBHA), is published on line and is accessible to members of the IBHA. The IBHA also publishes its bulletin, *Origins*. The bulletin contains members' accounts of their activities and views, announcements of conferences and meetings of interest, and other communication of interest to IBHA members.

We encourage readers of the JBH to send letters and responses to jbh@bighistory.org to the address below:
Brooks College of Interdisciplinary Studies
Grand Valley State University
1 Campus Drive
Allendale MI 49401-9403
https://bighistory.org/
616-331-8035
jbh@bighistory.org

Cover photo by Erick Gustafson

Please visit journalofbighistory.org; please submit manucripts and other material to *Journal of Big History* (JBH) at jbh@bighistory.org.
The views and opinions expressed in JBH are not necessarily those of the IBHA Board.
JBH reserves the right to accept, reject or edit any material submitted for publication.

The IBHA appreciates its members for their participation in the IBHA. Membership dues all go towards the administration of the association, but do not by themselves cover our costs. The only paid position is a part time administrative assistant. Other costs are for our website and for the JBH, for example. Please consider a tax deductible (in the US) gift to our 501(C)3 and please consider remembering the IBHA in your will.

Contents

Articles

Book Review

Guerra, Ética, Etologia

fundamentos evolucionários do conflito e da cooperação na linhagem do homem

Daniel Barreiros
Universidade Federal do Rio de Janeiro - UFRJ - Brasil

RESUMO

O objetivo deste artigo é o de estabelecer uma narrativa macrohistórica acerca da emergência da guerra e da ética social como condições simplesiomórficas na linhagem de *Homo sapiens*. Isso significa que esses dois aspectos comportamentais, que representam um ramo bastante seleto na árvore filogenética da ordem dos Primatas, são compartilhados por duas linhagens de grandes símios africanos que divergiram de um ancestral comum por volta de seis milhões de anos atrás, resultando em humanos e chimpanzés extantes. Dessa forma, o artigo propõe uma interpretação etológica para a guerra e para a ética social, sendo ambas inatas de uma mente social modular altamente especializada, presente tanto nas espécies do gênero *Homo* quanto do gênero *Pan*. Não obstante essa interpretação, o artigo conclui que restrições comportamentais à violência interssocietária coalizacional parecem ser um aspecto exclusivo da cognição modular transdominial que caracteriza os humanos modernos. Assim, se na longa duração evolucionária, a guerra e as restrições à violência intrassocial aparecem em certa medida como traço etológico comum a humanos e chimpanzés, uma ética da guerra – e a capacidade cognitiva para a paz interssocietária – parece ser uma capacidade exclusivamente humana.

Correspondence | Daniel Barreiros, daniel.barreiros@ie.ufrj.br
Citation | Barreiros, D. (2018) Guerra, Ética, Etologia: fundamentos evolucionários do conflito e da cooperação na linhagem do homem. *Journal of Big History*, II(2); 1 - 18.
DOI | http://dx.doi.org/10.22339/jbh.v2i2.2299

O objetivo deste artigo é o de estabelecer uma narrativa macrohistórica acerca da emergência da guerra e da ética social como condições simplesiomórficas na linhagem de *Homo sapiens*. Isso significa que esses dois aspectos comportamentais, que representam um ramo bastante seleto na árvore filogenética da ordem dos Primatas, são compartilhados por duas linhagens de grandes símios africanos que divergiram de um ancestral comum por volta de seis milhões de anos atrás, resultando em humanos e chimpanzés extantes. Essa narrativa não pretende substituir qualquer interpretação consolidada acerca da guerra e da paz defendida pelas ciências sociais e humanidades em geral; pelo contrário, seu objetivo é o de proporcionar uma nova profundidade a essas interpretações, a partir de um ponto de vista evolucionário. Esperamos que todas as contradições geradas por esse "jogo de escalas" proporcionem reflexões dialéticas a respeito da relação entre ação social e a duração histórica, no caminho trilhado por Braudel (2009) e Christian (2005).

Muitas são as definições correntes sobre os fenômenos da guerra e da ética nas relações humanas, mas que, na perspectiva deste ensaio, acabam por comungar de um mesmo limite. A generalidade dessas dimensões do comportamento humano, verdadeiro estimulante da imaginação – que faz surgir um *homo ethicus* ou, na face oposta, *homini lupus homini* – deixa escapar o quão raras foram as condições que permitiram eclodir o conflito entre sociedades, e mais ainda, mecanismos comportamentais voltados para facilitá-lo ou preveni-lo.

Na longuíssima duração do tempo evolucionário, e considerando a trajetória das espécies na ordem dos Primatas, a violência coalizacional intersocietária – definição que torna todas as guerras humanas casos particulares – e a complexa rede de mecanismos etológicos[1] que dá fundamento, no campo do inconsciente humano, aos saberes éticos, são pontos fora da curva. Não nos referimos aqui à violência interpessoal, esse traço comportamental comum e que, na condição de diferencial evolucionário, envolve normalmente a disputa por energia e oportunidades

1 Etologia é o estudo do comportamento animal.

reprodutivas. Certa sociabilidade primata, há pelo menos seis milhões de anos, gerou o contexto para a fixação de intrincados instrumentos cognitivos voltados para a resolução de conflitos intragrupo, envolvendo complexas hierarquias de *status*, violência não letal, ritualizações e estratagemas sociais; e simultaneamente, foi base para esse raro fenômeno, consubstanciado na patrilinearidade cooperativa masculina, a partir do qual emergiu a violência letal extragrupo.

Fazem a guerra, nessa macroperspectiva, humanos e chimpanzés-comuns. Também são eles atores em complexas tramas sociais cotidianas, nas quais o equilíbrio de força e prestígio entre "competidores cooperativos" é volátil, e o potencial de violência letal fratricida tem de ser mantido sob controle pela ação de uma mente modular social altamente desenvolvida. Compartilham da guerra, de freios etológicos, e de 98,8% de seus genes. Se todas essas características não emergiram independentemente nas linhagens que resultam nessas duas espécies, precisam ter estado presentes também, ao menos, no último ancestral comum entre humanos e chimpanzés. Seria filogenético,[2] desse modo, o potencial comportamental para a projeção de poder externa e para a contenção à escalada do conflito intrassocial? Numa perspectiva evolucionária, seriam a guerra e a ética filhas do mesmo ventre? E o que dizer de uma ética da guerra?

1. Sociabilidade multissexual instável no Eoceno Inicial

A eclosão da sociabilidade entre determinadas espécies primatas, estimada em 52 milhões de anos atrás, não parece ter sido elemento suficiente para engendrar o contexto evolutivo específico que possibilitaria fundar as bases etológicas para a guerra e a ética. O Eoceno Inicial[3] trouxe, com algumas prováveis novas espécies, a inovação comportamental na formação de agremiações multissexuais instáveis, em detrimento dos hábitos solitários, que deviam seguir marcando outras tantas espécies. Provavelmente, a tônica da sociabilidade entre esses antepassados devia girar em torno de uma profunda volatilidade quanto à composição interna dos membros do grupo, com alto nível de fusão-fissão, respondendo à saturação demográfica e/ou oferta de alimentos. Nesse quadro, tendem à migração periódica tanto machos quanto fêmeas.

Se considerado o contexto climático de aquecimento global, de homogeneidade ambiental, de expansão das florestas tropicais pelos continentes (o que incluía a Antártida) e de ampliação da oferta energética, a dispersão territorial dos indivíduos nas unidades sociais tenderia a ser igualmente ampla, considerando-se o potencial de forrageamento individual eficiente por vastas áreas. A adoção de hábitos diurnos pode ter estado associada tanto ao desenvolvimento dessa sociabilidade instável quanto ao da visão estereoscópica, que expande a percepção de profundidade. Em espécies arbóreas, capacitava esses organismos a localizar frutos de maior valor nutricional em condições de poluição visual (florestas fechadas, com pouca luz), permitindo a exploração desse rico nicho criado pela expansão da vegetação angiospérmica. Ao mesmo tempo, a diuturnidade gerava um considerável passivo para essas espécies, provindo da maior exposição ao risco de predadores. Contra esse risco, a gregariedade atuava como estratégia de equilíbrio, aumentando o número de unidades sensoriais dispostas a simultaneamente detectar a presença da ameaça, e a disseminar essa informação para o benefício coletivo.

É assim que, possivelmente, a sociabilidade instável entre os primatas deva ter se constituído: na condição de estratégia antipredatória, meramente pragmática, sem gerar vínculos duradouros, tampouco cooperação complexa entre seus atores (Groves, Cameron, 2004, p. 36; Ladeia, Ferreira, 2015, pp. 56-58; Shultz, Opie, Atkinson, 2011, pp. 219; 222).

2 Diz-se filogenético um traço herdado por uma espécie em uma linha de ancestralidade que remete a outras espécies.

3 Eoceno foi o período geológico compreendido entre 56 e 33,9 milhões de anos atrás, de acordo com a Comissão Internacional de Estratigrafia (ICS). Mais informações podem ser obtidas em <http://www.stratigraphy.org/index.php/ics-chart-timescale>. A geocronologia empregada neste ensaio segue a convenção internacional.

2. Mudança climática do Oligoceno ao Mioceno Inicial: *Proconsul* e a sociabilidade matrilinear feminina estável cooperativa

Com a relativa homogeneidade climático-ambiental do Eoceno dando lugar, a partir do Oligoceno,[4] a um progressivo resfriamento global e à aridificação, o tabuleiro evolucionário africano foi significativamente perturbado, abrindo-se mais um contexto rico para especiações e extinções. A essa transformação climática somavam-se, no Mioceno Inicial,[5] intensa atividade tectônica e mudanças orográficas dela advindas, que resultaram no soerguimento do Himalaia, do Planalto Tibetano e do Altiplano Etíope. O relevo preveniu a entrada, no continente africano, de correntes de ar úmido vindas do Oceano Índico; tal fato teve forte impacto na África oriental, no sentido de gerar uma "colcha de retalhos" ambiental com tendência à aridificação. A multiplicidade de nichos contribuiu para a fixação de soluções evolucionárias inovadoras no âmbito das espécies primatas.

Proconsulidae é o nome dado à família que reuniu cerca de uma dezena de espécies miocênicas, surgidas por volta de 23 milhões de anos atrás, das quais *Proconsul africanus* é a espécie mais conhecida. A anatomia dentária grácil comum aos proconsulídeos sugere hábitos arborícolas em florestas tropicais e subtropicais fechadas, e o consumo de frutas e folhas macias, algo não diferente de seus antepassados eocênicos e oligocênicos. Dentição com fina camada de esmalte tende a ser mais propensa ao desgaste, tornando mais efetiva a ação cortante das cúspides; encontramos esse tipo de adaptação normalmente associado a primatas que adotam dietas de folhas e frutos macios, pouco abrasivos e que exigem pouca preparação e mastigação (Pampush et al., 2013, p. 218).

Devemos considerar, contudo, que as condições climático-ambientais na África oriental do Mioceno vinham criando "ilhas" de florestas cercadas por espaço savanizado, e que, se os proconsulídeos dependeram mesmo da exploração dos recursos florestais, a circunscrição de seu alcance espacial nesses territórios insulados decerto impactou em suas estratégias sociais. Além disso, outro aspecto relevante consiste da possibilidade, sugerida a partir do registro fóssil, de um acentuado dimorfismo sexual entre essas espécies, que se expressa pela diferença morfológica entre os sexos no que tange a massa corporal e formato / dimensões dos dentes caninos. Em suma, machos mais corpulentos e dotados de armas naturais podem representar um indício relevante de competição reprodutiva e territorial envolvendo comportamento agonístico (intimidação e violência interpessoal).

Num quadro de mudança climática, de vida arbórea, de insulamento territorial dos recursos naturais exploráveis (em manchas florestais ricas, cercadas por terreno aridificado) e de evidência de dimorfismo sexual, podemos sugerir que os proconsulídeos pertenceram ao conjunto dos primeiros primatas a viverem em sociedades estáveis. Considerando-se o alto custo reprodutivo e energético representado pela gestação intrauterina e pela lactação, o acesso a recursos nutricionais de alta qualidade e com oferta regular é uma exigência evolucionária de primeira ordem na etologia energética feminina. Esse é um elemento central de definição dos padrões de territorialidade primata, o que significa dizer que, na África oriental do Mioceno, onde as florestas mais ricas se tornavam cada vez mais isoladas por "mares" de savana, fêmeas de espécies arborícolas tenderiam a se concentrar nesses espaços segregados.

Enquanto a qualidade nutricional desses recursos concentrados foi grande, o contexto ambiental privilegiou a exploração conjunta por fêmeas aparentadas, que cooperavam para garantir acesso ao alimento a todas na comunidade genética matrilinear, e excluir outras fêmeas não aparentadas. Já para os machos, as exigências energéticas e reprodutivas são mínimas (inclusive no que tange à produção gamética), de modo que o principal desafio enfrentado tende a ser garantir o acesso às próprias fêmeas. É desse modo que a territorialidade é dada pelas estratégias femininas, já que os machos apenas acompanham os perfis

4 Oligoceno foi o período geológico compreendido entre 33,9 e 23,03 milhões de anos atrás.

5 Mioceno foi o período geológico compreendido entre 23,03 e 5,33 milhões de anos atrás.

de dispersão dos coletivos matrilineares. São eles que tenderão a migrar de seus grupos de origem quando atingida a maturidade, buscando enfrentar outros machos por oportunidades sexuais longe de sua comunidade genética.

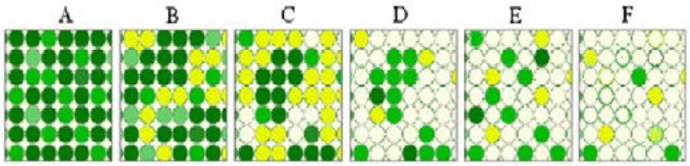


Figura 1. Distribuição de Recursos

Nesse modelo simplificado, quanto mais intensa é a cor verde, mais ricos são os nutrientes presentes. A) Recursos de alta qualidade, desconcentrados no espaço. Não há necessidade de concentração de fêmeas em patches; isso faz sobressair a vantagem, para as fêmeas, de forragearem sozinhas, e garantirem para si recursos nutricionais, evitando inclusive a competição com fêmeas aparentadas. Grupos instáveis se formam para segurança e observação. Machos dispersam de seus grupos de origem chegando à maturidade, bem como as fêmeas; B) Recursos de alta qualidade concentrados em manchas homogêneas, uniformes, e de grande extensão, acompanhados de recursos de qualidade mediana. Fêmeas aparentadas se concentram nessas manchas de alta qualidade, que são suficientes para que se alimentem juntas, desde que mantidas distantes as fêmeas não aparentadas. Prevalece a defesa do pool genético. Machos não aparentados ficam ao redor dessas fêmeas; C) Recursos de alta qualidade estão intensamente concentrados em manchas uniformes. Fêmeas aparentadas se reúnem; nessas condições, um macho sozinho é capaz de controlar o território de forrageamento desse grupo de fêmeas, e estabelecer um harém. Perfil presumido em Proconsul sp., Afropithecus sp., Gryphopithecus sp., Kenyapithecus sp. e Graecopithecus sp.; D) A qualidade dos recursos diminui, sendo mantido o padrão de distribuição. Grupos de parentesco feminino, se presentes, conduzem à disputa por energia entre fêmeas aparentadas, e por isso são desprivilegiados evolucionariamente. Fêmeas dispersam espacialmente em busca de patches de energia de média qualidade. Solidariedade entre as fêmeas decai. Oportunidade para haréns continuarem, dessa vez com patrilinearidade. Coincide com o padrão de sociabilidade de Gorilla sp.; E) Distribuição dos recursos se torna heterogênea, e desconcentrada no espaço. Grupos de parentesco feminino, já esgotados anteriormente, tornam-se ainda menos possíveis, bem como os haréns. Não é possível para um macho solitário dominar as fêmeas numa mesma área, já que se espalham para aproveitar os recursos de qualidade mediana, e os poucos de alta qualidade. Controle por um macho se torna impossível, mas patrilinearidade se preserva. Coalizões de machos aparentados para dominar as fêmeas dispersas. Coincide com o padrão de sociabilidade de P. troglodytes; F) Distribuição dos recursos por demais rarefeita condena a sociabilidade permanente. Entre Pongo sp., leva à dispersão feminina, à ocupação individual, pelas fêmeas, de zonas dotadas de recursos, e à formação de superterritórios masculinos

Esse quadro não dá conta da emergência da violência coalizacional intersocietária como traço comportamental, e é razoável assumir que ela não existiu no Mioceno Inicial entre os primatas, mesmo que consideremos a emergência da sociabilidade estável feminina. Entre os proconsulídeos, é provável que grupos de machos não aparentados e não cooperativos tenham sido tolerados pelos coletivos matrilineares na circunstância de o território e os recursos serem compatíveis com as exigências das estratégias reprodutivas e energéticas femininas. Se considerarmos, ainda, o avançar do agravamento climático e a possibilidade de que a oferta e a concentração dos recursos naturais tenham alcançado certa massa crítica (Barnosky, Kratz, 2006, p. 528), a ponto de tornar a presença de muitos machos adultos não aparentados um fardo para os coletivos matrilineares, poderíamos nesse caso postular a hipótese de desenvolvimento de uma sociabilidade harênica, na qual um macho dominante se torna capaz de estabelecer exclusividade de acesso às oportunidades reprodutivas representadas pelo grupo estável de fêmeas aparentadas.

Tal exclusividade de acesso ocorre através da viabilidade física de monitoramento sensorial do território por onde as fêmeas forrageiam – estando as florestas cada vez mais concentradas –, e com isso, pelo desenvolvimento de formas de contenção agonística das investidas de outros machos em busca de oportunidades reprodutivas. Para as fêmeas, a redução do número de organismos alheios à sua comunidade genética, e que sacam contra recursos energéticos decisivos – consubstanciada na redução do número de machos coabitantes – é algo que privilegia estrategicamente seu *fitness* reprodutivo em nível de grupo.

Haréns eram uma forma de sociabilidade verossímil entre os proconsulídeos, e na circunstância de terem se dado, devem ter provocado alto grau de tensão e violência interpessoal masculina – conclusão que encontra suporte na condição dimórfica comum nessa família. Nesses termos, a disputa agonística entre machos migrantes, não aparentados e não cooperativos, sinalizava negativamente em relação às coalizões masculinas patrilineares, matéria-prima da guerra e dos freios etológicos ligados à resolução de conflitos internos (Cameron, Groves, 2004, pp. 38-40; Foley, 2008, p. 220-227; Ladeia, Ferreira, 2015, p. 75; Nordhausen, Oliveira Filho, 2015, pp. 36-37; Wrangham, Peterson, 1996, pp. 131; 174-175).

3. *Afropithecus*: matriarcados estáveis e a savana

Tendo provavelmente divergido de uma ou mais espécies proconsulídeas por volta de 18 milhões de anos atrás, outro conjunto de espécies primatas, organizadas na família Afropithecidae, potencialmente levou adiante as inovações comportamentais desenvolvidas por seus ancestrais, ainda que em contexto ecológico um tanto distinto. Essa herança se torna relevante quando consideramos que um afropitecídeo foi o mais provável ancestral de todos os hominídeos, o que inclui, *lato sensu*, os seres humanos.

Em linhas gerais, a família comportou espécies morfologicamente bastante distintas do perfil de seus ancestrais proconsulídeos – com quem coexistiram, ressalte-se. Foram prováveis braquiadores,[6] o que significa que sua estratégia motora consistia em locomoção suspensória em ambiente florestal, tal como fazem gibões, chimpanzés e outros símios ainda existentes. Seu plano corporal invocava postura mais vertical que a demonstrada pelos proconsulídeos. A massa corporal média presumida dos afropitecídeos é maior do que a considerada compatível com os hábitos de primatas essencialmente arborícolas, o que faz invocar a possibilidade de vida semiterrestre. Considerando o plano corporal adaptado à braquiação, sua estratégia motora terrestre deve ter se baseado na nodopedalia (movimento com apoio nos membros posteriores e nos nós dos dedos dos membros anteriores), tal como fazem gorilas e chimpanzés, aspecto que traria consequências futuras quanto ao exercício da territorialidade. Assumindo a possibilidade de comportamento semiterrestre, somada à evidência de arquitetura facial robusta, de esmalte dentário espesso e de potentes dentes mastigadores, fica sugerido que, ao contrário de seus ancestrais, os afropitecídeos ocuparam habitats marginais às florestas úmidas, com regulares incursões aos espaços savanizados, de onde podiam obter alimentos reserva (*fallback foods*) em sintonia com as flutuações ambientais. Oferecendo a savana recursos notadamente secos, abrasivos e rígidos, que exigem certa preparação mastigatória, os afropitecídeos estavam adaptados motora e odontomorfologicamente para a exploração desse nicho aberto pela aridificação.

A ampliação das oportunidades alimentares para essas espécies – na medida em que eram capazes não só da exploração de recursos florestais como também daqueles oriundos das savanas, bosques e matagais – alterou a relação entre sociedade e espaço para esses primatas. Devemos considerar que a espacialidade dos recursos é dinâmica, e é função da capacidade de

6 Braquiação é a estratégia de locomoção suspensória na qual o animal se move dependurado entre os galhos das árvores, na posição vertical.

processamento / eficiência termorregulatória[7] das estratégias motoras correntes. Assim sendo, se para os antepassados proconsulídeos, arbóreos e de dentição grácil, os habitats aridificados significaram barreiras intransponíveis – o que conduziu suas sociedades à concentração em territórios segregados, com as consequências em termos de organização social acima sugeridas –, para os mais robustos afropitecinos, a savana, em maior ou menor medida, foi também espaço de forrageamento e exploração, o que teria contribuído para reduzir os efeitos da concentração demográfica sobre suas estratégias sociais.

Dessa forma, talvez devêssemos supor que, na raiz evolucionária de todos os hominídeos, a sociabilidade estável – um dos pilares primitivos da guerra e da ética – passava a ser contraindicada pela desconcentração das oportunidades nutricionais no território, algo que conduziria à dispersão espacial das fêmeas, à contraproducência da manutenção de coletivos matrilineares permanentes e à inviabilidade da estratégia de exclusividade reprodutora masculina, expressa pelo comportamento de harém. Todos esses fatos tenderiam, se verdadeiros em alguma medida, a remeter nossos ancestrais novamente a um perfil de sociabilidade comum aos parâmetros eocênicos – as agregações multissexuais instáveis – ainda praticado por muitas espécies de macacos, especialmente platirrinos[8] do Novo Mundo.

A despeito dessas possibilidades, a hipótese filogenética de transmissão da sociabilidade estável entre os hominídeos não deve ser descartada prematuramente, já que dispomos de bons exemplos de primatas também capazes de amplas radiações territoriais, mas que preservam formas alternativas de comportamento harênico e de coletivos femininos matrilineares estáveis, como é o caso de algumas espécies de papioninos (especialmente babuínos). Acontece que entre estes últimos, os haréns se preservaram (potencialmente a partir de uma herança proconsulídea, tal como em *Afropithecus* spp.), mas, dada a ampliada espacialidade relativa dos recursos naturais, tornou-se possível a constituição de *tropas*, verdadeiras confederações de haréns, formadas por muitas unidades, cada qual composta por fêmeas aparentadas e um macho dominante, eventualmente acompanhados por poucos machos subalternos.

Esse padrão etológico entre os papioninos representa uma face da flexibilidade morfocomportamental presente em graus diferenciados em todas as espécies, o que significa dizer que existem bandas de acomodação nas quais os aspectos herdados podem ser replicados a despeito da transformação do meio. Então não podemos descartar a tese de que os afropitecídeos tenham preservado padrões sociais etológicos herdados de espécies anteriores, especialmente se considerarmos que eles persistem entre primatas africanos extantes, que descendem dos primeiros (Cameron, Groves, 2004, p. 39; Barnosky, Kraatz, 2007, p. 525; Foley, 2008, pp. 150-151; 178-179; 183-184; Pampush et al., 2013, p. 222; Wrangham, Peterson, 1996, pp. 56-59).

4. O Mioceno Médio e as migrações no eixo afro-asiático

A chegada do Mioceno Médio, por volta de 15 milhões de anos atrás, coincidiu com a expansão das calotas polares e com maior retração do nível dos mares, em um pico de intensidade no já corrente processo de resfriamento e aridificação globais. Nessas condições, as porções setentrionais da Eurásia se tornavam inóspitas para primatas em geral, ao passo que se formava uma zona biogeográfica de clima mais ameno e de relativa homogeneidade no Saara, na África Centro-Oriental, na Europa meridional e no Levante. São conhecidos os movimentos de muitas espécies africanas endêmicas em direção ao norte, passando pelas costas mediterrânicas, e dessas mi-

7 Primatas nodopedálicos em deslocamento terrestre por campo aberto (como em espaço savanizado, com poucas árvores), expõem maior superfície corporal (cabeça, ombros, dorso) à incidência direta dos raios solares do que primatas bípedes, como os humanos modernos.

8 Platirrinos compreendem uma parvordem de primatas que reúne todos os macacos das Américas. Ao contrário dos catarrinos do Velho Mundo (como os humanos, babuínos, gorilas, chimpanzés, etc.), que possuem narinas protuberantes e voltadas para baixo, os platirrinos possuem narinas achatadas e voltadas para os lados. Algumas de suas espécies possuem cauda preênsil, ausente em todos os catarrinos.

grações também participaram os antropoides. A partir das radiações para fora da África, processos especiativos foram se sucedendo entre os grandes símios na Europa e na Ásia menor, sem que os aspectos morfológicos e, possivelmente, comportamentais, tenham se afastado significativamente daqueles herdados dos antepassados afropitecídeos. Apesar disso, o período de dez a sete milhões de anos atrás foi marcado por redução na diversidade dos grandes primatas, representando para eles, talvez, um primeiro gargalo evolucionário significativo. As extinções que marcam o período denotavam que o portfólio comportamental e morfológico montado sobre a tríade dentição robusta/locomoção semiterrestre/exploração ocasional da savana tinha já encontrado seu limite de acomodação. Foi nesse contexto, então, que *Graecopithecus freybergi* ou algum outro símio dele derivado alcançou as costas mediterrânicas mais uma vez, escapando do agravamento climático no norte. A relevância dessa espécie está no fato de que talvez tenha sido o pivô das radiações dos grandes antropoides europeus em direção ao eixo afro-asiático, onde o meio ambiente iria submeter à prova final as estratégias sociais dos recém-surgidos hominídeos (Cameron, Groves, 2004, pp. 41-42; 55-57; Ladeia, Ferreira, 2015, pp. 76-77).

5. Os pongíneos na Ásia: portas fechadas para a guerra

As migrações pela Ásia meridional até o Extremo Oriente, na condição de vetores de expansão latitudinal, foram marcadas pela incidência de condições climáticas razoavelmente constantes, ainda que diretamente agravadas pelo clima glacial das porções setentrionais da Eurásia. Essa homogeneidade representou um conjunto de desafios evolucionários compatíveis, da Anatólia ao Sudeste Asiático, algo igualmente sugerido pela similaridade morfológica entre os grandes primatas asiáticos extintos e *Pongo* sp. (as duas espécies de orangotangos), que vivem hoje somente nas ilhas de Sumatra e Bornéu. Compartilham de perfil odontomorfológico herdado dos antepassados europeus que perfizeram a rota em direção às zonas meridionais – esmalte dentário espesso, molares robustos em comparação com o restante da dentição –, e, como já dito, esse perfil provavelmente já não mais garantia a sobrevivência contra a sazonalidade e a aridificação planetária. Desse modo, os pongíneos (os grandes símios asiáticos) devem ter se adaptado à extrema rarefação das florestas tropicais, ao empobrecimento da oferta de nutrientes e à desconcentração espacial radical dos recursos na Ásia meridional (bosques cercados de planícies), através de padrões de comportamento que identificamos, hoje, nas últimas espécies extantes da subfamília.

É provável que os coletivos femininos matriarcais estáveis tenham se tornado inviáveis, já que, com recursos escassos e muito fragmentados, ou as fêmeas precisariam se espalhar no espaço em busca de alimento, ou deveriam competir entre si num território restrito e incapaz de comportar as exigências calóricas de todas. Isso significaria inevitável prejuízo em termos de *fitness* reprodutivo para um número razoável de fêmeas consanguíneas, ao passo que, na circunstância de cada uma delas migrar para zonas diferentes, emergiria a possibilidade de sucesso individual sem ter por consequência o fracasso de suas parentes. É o que fazem os orangotangos: entre eles colapsa a cooperação feminina matrilinear, já que cada fêmea com seus filhotes impúberes se fixa em determinado núcleo arbóreo, apartado dos demais, com recursos suficientes para manter essa unidade familiar estável. A sociabilidade grupal é igualmente dissolvida, e os consórcios entre machos e fêmeas se tornam temporários, criando apenas uma frouxa rede de relacionamentos por um território amplíssimo.

Para os machos, a territorialidade feminina previne a estratégia de exclusividade sexual tradicional (harém), dada a incapacidade prática da prevenção do acesso a machos competidores. Não obstante, a etologia da exclusividade sexual mantém-se relativamente viva, na medida em que os machos dominantes circulam terrestrialmente (as fêmeas raramente deixam o topo das árvores) por vários núcleos espaciais femininos (criando um superterritório), tentando assim garantir seus privilégios reprodutivos e eliminar competidores, formando algo como um harém lasso. O grau de incerteza sobre a paternidade gerado pelo modo de ocupação das fêmeas no espaço conduz a alto nível de pressão competitiva entre machos adultos, que se expressa por notório dimorfismo sexual (marcadores fenotípicos como a massa corporal e bolsas de gordura nas laterais da face, indicadoras de maturidade sexual), comportamento de vocalização (sinalizando a presença no território) e intenso nível de violência interpessoal. Assim, os pongíneos representaram um beco sem saída evolucionário no que tange à etologia da guerra e da ética: está prevenida entre eles a formação de coletivos matrilineares, bem como de qualquer grupo social cooperativo masculino, patrilinear ou não. O alto grau de agonismo,[9] e de violência física em particular, não advoga favoravelmente para um comportamento guerreiro, nem gera contexto que torne cruciais determinados instrumentos etológicos de resolução de conflitos (Cameron, Groves, 2004, 75-77; Foley, 2008, p. 218; Nordhausen, Oliveira Filho, 2015, p. 29; Wrangham, Peterson, 1996, pp. 133-134).

6. Gorilíneos na África: patrilinearidade não cooperativa proscreve a violência coalizacional intersocietária

Os pongíneos não interessam diretamente ao problema da violência coalizacional intersocietária e da ética entre os humanos modernos, já que consistiram em linhagem divergente daquela que resultaria nos grandes antropoides africanos, tendo seu último ancestral comum vivido por volta de 12 milhões de anos atrás, justamente no período de radiação para fora da Europa, e em direção ao eixo afro-asiático. São os hominídeos que se especiaram no caminho para a África que nos interessam em particular, visto que entre eles a etologia da guerra e da ética daria mais um passo decisivo, preservando a hipótese filogenética de transmissão desse portfólio comportamental.

As mais antigas espécies alocadas sob o manto da subfamília Gorillinae surgiram aproximadamente por volta de dez milhões de anos atrás, e estão entre aquelas que apareceram em solo africano durante o Mioceno Tardio. A migração de antropoides europeus para o sul, em sentido longitudinal, significou o movimento de espécies parcialmente adaptadas a condições de temperatura e aridez mais severas do que aquelas encontradas na África Centro-Oriental. Assim, em ambiente menos rigoroso se comparado à faixa latitudinal que vai da Ásia mediterrânica à Península Malaia, entre os gorilíneos e seus antepassados diretos africanos preservou-se algum tipo de sociabilidade grupal permanente – e se tomarmos os gorilas contemporâneos por referência, uma sociabilidade harênica.

Apesar de encontrarem biomas menos afetados pela mudança climática, também para os gorilíneos a estratégia primitiva, calcada na dependência de dentição robusta, da nodopedalia terrestre e da exploração oportunista da savana, parece ter sido prescrita. Seu perfil odontomorfológico difere significativamente daquele de seus ancestrais europeus, indicando especialização dietária nos espaços florestais em franco retrocesso: temos então uma dentição mais grácil, com esmalte de fina espessura, demonstrando que seguia sendo uma opção com retornos de curto prazo a "aposta" em nichos em desaparecimento. Com uma espacialidade dos recursos naturais menos rarefeita que na Ásia meridional, ainda que com qualidade nutricional inferior àquela desfrutada pelos antiquíssimos primatas do Eoceno, os haréns devem ter se mantido até o surgimento de *Gorilla* sp., ainda que não a matrilinearidade feminina cooperativa estável. Com a qualidade e quantidade dos recursos diminuída, ainda que seguissem distribuídos com alguma uniformidade e concentração, a etologia energética feminina foi diretamente ameaçada – como, de forma análoga, entre os pongíneos.

A sociabilidade feminina seguia sendo vantajosa – e já conhecemos as vantagens da gregariedade desde os tempos eocênicos –, mas não mais entre fêmeas aparentadas. Então, entre os gorilíneos, são as fêmeas que principalmente migram para outros grupos ao atingirem a maturidade – isso significa, do fim ao cabo, que a cooperação consanguínea se inviabiliza, e a gregariedade se dá entre indivíduos

9 Caracteriza-se como agonista toda forma de comportamento de conflito envolvendo violência física ou intimidação.

não aparentados. O colapso dessa cooperação entre fêmeas aparentadas significa que a capacidade de autoproteção feminina também colapsa, abrindo espaço para o avanço da agenda reprodutora masculina no âmbito dessas sociedades harênicas. O exercício do poder e da exclusividade sexual por um macho dominante não é mais incompatível com a manutenção de um ou mais de seus filhos adultos no grupo (embora também eles tendam a migrar e disputar privilégios sexuais em outros grupos), sacando contra recursos nutricionais já escassos, e assim, desprestigiando a etologia energética feminina (sendo as fêmeas não aparentadas, cada descendente do macho dominante forma comunidade genética com somente uma delas, em exclusão das demais), o que reforça o imperativo de migração.

Não obstante a formação de vínculos patrilineares entre os grandes primatas africanos desde, provavelmente, dez milhões de anos atrás, a ela não se segue a geração de laços cooperativos entre esses aparentados. A exclusividade sexual do macho dominante, típica dos regimes harênicos, reproduz-se nos gorilíneos, estando submissos todos os demais machos a essa ordem política. Trata-se, assim, de dominância em sentido estrito, e os filhos de um *silverback*[10] somente terão acesso ao seu harém após sua morte. A dominação também se estende às fêmeas, cujos laços de solidariedade foram rompidos com o fim da matrilinearidade: seus conflitos, envolvendo inclusive o acesso à energia, são reprimidos. Existe, por assim dizer, um alto grau de paz interna e de submissão ao poder constituído, e não se configura, sob nenhum aspecto, uma etologia da rebelião, como viria a se tornar comum milhões de anos depois, no ancestral comum entre humanos e chimpanzés. Não há contestação à ordem, nem tentativas de tomada de poder por parte dos membros do grupo; as ameaças à dominância geralmente ocorrem a partir da chegada de jovens migrantes, que tentam usurpar o harém de um *silverback* estabelecido. A exclusividade sexual, como é de praxe, dá margem a alto grau de violência interpessoal masculina e de dimorfismo sexual. Mais uma vez, a despeito de duas das matrizes da etologia da guerra e da ética se apresentarem – a sociabilidade estável e a patrilinearidade –, o regime harênico e a falta de cooperação masculina intragrupo proscrevem-nas (Foley, 2008, p. 224; Pampush et al., 2013, p. 217; Wrangham, Peterson, 1996, pp. 147-149).

7. O Mioceno Tardio e o último ancestral comum entre humanos e chimpanzés: hierarquias complexas e patrilinearidade cooperativa masculina

O planeta seguia sua trajetória em direção à glaciação, e alguns milhões de anos após a emergência dos gorilíneos, a disrupção climática alcançava novo limiar crítico, fazendo com que grandes primatas altamente conservadores, como os gorilas (em termos de ocupação de nichos ecológicos), migrassem coletivamente acompanhando a retração das zonas florestais – no longuíssimo tempo evolucionário, acrescente-se –, enquanto populações marginais se adaptavam à mudança em processos especiativos. A espacialidade relativa dos recursos para populações antropoides de dentição grácil e nodopedálicas não se esgarçou como na Ásia meridional, mas ainda assim, o perfil de concentração teoricamente viável para a preservação das sociedades harênicas parecia se replicar com dificuldade. Com recursos alimentares cada vez mais rarefeitos no território, essas populações viventes nas regiões limítrofes entre a savana e a floresta precisavam cada vez mais se disseminar pelo território, afastando-se do núcleo organizacional dos grupos; e como sabemos, tal fato, em última instância, significa dizer que as fêmeas não aparentadas, perseguindo sua agenda energética, forrageiam cada vez mais distantes umas das outras. Assim como ocorreria entre os pongíneos em sua migração latitudinal pelo sul da Ásia, a espacialidade feminina formava um perímetro incompatível com a possibilidade concreta de vigilância de um macho altamente dimórfico e disposto a manter sua exclusividade sexual em prejuízo de todos os demais.

10 A designação *silverback* (dorso prateado) está consolidada na literatura para referir-se ao macho dominante em sociedades de gorilas. O termo deriva da coloração clara que a pelagem desses primatas adquire com a consolidação da maturidade sexual.

Lembremo-nos, ainda, que esses machos deviam herdar, desde pelo menos dez milhões de anos atrás, comportamento de formação de linhagens patrilineares e fundamentalmente patrilocais. O agravamento climático colocava sobre a balança as vantagens, para os machos, em termos de defesa de sua comunidade genética (a patrilinearidade), de um lado, e o exercício de uma agenda reprodutiva individual em contexto de inviabilidade do comportamento harênico, de outro. No caso dos pongíneos asiáticos, desprovidos da herança etológica da patrilinearidade, não havia dilema: a busca pela exclusividade sexual manteve-se, apesar das circunstâncias ambientais, com alto grau de comportamento agonístico entre machos não aparentados. Já no caso desses grandes primatas africanos, entre os quais provavelmente viveu o último ancestral comum humanos e chimpanzés, a herança da patrilinearidade fez as peças do tabuleiro evolucionário moverem-se em direção inédita.

Em lugar da dissolução desses laços entre machos aparentados, ao contrário, eles se tornam mais firmes, com o desenvolvimento de complexa cooperação masculina, condição etológica rara. Diante de perímetros de vigilância incompatíveis para um "panóptico" individual, provocados pela ocupação espacial feminina (no ato de forrageamento), rompem-se os haréns e dilui-se a perceptividade do lócus de dominância do macho alfa. Evitando o conflito sexual fratricida por oportunidades reprodutivas, esses coletivos patrilineares masculinos se organizam em complexas hierarquias de *status*. Desaparecida a dominância estrita, o acasalamento torna-se uma questão poliginândrica.[11] Sinalizando ritualisticamente o reconhecimento do lócus de prestígio de cada um de seus pares, cooperam eficientemente para o controle do território por onde as fêmeas do grupo se espalham, com o objetivo de negar a machos não aparentados, viventes em outros grupos sociais, acesso a elas. Aparece com nitidez a dimensão de *seleção de grupo*, multinível.

As agendas reprodutivas individuais são relativizadas pela dimensão coletiva e societária, de defesa da comunidade genética masculina. Com o comportamento agonístico entre machos abrandado na dimensão intragrupo, o que inclui a violência interpessoal, com a redução dos níveis de dimorfismo sexual e com o surgimento de mecanismos etológicos complexos de gerenciamento de conflitos, abre-se o espaço igualmente raro para a projeção da violência e do poder para o nível extragrupo, configurando o fenômeno da violência coalizacional intersocietária (Aureli et al., 2008, pp. 629-630; Foley, 2008, p. 230; Wrangham, Peterson, 1996, p. 52).

8. Sociabilidade pós-harênica, patrilinearidade cooperativa e a ética intuitiva

Os primatas contam, em graus variados, com uma eficiente inteligência geral para a resolução de problemas. Isso implica que, para além de simples conteúdos comportamentais herdados, em seu portfólio etológico constam mecanismos de aprendizado a partir da interação com o ambiente, partindo de regras genéricas, envolvendo tentativa e erro. Durante a maior parte da história dessas espécies, a acomodação aos desafios evolucionários parece ter sido possível empregando-se essa forma de inteligência, de menor custo. O contexto ambiental para o desenvolvimento de tipos mais especializados de cognição, energeticamente mais dispendiosos, parece ter emergido lentamente desde dez milhões de anos atrás, com as migrações de volta para a África e em direção à Ásia; e ter se tornado efetivamente visível com a modularização da cognição social, por volta de seis milhões de anos atrás, com a emergência do último ancestral comum entre humanos e chimpanzés.

A cognição social altamente especializada emerge como mecanismo de acomodação entre a competição e a cooperação, entre as agendas reprodutivas masculinas individuais e a patrilinearidade estável pós-harênica. Num contexto de disputa por oportunidades sexuais entre machos aparentados, a luta fratricida é reduzida a níveis evolucionariamente irrelevantes

11 Em regimes poliginândricos, machos e fêmeas selecionam parceiros sexuais ocasionais, sem estabelecimento de vínculos estáveis. Naturalmente, a poliginandria não implica oportunidades reprodutivas iguais, necessariamente. São favorecidos machos em situação hierárquica superior, mas, ao mesmo tempo, não há exclusividade reprodutiva, como no caso dos haréns.

por meio de uma complexa capacidade de análise do lócus de poder de cada membro do grupo em suas relações com os demais, e de formulação de hipóteses sobre as possibilidades de ascensão e de queda na pirâmide social de todos os agentes envolvidos na rede de relacionamentos, a partir das quais um indivíduo possa traçar suas estratégias de preservação ou de conquista de *status*. O reconhecimento de escalões de prestígio e poder, a aceitação provisória da própria condição social, e o desenho de estratégias para a contestação da hierarquia em benefício próprio e de seus aliados surgem como diretrizes etológicas fundamentais entre chimpanzés, e presumidamente estiveram presentes também em seu último ancestral comum com os humanos.

Chimpanzés e humanos especiaram-se a partir de sociedades estáveis, com patrilinearidade e patrilocalidade masculina, marcadas pelo alto grau de incerteza quanto aos privilégios e limites nas relações sociais internas ao grupo. A relativa simplicidade das hierarquias harênicas dos antepassados, nas quais o espaço de dominância e de privilégios sexuais era nítido, monocrático, e a subalternidade era condição comum a todos os demais membros do grupo, foi substituída por uma espécie de caos sistêmico em nível intragrupo, no qual múltiplos escalonamentos na pirâmide hierárquica emergem, e a luta pela ascensão social (masculina) se generaliza. Esse seria um contexto propício para a dissolução dos laços permanentes – nesse caso, a sobrevivência da patrilinearidade acabou sendo garantida pelo surgimento dessa cognição social modular, de alto custo, que coincide com a ampliação da alometria cerebral entre os panínеos (as duas espécies de chimpanzés conhecidas) e os hominíneos (todos os antropoides bípedes surgidos após a divergência com os chimpanzés), se comparados aos antropoides mais antigos na árvore evolutiva.

A emergência de módulos mentais dedicados ao gerenciamento das relações sociais significa dizer que os mecanismos cognitivos estereotipados e genéricos produzidos a baixo custo pela inteligência geral tornaram-se insuficientes para gerar respostas eficazes num contexto de excesso de informações e de partes moventes na mecânica social. O foco da modularidade não está em seus conteúdos inatos, mas na capacidade de *formulação de hipóteses testáveis* sobre o comportamento de terceiros, envolvendo ou não a presença do próprio observador, em uma extensão dos complexos cognitivos ligados à chamada teoria da mente, presente em diferentes graus de complexidade por toda família dos Primatas. O processo de avaliação do estado mental de terceiros assume como modelo as reações que o próprio sujeito pensante esperaria de si próprio estando ele, hipoteticamente, na situação do outro, o que envolve um razoável grau de desenvolvimento de habilidades empáticas. A extrapolação das hipóteses precisa, ainda, ter seus resultados adaptados às características de temperamento individual do analisando (que devem, por definição, ser previamente conhecidas) e ao campo do contingencial (leia-se, das circunstâncias da ação). Nesse caso, o emprego de regras de aprendizado genéricas, padronizadas, para a tomada de decisão social estratégica nas condições de complexidade presentes nestas sociedades antropoides pós-harênicas, resultaria em grande chance de erro.

É de difícil sustentação a ideia de que nas linhagens que provêm do último ancestral comum entre humanos e chimpanzés, a sociabilidade seja produto do aprendizado social. Chimpanzés podem ser ensinados por humanos a executar tarefas em cativeiro que, em seus habitats naturais, não seriam apreendidas, visto que não cumprem qualquer papel evolucionário relevante (linguagem de sinais, produção de ferramentas líticas, etc.). Para tais atividades, e como para todas as demais que executam, excetuando os jogos de *status*, chimpanzés empregam sua inteligência geral, que funciona como uma espécie de ferramenta multiuso de aprendizagem, produzindo resultados simples após período de tentativas e erros, mas a baixo custo energético. Já no que tange ao comportamento social modularizado, não há nada efetivamente que se possa ensinar a um chimpanzé, ou que devam ensinar uns aos outros: mesmo indivíduos nascidos em cativeiro desenvolvem intuitivamente, na idade certa, as competências sociais necessárias para o intenso "jogo maquiavélico" das disputas de *status*, o que demonstra sua inatidade.

Em suma, nas linhagens de homens e chimpanzés, módulos mentais dedicados, capazes de compreender o funcionamento das hierarquias sociais, e de formular estratégias de posicionamento nessa pirâmide, emergem com a idade, tal como os dentes definitivos.

A modularidade da cognição social parece permitir que os chimpanzés (e, presumidamente, seu último ancestral comum com os humanos) desenvolvam consciência de si e dos membros do grupo (algo sugerido por resultados positivos em teste de autorreconhecimento em espelhos). Entretanto, circunscrita ao âmbito da etologia, está longe de equivaler ao *self*, à consciência holística transmodular e transdominial que só recentemente emerge na história evolutiva de *H. sapiens*. A inteligência social modular em chimpanzés parece isolada da inteligência geral, incapaz de interagir de forma fluida com outros domínios cognitivos não modularizados, de modo que esses grandes símios se mostram conscientes de si e dos outros somente *enquanto atores sociais*, e na ocasião do exercício das relações sociais.

Não existem evidências substantivas do emprego da cultura material – ligada aos mecanismos gerais de cognição técnica, imersos na inteligência geral – para a obtenção de vantagens nas disputas de *status*. Não há, ainda, qualquer dimensão simbólica da cultura material que seja instrumentalizada de modo a transmitir informação social ao coletivo, que sinalize acerca do lócus hierárquico ocupado por determinado indivíduo, nem que sirva para dissimular a ocupação de um escalão inferior nos esquemas de estratificação. Com a mente modular social incapaz de acessar outros domínios cognitivos e colocá-los a seu serviço, chimpanzés não parecem capazes de simulações mentais complexas a respeito de questões ligadas ao forrageamento e à produção de ferramentas envolvendo outros de seus coespecíficos. Alheia ao âmbito do social, a inteligência geral opera em domínios inconscientes de si próprios, incapazes de produzir percepções mentais e autorrepresentações daquilo que sabem. Essa condição, embora altamente derivada quando comparamos chimpanzés e outros primatas, é primitiva diante da transdominialidade cognitiva dos humanos modernos.

Não obstante sua insularidade, a mente modular social permitiu a fixação de padrões inatos e de normas ritualísticas na luta pelo poder intragrupo, algo que, por sua vez, se traduz na disputa por vantagens no acesso a oportunidades reprodutivas (ressaltando-se a inexistência de uma condição etológica de busca por exclusividade sexual). A partir de observação do comportamento de chimpanzés em habitat natural, sabemos que esses confrontos por dominância entre dois machos adultos podem durar muitos meses, e serem marcados por demonstrações intensas de agonismo. É comum que o macho desafiante se recuse a realizar rituais de submissão ao dominante, rituais esses que são regularmente cumpridos pelos demais membros do grupo como forma de reafirmação dos laços de lealdade, de reconhecimento de seu lugar na hierarquia, e da estabilidade do corpo social.

Esses rituais envolvem determinadas posturas corporais e gestos, como dar as costas ao líder, curvar-se ou abaixar-se diante dele, ou ainda, demonstrar aquilo que alguns primatólogos designam por "sorriso assustado". Sendo parte da sinalização de poder o ato de estender o braço e tocar com a mão o ombro de outro chimpanzé de *status* inferior, o desafiante, durante sua campanha pelo poder, tende a não permitir que o macho dominante o realize consigo. Normalmente, essas demonstrações de intimidação, agressividade e poder são observadas atentamente por todos os membros do grupo, que, ao longo do tempo, tendem a se posicionar na disputa, em apoio a cada um dos contendores.

Durante o processo de estranhamento entre facções, as coalizões que se formam cotidianamente para tarefas específicas – forrageamento, *grooming,*[12] etc. – tendem a se tornar mais voláteis. Tanto o macho dominante quanto seu desafiante buscam intimidar as fêmeas do grupo e, para tal, formam alianças com outros machos subalternos. Ao intimidá-las, o que buscam os disputantes é o apoio político das

12 Também chamada de catação, o *grooming* é um importante ritual social entre os primatas. Seu objetivo primário é o de remoção de parasitas e detritos da pelagem, em benefício da higiene. Contudo, o papel do *grooming* enquanto ato social transcende essa dimensão, funcionando como importante instrumento de reforço dos laços afetivos.

próprias fêmeas, sem o qual a tomada / conservação do poder não se conclui. A luta por suporte também é reforçada pela ampliação do tempo social gasto com cada fêmea e com seus filhotes, através da prática do *grooming*. Nesse caso, podem entrar em cena machos subalternos aliados, que afastam as fêmeas partidárias do adversário, para que elas não interfiram na estratégia de socialização e conquista de apoio. Esses machos inferiores buscam, com isso, galgar degraus na hierarquia a partir da vitória de seu candidato, o que geralmente lhes garante um acrescido grau de vantagens reprodutivas. Então, após intensa dedicação na luta pela dominância, os membros do grupo tendem a convergir para o apoio a um dos competidores, e ao isolamento social do outro, encerrando o processo. A partir daí, as demonstrações de comportamento agonístico por parte do macho vitorioso tendem a se reduzir; o líder assume postura conciliadora e pacificadora, mediando conflitos entre as fêmeas e auxiliando machos mais fracos ou com menos prestígio contra adversários mais fortes. Em algum tempo o processo de contestação da hierarquia recomeça, o que não raro envolve radicais recomposições de alianças.

É o alto grau de incerteza provocado pela expansão do tamanho populacional dos grupos sociais, pela ruptura da dominância harênica e pela preservação da patrilinearidade masculina – com a possibilidade de disputa letal fratricida –, o que leva, entre os grandes primatas africanos, ao desenvolvimento de uma ética intuitiva. Ela, funcionando como verdadeiro freio etológico prossocial, determina as normas e procedimentos da disputa interna pelo poder, reduz o grau de violência interpessoal masculina letal (embora não a elimine) e estabelece quando e como a luta se dá por encerrada, até que o ciclo se reinicie. Ao contrário das culturas chimpanzés, em que determinados grupos sociais desenvolvem práticas e comportamentos transmitidos socialmente, pelo aprendizado observacional, e que não se repetem em quaisquer outros grupos, um núcleo comum de parâmetros de sociabilidade e de resolução de conflitos interna se reproduz de forma inata em todos os grupos de chimpanzés, na natureza ou em cativeiro .(Aureli et al., 2008, pp. 632; 636-637; Bauernfeind et al., 2013, pp. 263-264; 271-273; Foley, 2008, pp. 207-210; Mithen, 2002, pp. 67-71; 102-111; 126-131; 139-142; Nordhausen, Oliveira Filho, 2015, 36-38; Wrangham, Peterson, 1996, pp. 128; 143-144; 186).

9. O conflito intersocietário e o esgotamento da cognição social

Se o desenvolvimento da cognição social resulta em padrões éticos inatos nas relações intragrupo, o espaço das relações intersocietárias é justamente o de ausência não só de freios etológicos, como também da ação dos mecanismos cognitivos de gerenciamento do jogo hierárquico. Entre os chimpanzés comuns (*Pan troglodytes*) e, provavelmente, mais uma vez, no último ancestral comum com os humanos modernos, a comunicação entre grupos sociais resume-se à violência coalizacional, cujo objetivo é o de eliminação dos machos estrangeiros, de abdução das fêmeas férteis e de desarticulação das comunidades inimigas. Não existem quaisquer mecanismos cognitivos de pacificação ou de contenção do conflito letal entre essas sociedades.

Está na origem da violência coalizacional intersocietária a fragmentação dos grupos sociais provocada, ocasionalmente, pelas lutas por dominância em âmbito interno. Determinados limites socioambientais se impõem para a coesão dos grupos de chimpanzés, o que envolve um equilíbrio delicado entre rarefação espacial dos recursos naturais e o contingente populacional ascendente. Quando esses limites são ultrapassados e o ciclo de disputa por poder se reinicia, o fracionamento da macrounidade social pode ser o resultado. O primeiro indício de que a luta política provocará secessão está na segmentação do grupo em facções de relacionamento: a tendência à alternância de parceiros no forrageamento e no *grooming* dá lugar a escolhas mais limitadas e repetidas, com indivíduos buscando alimento e reforçando seus laços sociais mais frequentemente com determinados coespecíficos que com outros.

Um importante sinal de secessão em progresso está no alinhamento das facções a seus líderes, sem defecções, ao contrário do que ocorre normalmente nas lutas por dominância, nas quais um dos competidores vai sendo abandonado por seus apoiadores e progressivamente isolado por uma maioria crescente. A fratura social vai se tornando visível na própria configuração da espacialidade grupal: os indivíduos de cada facção, ainda que coabitem as mesmas localidades, tenderão a se arranchar em campos opostos. O processo segue com a dissociação entre as áreas de forrageamento de cada uma das facções, circunstância a partir da qual a cisão entre dois grupos sociais pode ser considerada completa, em termos formais.

Uma vez separadas as duas unidades sociais com seus respectivos machos dominantes e pirâmides hierárquicas, logo pequenos subgrupos de caráter temporário se constituirão para a realização de atividades cotidianas, mas também para o exercício da violência coalizacional intersocietária. Provocados etologicamente pelo macho dominante (na maior parte dos casos) por meio de gestos e vocalizações, com finalidade de se organizarem para a luta, companhias de machos adultos podem ser formadas e partir em marcha para o território do grupo inimigo, muitas vezes acompanhados de algumas fêmeas jovens e sem filhotes. Não se trata sobremaneira da organização de patrulhas defensivas, reativas, nem do exercício da violência como subproduto ocasional da prática do forrageamento, por exemplo. Deixando para trás oportunidades alimentares, as marchas ocorrem com a finalidade única e exclusiva de levar a violência letal ao "outro". Durante as incursões, a detecção de sinais sensíveis da presença do oponente (sons, em particular) provocam reações de ansiedade, controladas por meio de gestos que asseguram confiança e lealdade (toques, abraços).

Como nas condições ecológicas em que viveram chimpanzés e seu último ancestral comum com os humanos, a rarefação dos recursos no espaço leva à fissão temporária dos grupos permanentes para a prática do forrageamento, e isso cria a oportunidade aguardada por uma companhia agressora: apanhar um macho inimigo solitário e incauto, enquanto se alimenta. Diferentemente das guerras entre unidades políticas estatais em *H. sapiens* (mas não essencialmente diferente dos conflitos entre unidades políticas humanas não estatais), os choques entre chimpanzés são necessariamente assimétricos. Em caso de uma avaliação errada das circunstâncias, que promova o encontro entre um companhia e um grupo numericamente equivalente de inimigos, os ataques são abortados com a fuga imediata de volta ao território de origem.

Mas, caso a situação seja propícia, a companhia de machos coopera eficientemente para isolar e levar o oponente à morte. Se mais de um adversário é encontrado e a vantagem numérica seguir sendo inequívoca para os atacantes, o reide poderá ser empreendido a partir da tática de negar aos inimigos a capacidade de cooperarem, por meio da garantia de seu isolamento no terreno. Fêmeas e machos jovens que acompanhem o grupo invasor normalmente observam a ação sem se engajar nela. As incursões geralmente terminam uma vez assegurada a morte do oponente, e podem envolver demonstrações etológicas bastante peculiares, como a emasculação de oponentes caídos, mas ainda vivos, ou o consumo de seu sangue. Há descrições de que os efluxos de um adversário foram compartilhados entre um macho experiente e outro mais jovem, pertencentes ao mesmo bando.

O recuo de volta ao território de origem não ocorre antes de se empreender coerção sobre eventuais fêmeas jovens do grupo inimigo (envolvendo violência física não letal e intimidação), para que se juntem ao grupo vencedor (não é incomum que fêmeas mais velhas sejam eliminadas). A abdução das fêmeas ocorre ou por sua migração ou pela incorporação do território por onde forrageiam, a partir do momento em que forem poucos os machos adversários capazes de negar futuro acesso dos invasores. Ressaltemos que, num contexto de fim da exclusividade sexual (ainda que ela seja distribuída de modo desigual, com base no escalonamento hierárquico intragrupo), a incorporação de novas fêmeas à sua macrounidade social garante a todos os machos engajados em violência intersocietária o aumento potencial de seu *fitness* reprodutivo em algum grau, caso cooperem.

A dimensão reprodutiva da violência coalizacional ganha ainda visibilidade pela prática do infanticídio após a abdução: os primeiros filhotes nascidos de fêmeas recém-incorporadas tendem a ser mortos pelos machos quando nascem, enquanto as gerações seguintes são preservadas. Em um regime poliginândrico, não há garantia de paternidade intragrupo, e esse fato tende a coibir os atentados masculinos contra infantes; mas, no caso da absorção de fêmeas estrangeiras, a possibilidade de não paternidade da primeira geração é razoável, e o infanticídio visa assegurar a pureza da comunidade genética patrilinear.

A agressão letal não se configura exatamente como um fenômeno etológico raro entre os mamíferos, mas a parte do leão nesses casos envolve infanticídio, praticado em nível individual por machos ou fêmeas; ou ainda a disputa por recursos naturais em situação de escassez. Na competição reprodutiva, duelos interpessoais são igualmente comuns, e podem resultar em letalidade, embora não seja a regra. Em termos etológicos, a agressão letal entre adultos é um comportamento com altíssimo custo: em circunstâncias de simetria de poder, pode resultar tanto na morte da vítima quanto do agressor. Desse modo, a letalidade pode se fixar como traço de comportamento agonístico quando: 1) amplia o *fitness* reprodutivo do agressor; 2) ocorre em condições nas quais os riscos são controlados.

Essa é a lógica que pauta o infanticídio, sem dúvida a categoria mais comum de violência com morte: os riscos de contra-ataque serão nulos caso os infantes não sejam protegidos por fêmeas altamente dimórficas (o que não é o caso entre os antropoides), por machos dominantes em haréns, ou pela cooperação de múltiplos machos em regimes sociais como os dos chimpanzés (nestes últimos dois casos, considerando exclusivamente agressão perpetrada por agente externo ao grupo social). Desse modo, o que faz a violência coalizacional intersocietária rumar pelo caminho da agressão letal é justamente o desequilíbrio de poder, a assimetria explorada pela tática cooperativa. O equilíbrio de poder é etologicamente um mecanismo eficaz de contenção da violência, e as coalizões masculinas rompem justamente esse equilíbrio: na medida em que os agressores nos reides de chimpanzés raramente sofrem qualquer dano, a eliminação física de machos considerados externos à comunidade genética patrilinear acaba por garantir seus frutos em termos de agenda reprodutiva.

Mas as relações intersocietárias entre nossos parentes evolucionários mais próximos, podem elas ser objeto da etologia de resolução de conflitos, daquela ética intuitiva da qual falamos? Teria o ancestral comum entre homens e chimpanzés sido capaz de se comportar, no âmbito intersocietário, pautado por instrumentos cognitivos de contenção da violência? Temos que, no âmbito da inteligência social modular – mecanismo exemplar de ordenamento das relações em contexto de caos sistêmico intragrupo –, existem limites de processamento claros, relacionados à capacidade cerebral. O quociente de encefalização,[13] o volume neocortical, a demografia dos grupos e o tempo de *grooming* são variáveis associadas. Quando o volume de informação social produzido pelo número crescente de relacionamentos simultâneos supera os limites de processamento da mente modularizada, a coordenação e a cooperação se tornam cada vez menos viáveis. Então, quanto maiores os grupos, mais tempo social é necessário para se reforçar os laços, e maior é a demanda sobre o aparato cognitivo no sentido de coletar informações sobre o *status* alheio, construir hipóteses sobre as estratégias de ascensão social de terceiros e, com base nelas, posicionar-se tendo em vista, pelo menos, a preservação do próprio *status*. A sobrecarga de informação social faz com que o reconhecimento e a análise do lugar hierárquico de certos indivíduos se tornem vagos, criando situação anômala na qual os instrumentos de manejo de conflitos perdem eficácia. O facciosismo que aos poucos se instaura nos grupos em ruptura expressa justamente a capacidade de se identificar o *status* dos indivíduos com quem os laços seguem firmes e a dificuldade de se compreender onde os "outros" se encaixam nesse maquinário social.

Uma vez intensificada a demanda de processamento mental para além da capacidade cognitiva desses primatas, e considerando-se que tanto os mecanismos de resolução de conflitos quanto os jogos

13 O quociente de encefalização expressa a razão entre o volume cerebral médio em uma espécie e o volume esperado para o cérebro, caso esse órgão se desenvolvesse em condições isométricas (proporcionais) com o restante do corpo.

de *status* são aspectos incontornáveis do exercício de uma sociabilidade pós-harênica, o instrumental cognitivo seguirá sendo demandado, o que resultará em comportamento patológico e nítido sofrimento emocional. A fissão definitiva da macrounidade social atua, então, como fenômeno homeostático, reconduzindo o funcionamento dos módulos mentais especializados ao equilíbrio. Uma vez fracionados, os dois grupos recém-formados terão reconduzido seu contingente demográfico a limites cognitivamente manejáveis (Aielo, Dunbar, 1993, pp. 184-185; Aureli et al., 2008, p. 627; 637; Bauernfeind et al., 2013, pp.275-276; Ferguson, Beaver, 2009, p. 291; Mithen, 2002, pp. 140-141; Wrangham, Peterson, 1996, pp. 5-18; 158-159; 162-170; 179).

Considerações finais: a ética da guerra é fruto de condição apomórfica em *H. sapiens*

É dessa forma que, em tese, desde o último ancestral comum, chimpanzés e hominíneos seriam incapazes de incorporar suas relações intersocietárias ao campo da ética política intuitiva, algo que torna bastante incomuns os desenvolvimentos mais recentes na história evolucionária recente de *Homo sapiens*, com a eclosão da modernidade comportamental, da consciência transdominial e do pensamento abstrato, por volta de 40 mil anos atrás. Nessa longa trajetória evolucionária desde seis milhões de anos, rediviva sob certo aspecto nas sociedades contemporâneas de *P. troglodytes*, a classificação de um coespecífico como estrangeiro é fruto do descarte de informação social. O "outro", uma vez desligado de uma macrounidade social, passa a não ocupar qualquer lugar na hierarquia interna, deixando então de ser objeto dos processos cognitivos inatos dedicados ao gerenciamento de conflitos.

Os sinais somáticos demonstrados por chimpanzés em seus encontros com o inimigo sugerem que, ao contrário de terem eles sua "chimpanzeidade" reconhecida, são tratados como animais de caça. Em um ato de violência coalizacional intersocietária, os agressores emitem sinais vocais e gestuais que coincidem com a prática de encontrar e perseguir uma presa em fuga. A "deschimpização", ou seja, o processo cognitivo de ressignificar a natureza de um coespecífico, é universal e etológico, e não um procedimento socialmente aprendido para, supostamente, controlar uma aversão inata ao assassinato. Se assim fosse, deveria ser restrita a determinados grupos de chimpanzés, como o são as diversas faces da cultura material nessa espécie.

O reenquadramento da condição do estrangeiro funciona como artifício etológico voltado para deflagrar respostas do sistema nervoso simpático associadas ao exercício da violência letal, como na caça, e isso não está relacionado, de forma alguma, a qualquer reação defensiva. Chimpanzés são capazes de ignorar a presença de outros primatas potencialmente perigosos, como babuínos, com os quais ocasionalmente disputam alimentos. Esses primatas não pressionam os gatilhos etológicos ligados à violência coalizacional intersocietária, embora sejam uma ameaça em potencial. O babuíno não é o inimigo, mas um chimpanzé pertencente a outra macrounidade social.

É sintomático, então, que o estrangeiro represente a incerteza em seu mais alto grau, já que pertence ao campo do não social, ausente seja da base, seja do topo da pirâmide, ignorando tanto a dominância quanto a submissão. Sendo impossível a identificação de seu lócus hierárquico, não há estratégias sociais possíveis de serem traçadas a seu respeito, tornando a cognição modular inócua. Uma vez avessos à ordem e representando o caos exterior de um mundo desprovido de instrumentos de manejo de conflitos, ao estrangeiro restará a aniquilação, facultando-se somente às fêmeas a oportunidade de integração ao campo ordenado das relações sociais etologicamente controladas. Na circunstância de os grandes primatas contarem com mecanismos sensoriais de identificação de consanguinidade – o que em *H. sapiens* integra o campo do inconsciente pessoal, algo potencialmente sugerido pelo mito edipiano –, devemos compreender o quão forte precisa ser a pressão ambiental e cognitiva para a cisão de comunidades patrilineares, bem como para a ressignificação da natureza do "outro", com quem, em última instância, pode-se guardar relações de parentesco. Esses processos, resultando em violência letal, acabam por autorizar etologicamente o fratricídio, num equilíbrio complexo e instável com o próprio exercício da sociabilidade patrilinear cooperativa (Mithen, 2002: 308-309; Ferguson, Beaver, 2009: 287; Roscoe, 2007: 485-486; 491).

Consideremos, assim, a possibilidade desses mecanismos cognitivos de contenção e gerenciamento de conflitos sociais intragrupo (que, na interação entre a consciência transdominial, o inconsciente pessoal e o vasto universo etológico do inconsciente coletivo, chamaríamos de "pensamento ético social humano") serem uma simplesiomorfia[14] entre humanos e chimpanzés, sujeitos, é claro, a disrupções de natureza patológica. Da mesma forma, a capacidade do exercício da violência coalizacional intersocietária (com garras, dentes, espadas ou armas nucleares) parece manifestar-se como condição simplesiomórfica nas espécies dessas duas linhagens.

O que emerge desta reflexão como fruto de algo substancialmente apomórfico[15] é a capacidade de *H. sapiens* de dispor de uma ética da guerra, do poder de formular normas abstratas que determinem os limites e os parâmetros do exercício da violência intersocietária, e que, em última instância, possam eventualmente culminar na sua negação. Não obstante, o pacifismo e as normas da guerra, nesse caso, parecem longe de se configurarem como condição etológica, depositada no inconsciente coletivo humano, e depender, exclusivamente, do exercício da consciência transdominial. Já a violência intergrupal e a desumanização, ainda que combatidas firmemente nos domínios do consciente, encontram repouso firme nos recônditos do inconsciente coletivo, herdados de um turbulento passado evolucionário.

14 Simplesiomorfia designa qualquer característica primitiva compartilhada por duas ou mais espécies. Essa característica não é distintiva de qualquer das espécies que dela compartilham. *H. sapiens* por exemplo, não possui cauda, tal como os gorilas; esta portanto não é uma condição que defina o humano moderno, tampouco os gorilas.

15 Apomorfia é uma característica inovadora presente em determinada espécie, e que a faz diferir de todas as suas ancestrais. A bipedia é uma provável apomorfia na linhagem dos hominíneos desde a divergência com o último ancestral comum entre eles e os chimpanzés.

Referências

AIELLO, L., DUNBAR, R. Neocortex size, group size, and the evolution of language. *Current Anthropology*, v. 34, n. 2, p. 184-193. 1993. DOI: 10.1086/204160.

AURELI, F. et al. Fission-fusion dynamics: new research frameworks. *Current Anthropology*, v. 49, n. 4, p. 627-654. 2008. DOI: 10.1086/586708.

BARNOSKY, A. e KRAATZ, B. The role of climatic change in the evolution of mammals. *BioScience*, v. 57, n. 6, p. 523-532. 2007. DOI: 10.1641/B570615.

BAUERNFEIND, A. et al. A volumetric comparison of the insular cortex and its subregions in primates. *Journal of Human Evolution*, v. 64, n. 4, p. 263-279. 2013. DOI: 10.1016/j.jhevol.2012.12.003.

BRAUDEL, Fernand. *Escritos sobre a História*. São Paulo: Perspectiva, 2009.

CAMERON, D., GROVES, C. *Bones, stones and molecules: 'out of Africa' and human origins*. San Diego: Elsevier, 2004.

CHRISTIAN, David. Macrohistory: the play of scales. *Social Evolution & History*, v. 4, n. 1, pp. 22-59, 2005.

FERGUSON, C. e BEAVER, K. Natural born killers: the genetic origins of extreme violence. *Aggression and Violent Behavior*, v. 14, n. 5, p. 286-294. 2009. DOI: 10.1016/j.avb.2009.03.005.

FOLEY, R. *Os humanos antes da humanidade: uma perspectiva evolucionista*. São Paulo: UNESP, 2003.

LADEIA, I., FERREIRA, P. A história evolutiva dos primatas. In: Neves, W. et al (orgs). *Assim caminhou a humanidade*. São Paulo: Palas Athena. p. 48-85. 2015.

MITHEN, S. *A pré-história da mente: uma busca das origens da arte, da religião e da ciência*. São Paulo: UNESP. 2002

OTTERBEIN, K. A history of research on warfare in anthropology. *American Anthropologist,* v. 101, n. 4, p. 794-805. 1999. DOI: 10.1525/aa.2000.102.4.834.

PAMPUSH, J. et al. Homoplasy and thick enamel in primates. *Journal of Human Evolution*, v. 64, n. 3, p. 216-224. 2013. DOI: 10.1016/j.jhevol.2013.01.009.

NORDHAUSEN, M., OLIVEIRA Filho, P. Nós, primatas. In: Neves, W. et al (orgs). *Assim caminhou a humanidade*. São Paulo: Palas Athena. p. 14-47. 2015.

ROSCOE, P. Intelligence, coalitional killing, and the antecedents of war. *American Anthropologist*, v. 109, n. 3, p. 485-495. 2007. DOI: 10.1525/aa.2007.109.3.485.

SHULTZ, Susanne; OPIE, Christopher; ATKINSON, Quentin. Stepwise evolution of stable sociality in primates. *Nature*, v. 479, n. 7372, p. 219-222. 2011. DOI: 10.1038/nature10601.

WRANGHAM, R., PETERSON, D. *Demonic males: apes and the origins of human violence*. Boston: Mariner, 1996.

Warfare, Ethics, Ethology

Evolutionary fundamentals for conflict and cooperation in the lineage of Man

Daniel Barreiros
Universidade Federal do Rio de Janeiro - UFRJ - Brasil

ABSTRACT

The aim of this article is to set a macro-historical narrative concerning the emergence of warfare and social ethics as symplesiomorphic features in the lineage of *Homo sapiens*. This means that these two behavioral aspects, representative of a very selected branch in the phylogenetic tree of the Primate order, are shared by the two lineages of great African apes that diverged from a common ancestor around six million years in the past, leading to extant humans and chimpanzees. Therefore, this article proposes an ethological understanding of warfare and social ethics, as both are innate to the social high-specialized modular mind present in the species of genera *Pan* and *Homo*. However behavioral restraints to intersocietal coalitionary violence seems to be an exclusive aspect of the transdominial modular cognition that characterizes modern humans. Thus, if in the evolutionary *long durée,* warfare and restrictions to intrasocial violence both appear to be ethologically common to humans and chimpanzees to a certain extent, an ethics of warfare - and, of course, the cognitive capability for intersocietal peace - seems to be distinctly human.

Correspondence | Daniel Barreiros, daniel.barreiros@ie.ufrj.br
Citation | Barreiros, D. (2018) Warfare, Ethics, Ethology: Evolutionary fundamentals for conflict and cooperation in the lineage of Man. *Journal of Big History*, II(2); 19 - 35.
DOI | http://dx.doi.org/10.22339/jbh.v2i2.2300

The aim of this article is to set a macro-historical narrative concerning the emergence of warfare and social ethics as symplesiomorphic features in the lineage of *Homo sapiens*. This means that these two behavioral aspects, representative of a very selected branch in the phylogenetic tree of the Primate order, are shared by the two lineages of great African apes that diverged from a common ancestor around six million years in the past, leading to extant humans and chimpanzee. This narrative is not intended to replace any well-established interpretations about war and peace sposed by social sciences and humanities in general; on the contrary, its objective is to add a new depth to these interpretations, from an evolutionary point of view. We hope that all contradictions brought forth by this “play of scales” can foster dialectical reflections about the relationship between social action and the historical durations, in the trail blazed by Braudel (2009) and Christian (2005).

There is a vast assortment of definitions - current or obsolete - for these widespread phenomena we call warfare and ethics in human relations. But, in the perspective of this essay, most of them - if not them all - end up sharing a common limitation. The universality of these supposedly contradictory aspects of human behavior is a genuine stimulant for our imagination - that gives birth to a *homo ethicus*, or, in the Janus face, a *homini lupus homini*. Nonetheless, that same universality makes us lose grasp of the uniqueness of conditions that made possible intersocietal conflict in primate societies, and, even further, that had set in motion neurocognitive mechanisms devoted to mitigate or prevent the escalation of intragroup agonistic behavior. In the *long dureé*, and considering the evolutionary history of primates in general, the intersocietal coalitional violence (a broad definition that makes all human warfare nothing but a particular example of a general phenomenon) and the complex array of ethological [1] mechanisms in the unconscious

1 Ethology, in general terms, is the study of animal behavior.

mind that gives foundation to ethical knowledge in *Homo sapiens* are absolutely exceptional. And we don't refer here to interpersonal violence, this common behavioral feature in the primate world, emerging as an evolutionary asset in the dispute over energy sources and reproductive opportunities. Some kind of primate sociability in the lineage of Man, emerging six million years before the present, has brewed up the context for the selection of intricate cognitive instruments, devoted to intragroup conflict resolution, based on complex status hierarchies, non-lethal violence, ritualizations and social stratagems. These phenomena, we believe, are mostly - but not exclusively - tributary to the rare development of masculine cooperative patrilineality, and, together, set intersocietal coalitionary lethal violence in motion. In this macro-perspective, humans and chimpanzees are the only species that make war. Furthermore, they are actors in a complex and daily social drama, in which the balance of power and prestige among "cooperative competitors" is highly volatile, and the potential for fratricidal lethal violence has to be kept under control by the operation of a well developed social modular mind. Humans and chimpanzees share warfare, ethological constraints to lethal violence among peers, and 98,8% of their genes. If this set of behavioral aspects hadn't emerged independently (homoplasically) in the two lineages that led, one to *Pan troglodytes* and other to *Homo sapiens*, it must have been manifest in species *before* the divergence, or, at least, in the last common ancestor (LCA) of humans and chimps. Would the behavioral potential for the projection of external power and for the deterrence of intrasocial conflict be phylogenetic [2] ? In an evolutionary perspective, would warfare and ethics be the offspring of the same womb? And what's to say about an ethics of warfare?

1. Multissexual unstable sociability in the Early Eocene

The emergence of sociability among primates, around fifty-two million years ago, didn't seem be a sufficient condition to generate a specific evolutionary context for the ethological foundations to warfare and ethics. Early Eocene [3] has brought, for some new putative species, the behavioral innovation of multissexual unstable groups, in detriment to the solitary life still led by other primate species, remnants of an even older paleocenic world. The main characteristic concerning these ancestral forms of sociability must have been an intense volatility regarding to the internal composition of social groups, with frequent processes of fusion-fission, in response to demographic saturation and availability of resources. In these terms, both adult males and females tend to migrate from their natal groups to others, and in most cases, more than once in a lifetime. If we consider the eocenic climatic context of (natural) global warming, worldwide environmental homogeneity, the expansion of rainforest across the continents (including Antarctica), and the expansion of the overall energetic supply in most of the world ecosystems, we should infer that dispersion of individuals over the territory would be high, inasmuch as the risks of foraging dispersal (predation) would be offset by the richness (in qualitative and quantitative aspects) of nutrients supply.

Diurnality in primates could be associated with the development of this unstable sociability, with stereoscopic vision and with the expansion of depth perception. In arboreal species, the latter would enable organisms to pinpoint fruits and other nutritional high-value resources in conditions of visual pollution (closed canopy forests, with scarce luminosity), therefore opening to primates this rich ecological niche, put forth by the expansion of angiospermic vegetation. At the same time, and as a trade-off, diurnality placed them at a disadvantage to these species living in daylight, consubstantiated

2 Are phylogenetic all characteristics inherited by a species from others in their direct ancestral line.

3 Eocene is the geological period comprehended between 56 and 33,9 million years before the present, according to the International Commission on Stratigraphy (ICS). For more information about this subject see http://www.stratigraphy.org/index.php/ics-chart-timescale. The geochronology followed in this essay follows the international convention.

in an increased exposure to predation risk. Against that, gregariousness acted as an equilibrium strategy, enhancing the number of sensorial units willing to simultaneously scan around for potential threats, and sharing that information for mutual benefit. In that way, possibly, the unstable sociability among primates could have emerged: as an anti-predator strategy, merely pragmatic, unable to form neither lasting bonds among individuals, nor complex forms of cooperation and coalitions (Groves, Cameron, 2004: 36; Ladeia, Ferreira, 2015: 56-58; Shultz, Opie, Atkinson, 2011: 219; 222).

2. Climatic change, from the Oligocene to the Early Miocene: *Proconsul* and the matrilineal cooperative female sociability

With the relative environmental homogeneity of the Eocene giving place to gradual global cooling and aridification in the Eocene-Oligocene [4] transition, the African evolutionary board was distinctively disturbed, kickstarting a new context for speciations and extinctions. This worldwide climatic transformation was simultaneous to (and reinforced by) intense tectonic activity and orographic changes in the Early Miocene [5], that resulted in the rising of the Himalayas, the Tibetan Plateau and the Ethiopian highlands. The geographic relief, thus, prevented the entry of moist air currents from the Indian Ocean, taking a heavy toll on East African ecosystems. As an environmental consequence, the territorial distribution of forest resources became even patchier, with the in-between spaces aridified in savannah-like form. The multiplicity of new niches ended up contributing to the selection of evolutionary innovations in the order of Primates.

Proconsulidae is a family of quadrupedal primates that encompasses a few miocenic species emerged around twenty-three million years ago, among which *Proconsul africanus* is the best known. Their dental anatomy is gracile; this suggests, in biomechanical terms, arboreal habits in tropical or subtropical closed forests, and a diet consisting of soft fruits and leaves, something very similar to their eocenic and oligocenic ancestors. In this species, the thin enamel layer all over the dentition is prone to wear due to abrasion (even in low levels), therefore increasing the shearing action of the cuspids. That is an adaptation commonly associated with primates occupying niches with plenty of soft and tender foods, that require little oral preparation and mastication (Pampush et al., 2013: 218).

Nonetheless, we should consider that environmental conditions on the Miocene of East Africa were creating a patchwork of forests (“isles”) surrounded by savanna, and if proconsulids really depended on the exploitation of arboreal resources, these populations were certainly trapped inside these patches, with very likely effects on their social strategies. Inferring from the fossil register, if we also take in consideration the possibility of a reasonable level of sexual dimorphism in terms of body mass and shape/size of canine tooth (in *Proconsul africanus*, at least), we can suggest that males, corpulent and heavily-armed, were involved in intense levels of reproductive and territorial competition, based on agonistic behavior (intimidation, vocalizations, interpersonal violence).

According to this scenario of climatic change, arboreal life, spatial insulation of the natural resources and sexual dimorphism, we can suggest that proconsulids probably belonged to a group of pioneer primate species that first lived in stable societies. If we go further, and consider the reproductive and energetic costs imposed by intrauterine gestation and lactation, the access to high-quality and regularly supplied nutritional resources is the most important evolutionary demand in female energetic ethology. This is a central issue in the patterns of primate territoriality and social structure, which means to say that, in the Miocene of East Africa, where and when the most valuable forests were becoming even more isolated by vastitudes of arid plains, females belonging to

4 The Oligocene was a geologic period comprehended between 33,9 and 23,03 million years in the past.

5 The Miocene was a geologic period comprehended between 23,03 and 5,33 million years in the past.

arboreal species like proconsulids would seek to occupy and defend those secluded spaces. While climate and aridification were not harsh enough to deplete the nutritional value of resources concentrated in these forest patches - expanding the African savannah even more -, the ecological context would favor the exploitation by groups of kin-related females, cooperating to guarantee the access to preferred foods in the benefit of the genetic matrilinear community, and to exclude other groups of non-kin females. In the case of males, the energetic and reproductive demands are minimal (including gametic production), so the main challenge is to guarantee access to females. In this way, primate territoriality ends up being conformed mainly by female energetic strategies, since males just follow the patterns of spatial dispersion shown by female collectives. Therefore, males tend to migrate from their natal groups after reaching sexual maturity, in order to confront other males for sexual opportunities far away from their own genetic community.

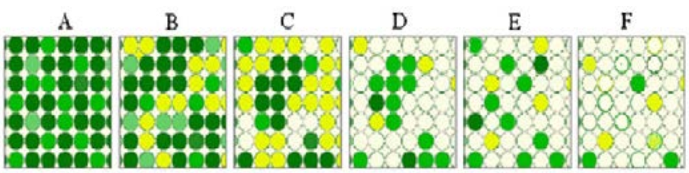


Figure 1. Resources distribution

In this simplified model, the darker the green, the richer the nutrients. A) High quality resources, deconcentrated in space. There is no circumstance for concentrating females in patches; and they forage alone and separate themselves from the others, avoiding competition with related females. Unstable groups emerge as anti-predation strategy. Males migrate from their natal groups when reaching maturity, as well as females. B) High and medium-quality resources are concentrated in homogeneous and large-scale patches. Kin-related females gather on these spots, which are rich enough for them to feed together, as long as unrelated females are kept at bay. The defense of the female genetic pool prevails. Unrelated males gather around these females, fighting for sexual opportunities. C) High quality resources are extremely concentrated in uniform patches. Under these conditions, a lone male, if tough enough, is able to control the female foraging territory, excluding all competing males and establishing a harem. D) Quality of resources decreases, and the distribution pattern is maintained. Females scatter spatially in search of medium quality energy patches. Solidarity among females decays, as kin-related groups becomes counterproductive. There's opportunity for harems to continue, this time organized in patrilineal lines. It coincides with the sociability pattern present in *Gorilla* sp. E) Distribution of resources becomes heterogeneous. Kin-related female groups, already exhausted, become even less possible, as well as the harems. A lone male becomes unable to prevent the access of competitors as females spread to take advantage of scarce and scattered best quality resources. Harems are impossible, but patrilineality is preserved. Coalitions of kin-related males are formed, to dominate over dispersed females. It matches the sociability pattern in *P. troglodytes*. F) Resources becomes too poor and scattered, condemning permanent sociability. In *Pongo* sp., leads to female dispersion, to individual occupation of patches, and the formation of male super territories.

This scenario does not account for the emergence of intersocietal coalitional violence as a behavioral aspect, and seems reasonable to assume that it didn't exist in any primate species in the Early Miocene, even taking into consideration the possibility of stable sociability among kin-related females. Among proconsulids, is probable that uncooperative groups of non-kin males have been tolerated by the matrilineal collectives only if the territory and resources were compatible with the extenuating reproductive and energetic strategies of females. However, with the ongoing climate crisis in Early Miocene, and the possibility that supply and spatial concentration of natural resources have reached a critical threshold (Barnosky, Kratz, 2006: 528), these turbulent and competitive collectives of males could have become an unbearable burden by pressing over the scarce food available and by the intense exercise of violence and harassment among themselves, generating increased levels of psychological stress, inducing to female reproductive suppression.

Also, that should be the case for the development of one-male groups, or harems, in which a dominant male shows himself capable of establishing exclusivity of access to reproductive opportunities, represented by the existence of a stable collective of females. That privilege only becomes possible if the harem holder is able to sensorially watch over the estrous females, which is that made easier by spatially concentrated forests. Consequently, being able to protect the foraging territory where females live, a single dominant male could then use his improved physical and psychological weaponry to prevent access to other postulant males, in their search for reproductive wayouts. To the kindred females, the reduction of the amount of organisms alien to their genetic community - the migrant males - should represent less pressure over decisive energetic resources, therefore enhancing female reproductive fitness in group level. Harems were a very possible form of sociability among proconsulids, considering their morphological characteristics and the paleoecological conditions in which they probably lived. If this were the case, the harem strategy would have provoked high level of tension and interpersonal masculine violence. That conclusion is, as a matter of fact, supported by the dimorphic condition in proconsulids. Hence, intense agonistic behavior among unrelated and uncooperative males (especially in the presence of harem holders) is something that would prevent the emergence of that very special condition for primate warfare and for the ethological restraints to intragroup masculine conflict: patrilineal male coalitions (Cameron, Groves, 2004: 38-40; Foley, 2008: 220-227; Ladeia, Ferreira, 2015: 75; Nordhausen, Oliveira Filho, 2015: 36-37; Wrangham, Peterson, 1996: 131; 174-175).

3. *Afropithecus*: stable matriachies and the savannah

Afropithecids, a distinct family, diverged in Africa around eighteen million years in the past, having some proconsulid species as their most likely direct ancestor. Fossil record suggests that this new sort of primates have adapted some behavioral innovations, probably developed by their evolutionary forefathers, to a paleoecological context somewhat distinct. This behavioral heritage becomes even more relevant when we consider that an afropithecid was the presumed ancestor of all hominids, which includes, *lato sensu*, modern humans. In general, this extinct family had species morphologically quite distinct from the profile of their proconsulid ancestors, with whom they have coexisted. Their upper body morphology suggests they were brachiators [6], which means that their main motor strategy consisted of suspensory locomotion in a forest environment, just like gibbons, and in a minor extent, chimpanzees and other extant great apes. The estimated average body mass of afropithecids is greater than what is expected from arboreal primates, which advocates in favor of semi-terrestrial habits. Considering the body plan adapted to brachiation (more vertical posture, different from quadrupedal primates like proconsulids), the motor strategy employed by afropithecids on the ground could have been based on knuckle-walking (like extant gorillas and chimpanzees), an aspect that could have future consequences for the exercise of territoriality.Assuming the possibility of semi-terrestrial behavior, coupled with the evidence for robust facial architecture, thick dental enamel and powerful chewing dentition, it is suggested that, unlike their ancestors, afropithecids occupied marginal habitats to the humid forests, with regular incursions into the savanna, from which they could have obtained fallback foods according to environmental fluctuations. Although the dry African

6 Brachiation is a form of suspensory locomotion in which the animal moves hanging under tree branches, in vertical position.

plains offered mostly abrasive and hard provisions, afropithecids were adapted for the exploitation of niches opened by the aridification.

The expansion of nutritional opportunities for these species - inasmuch as they were capable not only of exploiting forest resources, but also those present in savannas and woodlands - altered the relationship between society and space for these primates. We must consider that the spatiality of resources is a dynamic phenomenon, and, thus, depends on energy processing and of thermoregulatory efficiency [7] of the motor strategies employed. Therefore, if for arboreal and thin-enameled proconsulids these aridified habitats meant insurmountable barriers - which led their societies to concentrate in aisled forest patches -, for the more robust afropithecids, the savannah, to a greater or lesser extent, was also a space for foraging and exploration, which would have contributed to reduce the effects of demographic concentration on its social strategies.

Perhaps we should assume that, in this case, stable sociability - one of the primitive pillars for warfare and ethics - would have been contraindicated by the deconcentration of nutritional opportunities in the territory, which should have led - at least theoretically - to the spatial dispersion of females. That dispersion would be certainly counterproductive to the maintenance of permanent matrilineal female collectives, and hence, to the masculine sexual exclusivity strategies expressed, at the limit, by a successful harem behavior. All these facts would tend, if true to some extent, to conduct our afropithecid ancestors again to a profile of sociability very common to Eocene parameters - the unstable multisexual aggregation -, still practiced by many species of extant monkeys, especially New World platyrrhines [8].

In spite of this possibility, the phylogenetic hypothesis of the transmission of stable sociability from proconsulids to hominids should not be discarded prematurely, since we have good examples of primates also capable of occupying vast territories - as we believe afropithecids could have been - but which preserve alternative forms of harem behavior and matrilineal feminine collectives, such as some species of papionins. Among geladas (*Theropithecus gelada*) for example, harems have been preserved, but given the expanded relative spatiality of natural resources (considering geladas are grazing primates), it became possible to form troops, that consist in a sort of “confederations of harems”, composed of many reproductive units, each formed by related females and a dominant male, sometimes accompanied by a few subaltern males. This ethological pattern among geladas represents an aspect of morphological and behavioral flexibility present in variated degrees in all species. This means that there are “accommodation limits” through which an inherited aspect keeps being replicated - even in slightly modified forms -, in spite of the environmental change. Nonetheless, this limits can be broken by sudden or massive ecological transformation, giving rise to extinction / speciation processes. But, regarding afropithecids, we believe that is still defendable the hypothesis for the behavioral inheritance of matrilineal female collectives in one-male groups, especially if we consider that the same persists among some extant African primates, distantly descendant from that afropithecid radiation (Cameron, Groves, 2004: 39; Barnosky, Kraatz, 2007: 525; Foley, 2008: 150-151; 178-179; 183-184; Pampush et al., 2013: 222; Wrangham, Peterson, 1996: 56-59).

7 Knuckle-walking primates when moving terrestrially across open spaces (in savannah, with few trees), expose much of their body surface (head, shoulders, back) to the direct incidence of solar radiation, reasonably more than bipedal primates do.

8 Platyrrhines form a parvorder of primates that unites all New World monkeys. Unlike Old World catarrhines (such as humans, baboons, gorillas, chimpanzees, etc.), which have protruding noses facing downwards, platyrrhines have flat nostrils facing sideways. Some of its species have a prehensile tail, an evolutionary trace absent in all catarrhines.

4. The Middle Miocene and migrations across the Afroasiatic axis

The onset of the Middle Miocene, around fifteen million years ago, coincided with the expansion of polar ice caps and the lowering of sea levels, reaching a peak of intensity in the ongoing process of global cooling and aridification. Under these conditions, the northern portions of Eurasia became inhospitable to primates in general, while a milder and relatively homogenous biogeographic zone formed in the Sahara, East Central Africa, southern Europe, and the Levant. Many endemic African species are known to have migrated northward through the Mediterranean coast, and anthropoids also participated in these migrations. From these radiations out of Africa, many speciation processes have succeeded amongst the apes in Europe and Asia Minor without any significant impacts over the mainframe of social behavioral aspects, phylogenetically inherited by anthropoids from their afropithecid ancestors. Nonetheless, is important to note that from ten to seven million years ago, the moment of this primatological odyssey out of Africa, the reduction in diversity of the great apes species is impressive, something that probably have provoked the first major bottleneck in anthropoid evolutionary history. It is possible that this extinctionary context denoted that the behavioral and morphological portfolio based on robust dentition, semi-terrestrial locomotion and occasional exploitation of the savannah had met its limits of accommodation. In the final scene of that act, *Graecopithecus freybergi* or some close descendant species reached Mediterranean coast once again, escaping from the climatic aggravation in the north. The relevance of this species lies in the fact that perhaps it was the pivot of European anthropoid radiations across the Afro-Asian axis, where the harsh environment would put old social strategies once again to the test (Cameron, Groves, 2004: 41-42; 55-57; Ladeia, Ferreira, 2015: 76-77).

5. Pongins in Asia: closed doors to warfare

Migrations from southern Asia to the Far East, as vectors of latitudinal expansion, were marked by the incidence of partially constant climatic conditions, although progressively aggravated by the advance of glaciers through the northern portions of Eurasia. This homogeneity gave birth to a set of similar evolutionary challenges from Anatolia to Southeast Asia, something suggested by the morphological similarity between the extinct Asian great apes and *Pongo sp*. (the two species of orangutans, restrict today to the islands of Sumatra and Borneo). They share an odontomorphological profile inherited from European ancestors that migrated through meridional zones - thick dental enamel, robust molars compared to the rest of the dentition - although, as already mentioned, these characteristics, by themselves, probably no longer guaranteed survival against the harsh seasonality and aridification. Thus, the Asian great apes should have been able to adapt to the extreme rarity of tropical forests, the impoverishment of nutrient supply, and the radical deconcentration of resources in southern Asia (lowland plains) through behavioral patterns present today in the last extant species of the subfamily Ponginae.

It is likely that stable matrilineal female collectives have become unviable; with scarce and extremely fragmented resources, either females would need to scatter around for food, or compete with one another in a very restricted territory, poor enough to supply the energy necessary for them all. In terms of kin selection, coinhabiting related females would inevitably lose reproductive fitness in the circumstance of fiercely struggling for food, whereas, if each of them migrate to different areas, the possibility of individual success would emerge without the otherwise necessary failure of one or more relatives in the survival race. This is what orangutans do: among them, matrilineal female cooperation does not exist, since each female with her immature offspring is fixed in a small arboreal patch, separated from others females, only with sufficient resources to maintain this stable family unit. Group sociability is likewise dissolved, and consortia between males and females become temporary, creating a loose network of relationships over a vast territory.

For males, female territoriality prevents the traditional sexual exclusivity strategy (harem-holding) given the practical inability to prevent access to competing males. Nonetheless, the ethology of sexual exclusivity remains relatively alive, as dominant males circulate terrestrially through various female-controlled patches (creating a super territory), attempting to secure their reproductive privileges and eliminate competitors, and, in this way, forming something like a "loose harem". The degree of uncertainty about paternity generated by that kind of female territoriality leads to a high level of competitive pressure among adult males, which is expressed by notorious sexual dimorphism, vocalization behavior (signaling the presence in the territory) and intense level of interpersonal violence. Thus, pongins represented an evolutionary dead end with regard to the ethology of warfare and ethics: the formation of matrilineal collectives, as well as of any patrilineal male group, cooperative or not, is prevented. The high degree of agonism [9] and of physical violence in particular does not advocate in favor of warlike behavior, nor does it generate a context demanding conflict resolution strategies (Cameron, Groves, 2004: 75-77; Foley, 2008: 218; Nordhausen, Oliveira Filho, 2015: 29; Wrangham, Peterson, 1996: 133-134).

6. Gorillins in Africa: non-cooperative patrilineality inhibits coalitionary intersocietal violence

Pongins are primates not directly relevant to the phenomenon of intersocietal coalitionary violence and of ethics among modern humans, since they consisted of an evolutionary branch divergent from that which would result in the great African anthropoids. The last common ancestor of Asian pongins and African great apes lived about twelve million years ago, precisely in the period of radiations out of Europe. Rather, what interests us in particular are the hominids that have taken the way to Africa, coming from the north, since among them the ethology of warfare and ethics would take another decisive step.

The earliest species allocated under the Gorillinae subfamily emerged about ten million years ago, during the Late Miocene. These longitudinal migrations from Europe to the south represented the movement of species partially adapted to conditions of temperature and aridity more severe than those found in Central Eastern Africa. Thus, in a less rigorous environment than that existing all over the latitudinal range from Mediterranean Asia to the Malay Peninsula, some kind of permanent group sociability was still viable, and even more: if we take extant gorillas as reference, a harem sociability could be sustained. Although these migrant species found biomes less affected by climate change in African soil, the primitive strategy, based on robust dentition dependence, terrestrial knuckle-walking and opportunistic savannah exploitation, seems to have been abandoned. Odontomorphological profile of gorillins differs significantly from that of their European ancestors, indicating dietary specialization (again) in wet forest resources: once more among apes a graceful dentition emerges, with thin enamel, showing that reliance on this sort of retracting niche was still an option with evolutionary short-term returns. With natural resources less scattered than in southern Asian environments, although with inferior nutritional quality to that enjoyed by the ancient Eocene primates, the harem strategy could have been replicated by gorillins in Africa, but not the stable cooperative female matrilineality. With diminished quality and supply of resources, although distributed with some uniformity and concentration, female energetic ethology was somehow threatened. Female stable sociability continued to be advantageous - the advantages of gregariousness since Eocene times were well established at that time - but no longer among kin-related individuals. Among gorillins,

9 It is characterized as agonistic any form of conflict behavior involving physical violence or intimidation.

the females are the ones that mainly migrate to other groups after reaching sexual maturity. This means that kin cooperation becomes unfeasible, and gregariousness occurs between unrelated individuals. The collapse of the female kin cooperation means that the empowerment for self-protection against males also cedes, giving space to the advancement of masculine reproductive agenda within these harem societies. The exercise of power and sexual exclusivity by a dominant male becomes compatible with maintaining one or more of his adult sons in the group (although these young males may also leave the group and fight for sexual privileges elsewhere). This kin-related males means more pressure over already scarce nutritional resources, reinforcing the female migration imperative (they leave their natal group in order to avoid disputing over resources with their relatives).

The emergence of patrilineal links among African great apes in the Late Miocene does not signify the generation of cooperative relations among males. Sexual exclusivity, typical of harem regimes, is also a goal replicated in gorillin societies. To all kin-related males united in the same social group, except for the "silverback" [10], submission to reproductive exclusion is the only option. One of these subaltern males will have access to regular sexual opportunities only after the death of the harem holder (often, his father), or after achieving dominance elsewhere. Domination also extends to females, whose ties of solidarity have been severed by the collapse of matrilineality: to control an increased level of conflict over nutritional resources, product of a non-kin stable female sociability, it is not uncommon for the silverback the exercise of non lethal violence over their female protégées. There is, so to speak, a high degree of inner peace and submission to instituted power, and in no respect an "ethology of rebellion" is manifested, as would become usual millions of years later, to the common ancestor between humans and chimpanzees. There is no counter-hegemonic action, nor attempts to take power on the part of any members of the group; the threats to dominance usually happen in the circumstance of the arrival of young male migrants, who attempt to usurp the harem from an established silverback. Sexual exclusivity, as is customary, gives rise to a high degree of male interpersonal violence and sexual dimorphism. Again, despite the fact that two of the cornerstones of the ethology of warfare and ethics are present - stable sociability and patrilineality -, the harem regime and the lack of male intra-group kin cooperation makes this ethological phenomenon impossible (Foley, 2008: 224; Pampush et al., 2013: 217; Wrangham, Peterson, 1996: 147-149).

7. Late Miocene and the last common ancestor to humans and chimpanzees: complex hierarchies and male cooperative patrilineality

With glaciation increasing its intensity in ends of the Late Miocene, climatic disruption reached a new critical threshold, causing niche-conservative primates such as gorillas to migrate, following the retraction of forest zones. At the same time, some opportunity to adaptation emerged for primate populations living in more marginal environments. The relative spatiality of resources for gracile and knuckle-walking anthropoids did not become so disperse as happened in southern Asia, yet the concentration suitable for the preservation of harem societies seemed to replicate only with difficulty. With increasingly rare food resources in the territory, these populations living in ecologically marginal regions, between savannas and forests, needed to spread more and more across the territory. As we know, this ultimately means that unrelated females, pursuing their energetic agenda, had to forage farther apart from each other. As it would also be the case among the pongins in their latitudinal migration across southern Asia, female spatiality formed a perimeter incompatible with the exercise of surveillance by a highly dimorphic male, willing to maintain his sexual exclusivity.

10 "Silverback" refers to the dominant male in gorilla societies. The term derives from the silverish coloration displayed by the fur on the back of these primates after reaching sexual maturity.

Therefore, climatic extremes established for males a trade-off between the defense of their genetic community (patrilineality) and the exercise of individual reproductive agendas. In the case of Asian pongins, deprived of any inheritance of patrilineality behavior, this dilemma was null: the search for sexual exclusivity remained, despite the environmental circumstances, with a high degree of agonistic behavior among unrelated males. But, in the case of the last common ancestor to human and chimpanzees, the inheritance of patrilineality probably represented a factor that pointed the evolutionary arrow at unprecedented directions.

Instead of these ties between related males being dissolved, they become, on the contrary, even tighter, with the development of complex forms of male cooperation, a rare ethological condition. In the face of the impossible surveillance of the silverback over females, harems disappeared, and with them, the undisputed *locus* of male dominance. Avoiding fratricidal conflict for reproductive opportunities, in the circumstance of a "vacuum of power", these male patrilineal collectives organize themselves into complex status hierarchies. Without strict, gorillin-style dominance, mating becomes a polygynandrous issue [11], and male cooperation for the control of the foraging territory - and therefore, over non-kin females - arises. Kin-selection becomes clear, as individual reproductive agendas are relativized in face of the collective defense of male genetic community (against "alien" collectives of related males). With intragroup masculine agonistic behavior under control, interpersonal violence is reduced, as well as the level of sexual dimorphism. The emergence of complex ethological mechanisms for conflict management gives rise, then, to the rare phenomenon of intersocietal coalitionary violence (Aureli et al., 2008: 629-630; Foley, 2008: 230; Wrangham, Peterson, 1996: 52).

11 In polygynandrous regimes, males and females select occasional sexual partners without establishing stable bonds. Of course, polygynandry does not imply an equal distribution of reproductive opportunities. High-ranking individuals are favored, even though reproductive exclusivity (harem-like) is absent.

8. Post-harem sociability, cooperative patrilineality and intuitive ethics

Primates are imbued with an efficient general intelligence for problem-solving tasks. This implies that, in addition to simple inherited behavioral contents, their ethological portfolio includes learning mechanisms based on the interaction with the environment, with generic rules adjusted by trial and error. For most of the history of these species, accommodation to evolutionary challenges seems to have been possible by employing exclusively this sort of low-cost intelligence. The environmental context for the development of more specialized, energy-intensive types of cognition, seems to have emerged slowly, beginning ten million years ago, with the migrations back to Africa and Asia. These complex cognitive forms only become clearly visible with the modularization of social cognition, about six million years ago, with the emergence of the last common ancestor to humans and chimpanzees.

Highly specialized social cognition emerges as a mechanism of accommodation between competition and cooperation, between individual male reproductive agendas and stable post-harem patrilineality. In a context of dispute over sexual opportunities among related males, the fratricidal struggle is reduced to evolutionarily irrelevant levels through a complex capability to analyze the power *locus* of each member of the group in their relationships with the others, and to formulate hypotheses about the possibilities of rise and fall in the social pyramid for all the agents involved. From that natural proficiency in social strategy-making, an individual can trace his plans of preservation or conquest of status. The recognition of power and prestige levels of others, the temporary acceptance of one's own social condition, and the design of strategies for the contestation of hierarchy, in their own benefit and of their allies, appear as fundamental ethological guidelines among chimpanzees, and, presumably, also present in their last common ancestor with humans.

Chimpanzees and humans share a common

ancestor that most likely lived in stable societies with patrilineality and male patrilocality, marked by the high degree of uncertainty about the privileges and limits concerning each of the members of the social group. The relative simplicity of gorillinian harem hierarchies - in which male reproductive dominance was clear, monocratic, and subalternity was a common condition to all other members of the group - was replaced by a kind of intragroup "systemic chaos", in which multiple echelons in the hierarchical pyramid emerge, and the constant struggle for social ascension (masculine, especially) becomes widespread. This could have represented a favorable context for the dissolution of permanent sociability (given the potential for intragroup aggression), but in the lineages descending from the LCA [12], the survival of patrilineality ended up guaranteed by the emergence of this modular, high-cost social cognition, that coincides with the expansion of brain allometry in the Panini (the chimpanzee lineage of species, extinct and extant) and Hominini (all bipedal anthropoids emerged after the divergence with the first chimpanzees), when compared to the older anthropoids in the evolutionary tree.

The emergence of mental modules dedicated to the management of social relationships means that the stereotyped and generalized cognitive mechanisms produced at low cost by general intelligence have become insufficient to generate effective responses in a context of excessive "moving parts" (information) in social mechanics. The focus of modularity is not on its innate content, but on the ability to formulate testable hypotheses about the behavior of third parties, involving or not the presence of the observer. It is an extension of the cognitive complexes linked to the so-called "theory of mind" (ToM) present in different degrees of complexity throughout the Primate family. The ability to assess mental states is based on the modelization of reactions the subject would expect of himself if hypothetically immersed in a certain situation being observed but experienced by others, something that involves a reasonable degree of empathic skills. The extrapolation of these hypotheses must be calibrated according to the object's individual temperament (which is previously known, by definition) and to the circumstances of the action; the use of generic, standardized learning rules for strategic social decision-making, in the conditions of complexity present in these post-harem anthropoid societies, would result in a high probability of error.

It is difficult to support the idea that in the lineages of humans and chimpanzees, sociability is the product of social learning. Chimpanzees can be taught by humans to perform tasks in captivity that in their natural habitats would not be developed, since they do not fulfill any relevant evolutionary role (sign language and the production of lithic tools are two good examples). For such activities, chimpanzees employ their general intelligence, which functions as a kind of multi-purpose learning tool, producing simple results after some trial and error, but at low energy cost. As far as modularized social behavior is concerned, there is very little that can be taught to a chimpanzee, or that they should teach one another: even individuals born in captivity intuitively develop, at the right age, the social skills necessary for the intense "machiavellian" status games, which demonstrates their innateness. Thus, dedicated mental modules, that enable an individual to understand the functioning of social hierarchies and to devise status strategies, emerge with age, just like definitive teeth.

The modularity of social cognition seems to allow chimpanzees to develop some form of self-awareness, something suggested by mirror recognition tests. However, circumscribed to the scope of ethology, this awareness is far from equivalent to the *self*, the transdominial and transmodular holistic consciousness that only recently emerged in the evolutionary history of *Homo sapiens*. The modular social intelligence in chimpanzees seems isolated from the general intelligence, unable to interact fluidly with other non-modularized cognitive domains, so that these great apes become aware of themselves and of others only as social actors, and in the universe of social relationships. There is no substantive evidence of

12 Last common ancestor to humans and chimpanzees.

the use of material culture - related to the technical aspects of the general intelligence - for leverage in status disputes. There is not yet any symbolic dimension of material culture that is instrumentalized in order to transmit social information, to signal the hierarchical locus occupied by an individual, or to disguise the occupation of a lower echelon in the stratification pyramid. With the modular social mind incapable of accessing other cognitive domains, and placing them at its service, chimpanzees do not seem capable of complex mental simulations regarding foraging or tool-use issues, involving conspecifics. The general intelligence operates unconscious domains, incapable of producing mental perceptions and self-representations. This condition, though highly derived when comparing chimpanzees to other primates, is primitive in view of the cognitive transdominiality of modern humans.

Despite its insularity, the social modular mind has allowed the establishment of innate patterns and ritualistic norms in the struggle for intragroup power. From the behavior of chimpanzees in natural habitat, we know that these dominance clashes between two adult males can last for many months, and can be marked by intense demonstrations of agonism. It is common for the contestant male refusing to perform rituals of submission to the dominant male, rituals that are regularly attended by the other members of the group, as a form of reassertion of loyalty bonds, recognition of their place in the hierarchy, and stability of the social body. These loyalty rituals involve body postures and gestures, such as bowing down before the dominant male, or demonstrating what some primatologists call a "scared smile." As part of their power-signaling repertoire, dominant males tend to touch the shoulder of lower-status chimpanzees, and being touched this way is something that a contestant male tends to avoid at all costs in his struggle for ascension. These displays of intimidation, aggressiveness, and power are closely watched by all members of the group, who, over time, tend to give their support to one side or the other.

As the estrangement between factions builds up, daily coalitions for specific tasks - foraging, grooming [13], etc. - tend to become more volatile. Both the dominant male and his challenger seek to intimidate the females of the group, forming alliances with different subaltern males. What the disputants seek is the political support of the females themselves. The struggle for support is also reinforced by the increase in social time spent with each female and her offspring, through the practice of grooming. Allied subaltern males tend to help in keeping oppositionist females at bay, denying them the opportunity for hindering the adversary's efforts of socialization and support-gaining. What these lower-ranking males seek, therefore, is to climb steps in the hierarchy in virtue of the eventual victory of their "candidate", something that should provide increased reproductive advantages to them. In the end, after intense struggle for dominance, all the members of the group tend to converge to support one of the competitors, isolating the other. From there, the demonstrations of agonistic behavior on the part of the victorious male tend to slow down; the leader takes a conciliatory and pacifying stance, mediating conflicts between females and assisting weaker or less prestigious males against stronger opponents. At some point the process of contesting the hierarchy resumes, which often involves radical recomposition of alliances.

Provoked by the breakdown of harem dominance and by the preservation of male patrilineality, is the high degree of social uncertainty caused by the expansion of group size, the main reason for the development of an intuitive ethics in the lineages departing from the LCA to humans and chimpanzees? Functioning as a genuine prosocial ethological restrain, it rules the procedures of internal dispute for power, reduces the degree of male interpersonal violence (although it does not eliminate it), and establishes when and how a struggle cycle ends. Unlike chimpanzee cultures,

13 Grooming is an important social ritual among primates. Its primary purpose lies on removing parasites and dirt, for hygiene purposes. However, the role of grooming as a social act transcends that dimension, functioning as an important instrument for tightening affective bonds.

with socially transmitted practices and behaviors (by observational learning) exclusive to certain groups, a common core of sociability parameters and internal conflict resolution emerges innately in all groups of chimpanzees, in nature or in captivity (Aureli et al., 2008: 632; 636-637; Bauernfeind et al., 2013: 263-264; 271-273; Foley, 2008: 207-210; Mithen, 2002: 67-71; 102-111; 126-131; 139-142; Nordhausen, Oliveira Filho, 2015: 36-38; Wrangham, Peterson, 1996: 128; 143-144; 186)

9. Intersocietal conflict and exhaustion of social cognition

If the development of social cognition results in innate ethical standards in intragroup relations, intersocietal relations are precisely characterized by the absence of ethological restraints and rules concerning the management of social hierarchies. Among the common chimpanzees (*Pan troglodytes*) and, probably, the LCA, contact between social groups is restricted to coalitional violence aimed at eliminating foreign males, abducting fertile females, and disarticulating enemy communities. There are no cognitive mechanisms of pacification or containment of the lethal conflict between these societies.

Social groups occasionally split after struggles for dominance in the internal sphere. Socio-environmental limits are exerted over the cohesion of chimpanzees groups, something that involves a delicate balance between spatial distribution of natural resources and demographic factors. When these limits are exceeded, and a new cycle on the struggle for power begins, the breakup of the social macro-unit may be the result. One indication that political struggle may cause secession comes from the group segmentation into factions of relationship: the tendency to alternate partners in foraging and grooming gives way to more limited and repeated choices, with individuals reinforcing their social ties more often with certain partners than with others. As defections normally conclude a dominance struggle, with one of the competitors being abandoned by its supporters and progressively isolated by a growing majority, another important sign of secession in progress lies in the sustained allegiance of factions to their leaders. Then, the social fracture becomes visible in the very configuration of group spatiality: each faction, although inhabiting the same localities, will tend to rest apart from each other. Division goes on with the separation of distinct foraging grounds to each of the cliques, a situation considered of paramount importance in the secession process.

Being separated the two social units with their respective dominant males and hierarchical pyramids, small temporary subgroups will be formed in short time to fulfill daily tasks. One of these (not so daily) tasks consists in the exercise of intersocietal coalitionary violence. Incited by the dominant male (in most cases) through gestures and vocalizations, companies of adult males can be formed and march to the territory of the “enemy” group, sometimes accompanied by one or two nulliparous females. These initiatives are neither defensive measures or reactive patrols, nor byproducts of foraging expeditions. Leaving rich opportunities for obtaining food behind, these companies march with the sole purpose of bringing lethal violence to the “other.” Before raids begin, signals of the opponent’s presence (sounds in particular) provoke anxiety, only controlled by gestures that ensure trust and loyalty among fellow chimpanzees (touches, hugs). As we know, the ecological conditions in which chimpanzees live (and the LCA have lived, we presume), with scattered nutritional resources, leads to the temporary fission of permanent groups for the practice of foraging, and this fact creates the opportunity awaited by an aggressive band: to catch a lonely and unwary enemy male, distracted while feeding. Unlike warfare among human nations, chimpanzee clashes are necessarily asymmetric. With an erroneous assessment of circumstances that eventually results in contact between a company and a numerically equivalent group of enemies, the raids are always aborted, and the attackers flee in hurry, back to their homeland.

But if the situation is propitious, the companies of males are able to cooperate efficiently to isolate and lead the opponent to death. If more than one opponent is found, and the numerical advantage remains unequivocal for the attackers, the raid may continue with the tactic of isolating the enemy, denying them the opportunity to cooperate. Nulliparous females and young males accompanying the raiding party usually just observe all the action without engaging in it. Incursions ends after the opponent's death, and may involve peculiar ethological demonstrations, such as the emasculation of dying enemies or the consumption of their blood. Retreat to home territory does not occur before some coercion (non-lethal violence and intimidation) has been made on enemy young females (if available), in order to convince them to join the victorious party. The abduction of females happens either by their (forced) migration or the incorporation of their foraging territory, as soon as there are few enemy males capable of defending it. In a post-harem society with male patrilineal cooperation, the incorporation of new females into the social macro-unit ensures that all males engaged in intersocietal coalitional violence may find opportunities to enhance their reproductive fitness in some degree, if they cooperate. The reproductive dimension of coalitional violence also gains visibility through the practice of infanticide after abduction: the first offspring born to newly incorporated females tends to be killed by males, while subsequent generations are preserved. In a polygynandrous regime, there is no guarantee of paternity, and this fact tends to restrain male attacks against infants; but in the case of the absorption of foreign females, the possibility of alien paternity for the first generation is reasonable, and infanticide aims to ensure the "purity" of the patrilineal genetic community.

Lethal aggression is not exactly a rare ethological phenomenon among mammals, but the lion's share in such cases involves infanticide or the dispute over scarce natural resources, both practiced on individual level. In reproductive competition, interpersonal duels are equally common, and can result in lethality. In ethological terms, lethal aggression among adults is a very costly behavior: in circumstances of symmetry of power, it can result in the death of the victim or/and of the aggressor. Thus, lethality can be ethologically fixed as agonistic behavior when: 1) it increases the reproductive fitness of the aggressor; 2) it occurs under risk-controlled conditions. This is the rationale for infanticide, arguably the most common category of lethal violence: the risks involved will be null if infants are not protected by highly dimorphic females (which is not the case among anthropoids), by dominant males in harems or by the cooperation of multiple males in social regimes such as those of chimpanzees. Thus, what makes intersocietal coalitional violence follow the way of lethal aggression is precisely the imbalance of power, the asymmetry involved in the cooperative tactic. Ethologically, the balance of power is an effective mechanism of deterrence, and what male coalitions do is exactly breaking this equilibrium. Considering that chimpanzee raiders rarely suffer any harm, the physical elimination of competing males can bring advantageous fruits in terms of the reproductive agenda for the cooperative victors.

But can the intersocietal relations among our closest evolutionary relatives be subject to the ethology of conflict resolution, to that intuitive ethics we are talking about? Was the LCA able to behave, in the intersocietal universe, guided by cognitive instruments of violence management? In the context of a modular social intelligence - an exemplary mechanism for ordering relationships in an intragroup systemic chaos - there are clear limits related to brain capacity. Encephalization quotient [14], neocortical volume, the demography of the groups, and grooming time are associated variables. When the volume of social information produced by the increasing number of simultaneous relationships exceeds the processing limits of the modularized mind, coordination and cooperation become less and less viable. So, the larger the groups, the more social time is needed to strengthen the bonds. In these terms, greater will be the demand on the

14 The encephalization quotient expresses the ratio between mean brain volume and the expected volume for the brain in isometric (proportional) conditions with the rest of the body.

cognitive apparatus to collect information on the status of others, in order to build hypotheses about the social ascension strategies. The overload of social information makes most of the efforts to status assessment vague, creating an anomalous situation in which the instruments of conflict management lose effectiveness. Factionalism gradually emerges in the groups in response to the inability to identify correctly the status of once close individuals. As the demand for mental processing intensifies beyond the cognitive capacity of these primates, cognitive tools will continue to be pressed, something that results in pathological behavior and psycho-emotional suffering. The definitive fission of the macro-unit then acts as a homeostatic phenomenon, bringing the pressure over social mental modules to equilibrium. Once split, the two newly formed groups will have brought their demographic contingent to cognitively manageable limits back. (Aielo, Dunbar, 1993: 184-185; Aureli et al., 2008: 627; 637; Bauernfeind et al., 2013: 275-276; Ferguson, Beaver, 2009: 291; Mithen, 2002: 140-141; Wrangham, Peterson, 1996: 5-18; 158-159; 162-170; 179).

Final considerations: the ethics of warfare is an apomorphic condition in *H. sapiens*

Since the LCA, chimpanzees and most of the hominins would be unable to incorporate their intersocietal relations into the field of intuitive political ethics, something that makes the most recent developments in recent evolutionary history of *H. sapiens* quite unusual: the emergence of behavioral modernity, of transdominial consciousness and of abstract thought, about forty thousand years ago. In this long evolutionary trajectory since six millions years ago, classifying an conspecific as “foreigner” is a product of the disposal of excessive social information. The “other”, once disconnected from a social macro-unit, does not occupy any place in the internal hierarchy, and ceases to be the object of innate cognitive processes dedicated to conflict management. Somatic signals displayed by chimpanzees facing the “enemy” suggest that, unlike their ingroup fellows, aliens are treated as game animals. In an act of intersocietal coalitionary violence, aggressors make vocal and gestural signals that coincide with the act of pursuing an escaping prey. “Dechimpization”, that is, the cognitive process of re-signifying the nature of a chimpanzee conspecific, is something that seems to be universal and ethological in the lineages departing from the LCA, and not a socially learned procedure to control an supposedly innate aversion to murder. If so, it should be restricted to certain groups of chimpanzees, as are the various facets of material culture in this species. Reframing the alien’s condition functions as an ethological artifice aimed at triggering responses from the sympathetic nervous system associated with the exercise of lethal violence, such as in hunting, and this is in no way related to any defensive reaction. Chimpanzees are able to ignore the presence of other potentially dangerous primates, such as baboons, with which they occasionally compete for food. The presence of these primates does not trigger ethological reactions associated with intersocietal coalitionary violence, despite the fact that they pose a real potential threat. The baboon is not the enemy, but a chimp belonging to another social macro-unit is.

It is symptomatic that the foreigner represents the uncertainty in its highest degree, since it belongs to the universe of the “non-social”, absent either from the base and from the top of the pyramid, ignoring both dominance and submission. Being impossible to identify its hierarchical locus, there are no possible strategies to be planned concerning him, making social modular cognition useless. Averse to order and representing the outer chaos of a world devoid of instruments of conflict management, to aliens only annihilation is due, and to their females, the opportunity to be integrated in the universe of ethologically controlled social relations. Since primates rely on sensory mechanisms for identifying consanguinity - something that belongs to the personal unconscious in human mind, as potentially suggested by the Oedipus myth - we are invited to ponder about how strong environmental and cognitive pressures must be in leading to the breakup of patrilineal communities and to the "dechimpization" (or dehumanization) of the "other", with whom, once, kinship relations were recognized. When lethal violence comes to the intersocietal stage, it represents, somehow, ethological-sanctioned fratricide, in complex and unstable equilibrium with the very exercise of cooperative patrilineal sociability (Mithen, 2002: 308-309; Ferguson, Beaver, 2009: 287; Roscoe, 2007: 485-486; 491).

These cognitive mechanisms of intragroup conflict management (which, in the interaction between the transdominial consciousness, the personal unconscious and the vast ethological universe of the collective unconscious, would be called "ethical thought" in *H. sapiens*) are symplesiomorphic [15] in humans and chimpanzees, and subject to disruptions of pathological nature. Likewise, the capacity for intersocietal coalitionary violence (with claws, teeth, swords or nuclear weapons) seems to manifest itself as a symplesiomorphic condition in these two lineages departing from the LCA. What seems to be clearly apomorphic [16] is the ability of *H. sapiens* to have an ethics of warfare, the power to formulate abstract norms that determine the limits and parameters of the exercise of intersocietal violence, and, eventually, denying this own violence. Nonetheless, pacifism and the norms of warfare, in this case, seem far from being an ethological condition, deposited in the human collective unconscious: they depend exclusively on the exercise of transdominial consciousness. On the other hand, intergroup violence and dehumanization, although strongly opposed by domains of the conscious mind, find firm rest in the recesses of the collective unconscious, inherited from a turbulent evolutionary past.

15 Symplesiomorphy is a primitive characteristic shared by two or more species that is not crucial to define any of these species in particular. *H. sapiens*, for example, has no tail, such as gorillas; therefore this is not a condition that defines modern humans or gorillas.

16 Apomorphy is an innovative characteristic present in a particular species, that makes it differs from all its ancestors. Bipedal locomotion is probably an apomorphy in the lineage of the hominins, that makes them distinct from their last common ancestor with chimpanzees (and from chimpanzees themselves).

Bibliography

AIELLO, L., DUNBAR, R. Neocortex size, group size, and the evolution of language. *Current Anthropology*, v. 34, n. 2, p. 184-193. 1993. DOI: 10.1086/204160.

AURELI, F. et al. Fission-fusion dynamics: new research frameworks. *Current Anthropology*, v. 49, n. 4, p. 627-654. 2008. DOI: 10.1086/586708.

BARNOSKY, A. e KRAATZ, B. The role of climatic change in the evolution of mammals. *BioScience*, v. 57, n. 6, p. 523-532. 2007. DOI: 10.1641/B570615.

BAUERNFEIND, A. et al. A volumetric comparison of the insular cortex and its subregions in primates. *Journal of Human Evolution*, v. 64, n. 4, p. 263-279. 2013. DOI: 10.1016/j.jhevol.2012.12.003.

BRAUDEL, Fernand. *Escritos sobre a História*. São Paulo: Perspectiva, 2009.

CAMERON, D., GROVES, C. *Bones, stones and molecules: 'out of Africa' and human origins*. San Diego: Elsevier, 2004.

CHRISTIAN, David. Macrohistory: the play of scales. *Social Evolution & History*, v. 4, n. 1, pp. 22-59, 2005.

FERGUSON, C. e BEAVER, K. Natural born killers: the genetic origins of extreme violence. *Aggression and Violent Behavior*, v. 14, n. 5, p. 286-294. 2009. DOI: 10.1016/j.avb.2009.03.005.

FOLEY, R. *Os humanos antes da humanidade: uma perspectiva evolucionista*. São Paulo: UNESP, 2003.

LADEIA, I., FERREIRA, P. A história evolutiva dos primatas. In: Neves, W. et al (orgs). *Assim caminhou a humanidade*. São Paulo: Palas Athena. p. 48-85. 2015.

MITHEN, S. *A pré-história da mente: uma busca das origens da arte, da religião e da ciência*. São Paulo: UNESP. 2002

OTTERBEIN, K. A history of research on warfare in anthropology. *American Anthropologist,* v. 101, n. 4, p. 794-805. 1999. DOI: 10.1525/aa.2000.102.4.834.

PAMPUSH, J. et al. Homoplasy and thick enamel in primates. *Journal of Human Evolution*, v. 64, n. 3, p. 216-224. 2013. DOI: 10.1016/j.jhevol.2013.01.009.

NORDHAUSEN, M., OLIVEIRA Filho, P. Nós, primatas. In: Neves, W. et al (orgs). *Assim caminhou a humanidade*. São Paulo: Palas Athena. p. 14-47. 2015.

ROSCOE, P. Intelligence, coalitional killing, and the antecedents of war. *American Anthropologist*, v. 109, n. 3, p. 485-495. 2007. DOI: 10.1525/aa.2007.109.3.485.

SHULTZ, Susanne; OPIE, Christopher; ATKINSON, Quentin. Stepwise evolution of stable sociality in primates. *Nature*, v. 479, n. 7372, p. 219-222. 2011. DOI: 10.1038/nature10601.

WRANGHAM, R., PETERSON, D. *Demonic males: apes and the origins of human violence*. Boston: Mariner, 1996.

The Great Matrix of All Beings

William Grassie

ABSTRACT

Learning "the big ideas of all the big disciplines," as famed investor Charlie Munger suggests, is not really all that hard. Indeed, he claims it is both fun and profitable. He recommends that we build "a latticework of mental modules." Science offers a sixth sense—a way of seeing beyond the walls of Plato's cave into the realm of what is real. It isn't rocket science to sort-of-know how things actually work, how the pieces form an entire puzzle. Of course, god and the devil are both in the details, but it sure helps to have the picture on the puzzle box when trying to fit the pieces together.

Correspondence | William Grassie, william@grassie.net
Citation | Grassie, W. (2018) The Great Matrix of All Beings. *Journal of Big History*, II(2); 37 - 52.
DOI | http://dx.doi.org/10.22339/jbh.v2i2.2301

Learning "the big ideas of all the big disciplines," as famed investor Charlie Munger suggests, is not really all that hard. Indeed, he claims it is both fun and profitable. He recommends that we build "a latticework of mental modules."[1]

Another word for lattice is matrix. A matrix is a structure, medium, or environment in which something develops. Matrices can be physical, biological, social, and mathematical. It can be a mold that shapes organisms, for instance, with intercellular matrices in our body. In mathematical usage, a matrix is like a spreadsheet—a grid of quantities and formulas, charts and comparisons, statistics and accounts, or more abstractly, geometric possibility spaces. The etymology of the word *matrix* is from the Latin for "womb" or "mother." We might also title this chapter "the Great Womb of All Beings."

In all of these senses, including perhaps especially the creative maternal, we begin by considering the Great Matrix as discovered and elucidated by science. I find this metaphor and mnemonic a useful way to navigate a lot of science in diverse disciplines, as well as to introduce important philosophical issues into the Big History curriculum.

Everything that exists in the universe, every process that science has discovered, every power of nature, all that constitutes our human bodies and brains, our histories and cultures—all this and more—can be located within a number of natural hierarchies and scales—size, time, matter, energetics, electromagnetism, sound, information-ingenuity, sentience-consciousness, cultural constructs, and emergent complexity. There are other scales and hierarchies, to be sure, but these ten natural scales are the warp and woof of the universe upon which all complexity is woven—is the first step to applying Big History to understanding "all the big ideas of all the big disciplines."

1) Scales of Size

It takes some effort to grok the scales of size from the macrocosm to the microcosm as discovered by science.[2] How to internalize the new scales of science that extend far beyond our direct sense perceptions? In that sense, most of science is by definition counter-intuitive. You can't see much of what science is about—it's too large or too small, too fast or too slow. The existentialist in you may stay awake at night pondering how big the universe is, but it is perhaps more instructive to pay attention to how small entities can also be. The known universe is as many orders of magnitude smaller than us as it is larger. As physicist Richard Feynman observed, "there is a lot of room at the bottom."

The smallest unit in particle physics is the Planck

1 Charles Munger, "Usc Law School Commencement Address, May 1, 2007," Genius.com, http://genius.com/Charlie-munger-usc-law-commencement-speech-annotated.

2 The word "grok" was a coined by Robert Heinlein in his 1961 science fiction novel *Stanger in a Strange Land*. The word is now listed in the Oxford English Dictionary, meaning "to understand intuitively or by empathy, to establish rapport with" and "to empathize or communicate sympathetically (with); also, to experience enjoyment." https://en.wikipedia.org/wiki/Grok

scale (1.62 x 10^{-35} m). Beyond the Planck scale the concepts of size and distance break down and quantum indeterminacy becomes absolute—whatever that means. You will need to find a particle physicist to make more sense of these subatomic realms in which matter turns into miniscule fields and forces.

From the Plank scale we jump about 24 orders of magnitude in size pass the scale of subatomic particle to the scale of simple atoms. The diameter of a single hydrogen atom—one proton, one electron, and a lot of empty space—is approximately 0.1 nm (nanometers). Just to be clear, one nanometer (10^{-9}m) is a billion times smaller than one meter.

Amino acids—small molecules essential to cell chemistry and varying in atomic composition —are approximately 0.8 nm in size.

DNA, the molecule deoxyribonucleic acid, is about 2 nanometers wide but varies dramatically in length. The longest human DNA is found in chromosome 1 and consists of 220 million base pairs. When stretched out it would be about 85 millimeters long and very, very thin. In terms of length, human DNA has some 3 billion base pairs, but that's nothing compared to the marbled lungfish, which clocks in at a 133 billion base pairs. At the other end of the DNA spectrum is the diminutive bacteria *Carsonella ruddii* with the smallest known genome containing 159,662 base pairs. DNA, as we will explore below, has an aperiodic property that allows it to encode information for cell construction and replication.

Prokaryotes—single-cell bacteria and archaea, more about that later—are the smallest and ur-unit of life. The smallest bacteria are about 150-250 nm. The bacterium *E. coli*, part of your intestinal microbiome, measures up at approximately 2 µm (micrometers or microns). Here we have moved up in the scale a thousand fold, from nanos to microns, from 10^{-9} m to 10^{-6} m.

It takes imagination to grasp these orders of magnitude jumps in scale. At the other end of the size spectrum, we measure in light-years, which in spite of its name is a measurement of distance not time. One light-year is the distance that a photon travels in one Earth year—9.4 trillion kilometers in case you were wondering. The most distant thing that we know observationally in the universe is the background radiation from the big bang—13.8 billion light-years away from earth. 9.4 trillion kilometers times 13.8 billion is a really big number and an incomprehensible distance for mere humans to grasp—and yet we now do—though difficult to grok.

To get a feel for the relative size and scale of the universe, imagine holding a big beach ball one meter in diameter. It is summer and you're vacationing on Cape Cod, Massachusetts. If this beach ball were the size of the sun, then the next closest star, Proxima Centauri, would be another beach ball 6609 km away—the flight distance from Boston to Rome (the actual distance is 4.2 light years).

On the other hand, if your one-meter diameter beach ball were the size of our Milky Way galaxy, then the next closest galaxy—the Andromeda galaxy—would be another beach ball 25 meter away (the actual distance is 2.5 million light years). Indeed, at the scale of the galaxies, everywhere you looked on the beach that day, above, below, and all around, you would see beach ball galaxies in all directions—a 100 billion of them. In scale, galaxies are quite close to each other and gravitationally bound in a vast moving web, while the distance between stars within the galaxies is enormous and for the most part gravitationally not interactive with neighboring stars. When galaxies collide, as Andromeda and the Milky Way some day will, they pass through each other and nary a star will collide. The universe is a web of a hundred billion galaxies held together in a dynamic gravitational embrace.

The tendency is to focus on how puny we are in the scale of hundreds of billions of galaxies spanning billions of light years, but we should also remember how enormous we are when compared with the cells, molecules, atoms, and subatomic particles. Humans exist at scales measured in millimeters, centimeters, meters, and kilometers, which turn out to be about halfway between the very small and the very large. The human scale is also the only scale where certain types of complexity could exist.

The Cosmic Uroboros is a graphic representation of the hierarchy of size. The image of a snake eating its own tail is a symbol appearing in ancient Egyptian, Greek, Norse, and Indian art. The symbol is thought to represent self-reflectivity and self-regeneration. In a rendering of the uroboros by cosmologist Joel Primack and co-author Nancy Abrams, the hierarchy of size is plotted in powers of ten along the circle.[3] The universe is swallowing its own tail, because the very large scale of the universe loops back on the very small scale of subatomic particles. The particle accelerator at CERN recreates the intense energy of the early universe in order to study the constituents of matter. The sought-after "Grand Unified Theory" of fundamental physics (GUT in the image) would unite the four fundamental forces—gravity, electromagnetism, and the strong and weak atomic forces—connecting the subatomic and the cosmic dimensions, much as uroboros swallows its tail.

The first axis in the Great Matrix of all beings is size. Note again the curiosity and perhaps necessity that life exists halfway between the largest and smallest entities known to science. And take a moment also to cogitate and meditate on how your own being is built and functions bottom-up from small to ever larger combinatorials of atoms, molecules, and cells.

2) Scales of Time

The time scales of the universe are measured today in billions of years down to the nanosecond vibrations of cesium in atomic clocks. We have discovered a progressively true account of the 13.8-billion-year history of the universe, the 4.5-billion-year evolution of our planet, the 300,000-year rise of our species, and the 10,000-year accelerating drama of human civilization. This chronology is how Big History is generally approached in books and curricula, telling the story as an integrated narrative of complexity thresholds from the big bang to the 21st century.

In brief, our omnicentric universe began some 13.8 billion years ago in a fantastically small, dense, hot, and symmetrical state. This universe expanded, cooled, and evolved into more differentiated and complex structures—forces, quarks, hydrogen, helium, galaxies, stars, heavier elements, complex chemistry, and planetary systems.

Some 4 billion years ago, in a small second- or third-generation solar system, the intricate biophysical processes of life began on at least one small planet circling a diminutive, slow burning star. Animate energy-matter on Earth was a marvelous new intensification of the creative dynamic at work in the universe. Life adapted, evolved, complexified. Fuel by the power of the sun, Earth became a jewel of a planet with ever more complex forms and relations giving rise to the wonders of the natural world.

Then some 7 million years ago, proto-humans emerged on the savanna of Africa. By a million years ago, they mastered fire and made simple tools. About 80,000 years ago, a small band of *Homo sapiens* wandered out of Africa. With their tools, fire mastery, language, and collective learning, our ancestors spread across the planet. We are the only large mammals to inhabit so many diverse bioregions.

For most of our prehistory, our common ancestors were hunter-gatherers, living in small tribes, probably not larger than 150 individuals.[4] Our brains, instincts, emotions, and behavioral tendencies evolved to promote survival and reproduction in these social groups in diverse environments. Then 10,000 years ago, agriculture began, and with it growing populations of humans living in ever larger and more complex societies accumulating knowledge and knowhow, trading in genes, crops, animals, diseases, technologies, and ideas. The fossil fuel revolution began about 200 years ago and with it amazing growth in wealth and productivity. This unfolding leads us all the way to today, over 7

3 Joel R. Primack and Nancy Ellen Abrams, *The View from the Center of the Universe: Discovering Our Extraordinary Place in the Cosmos* (New York: Riverhead, 2006).

4 Robin Dunbar, *Grooming, Gossip and the Evolution of Language* (Cambridge: Harvard University Press, 1996).

billion of us collectively transforming the planet and ourselves. The wonder of it all is that each of us is also a collection of transient atoms, recycled stardust become conscious beings, engaged in this global conversation brought to you by ephemeral electrons and photons cascading through the Internet and bouncing off of satellites.

That is Big History, in about 300 words and with a little poetic license. The story is generally understood in thresholds of emergent complexity—the early universe gives rise to stars and galaxies, the stars give rise to the periodic tables of elements and complex chemistry, the chemistry gives rise to simple bacteria, the bacteria give rise to eukaryotic cells, which complex cells give rise to the evolution of organisms, the plants and animals give rise to complex ecosystems with ever more sentient organisms, and sentience in our lineage gave rise humans consciousness, tool-making, language, fire mastery, collective learning, agriculture, and the modern revolution. Throughout the chronology, increased complexity is hierarchical, built from the bottom-up. Each threshold realizes new emergent properties that previously did not exist.

Einstein's theory of relativity, of course, links space and time in a continuum, so separating chronology and size scale is technically an artifact of our minds and not fundamental physics. Moreover, many of the equations of fundamental physics work forward or backward irrespective of time. And when we talk about the very fast, the very dense, and the very hot, the very cold, concepts of space and time can become elastic, but in between these extremes, time matters. For practical purposes, it is necessary to keep space and time separate. Together space and time provide the first two dimensions in the Great Matrix.

3) Matter

Technically, matter is solidified energy, as reflected in the iconic equation E = mc2—energy equals matter times the speed of light squared. However, just as with space and time, matter and energy may reliably be understood as distinct, at least in the context of the Great Matrix that matters most to us here on Earth.

Particle physicists advance our knowledge of smaller and smaller realms. Cosmologists advance our knowledge of large and larger realms. Neither has all the answers and their best working theories—quantum mechanics and relativity—are mutually incompatible. The scale of atomic elements, however, is well established and central to all chemistry.

Matter turns out to be a scarce commodity in the universe. The universe is composed of about 70 percent cold dark energy. Another 25 percent consists of cold dark matter. Cosmologists don't really know what these are, ergo cold and dark, but infer their existence from the behavior of the web of galaxies in an expanding universe. Along with the cold dark energy and matter, there are invisible atoms. Hydrogen and helium, the atoms that fuel the stars, constitute about 0.5 percent of the universe. All the other visible atoms in the Periodic Table of Elements constitute less than 0.01 percent of the universe. Fortunately, we live in a chemically rich solar system on a very special planet. We are made of these remarkable and rare elements in this vast universe.

The character of elements is determined by the number of protons in their nucleus, along with isotopic variations in the number of neutrons and electrons bound to that nucleus. The discovery of periodicity and groupings of elements based on shared properties is one of the most remarkable achievements of science leading to all manner of new discoveries in physics and biology, not to mention engineering and technology. Chemistry may not have the same academic sex appeal as physics and biology, but it is the foundation upon which evolutionary and economic complexity builds.

4) Scales of Energetics

The intensity of energy flow is another axis in the Great Matrix. We can metaphorically equate energy with money. Energy does for nature what currencies do for human production and trade. And like all metaphors, it is literally not true. When Shakespeare calls time a beggar, he reminds us of our human finitude, but time is not literally a beggar. Metaphors are also important to science. Terms from computer science, for instance, are helpful in thinking about genetics, the neuroscience, and other fields.

Energetics is the currency of the Universal Central Bank, as guaranteed by the Laws of Thermodynamics. All of the regional banks—astronomy, chemistry, biology, technology, economics, culture—can create their own energy currencies and markets because of this omnipotent guarantee. The Universal Central Bank does not offer fixed rates of return, but rather variable rates of losses in a predictable flux that impacts every domain at every scale. Energy dissipates. Entropy grows. That's not a problem; that's the singular opportunity.

The First Law of Thermodynamics observes that energy can be transferred and transformed, but not created or destroyed. At the first instance of the big bang, the universe received its lifetime allotment of energy-matter. That energy allowance is still present, just redistributed and less concentrated as the universe expands.

The Second Law of Thermodynamics states that in a closed system energy always moves for hot to cold, from order to disorder, from concentrations to entropy. What concerns us is not energy per se, but the flow of energy and the work that is done along the way.

Energy is the ability to transform a system—to do work. The Standard International unit of energy is the joule—defined as the energy required to lift a 1 kilogram object 1 meter vertically from the Earth's surface (i.e. against a gravitational force of 1 Newton). That energy is also released when that object falls back down. Lifting the object is mechanical work. The falling back down is kinetic energy—the release of potential energy in a gravitational field.

Physics distinguishes energy and power. Power is energy used over time. The Standard International unit of power is the Watt defined as 1 Joule per second. That 100 Watt light bulb is consuming 100 Joules per second, about the same amount of energy that your body is consuming, albeit in a different form.

Energy comes in different forms and types—electromagnetic, chemical, thermal, kinetic, electrical, nuclear, gravitational—to name the major categories. There is no standard currency for energy in nature and society, though there are lots of standard equations for converting one energy form into the equivalent of another energy form.

For instance, we can convert the combustion of gasoline into the equivalent of human labor—11.6 days of labor to be precise. These calculations require multiple energy conversion, turning BTUs into kilowatts into 2500 food calories per day. We cannot, however, digest gasoline. It is a poison. The energy content of a single gallon of gasoline is equivalent to 28,977 food calories of energy or 51 Big Mac hamburgers to be precise.[5]

These kinds of conversions neglect important qualitative differences in energy types, but are nevertheless informative. Energy flow (i.e. power) is not the gasoline or the hamburgers—they are stored, potential chemical energy. Energy flow is the controlled burning of hydrocarbons by the machine or the body. A gallon of gasoline has a lot more potential chemical energy, than a gallon of you (which is mostly water), but the whole of you is able to harness the BTUs, Kilowatts, and Calories to amazing effects. We are the only entity that we know of that *constructs* complexity outside of our bodies in built devices using external energy sources. It goes all the way back to the controlled use of fire. The Prometheus myth got it right. Fire mastery is a god-like gift that separates us from all other animals.

The universe burns and humans have captured some of that energy. Physicists calculate the energy of the universe at the moment of the Big Bang as 10^{19} GeV (billion electron volts). At the opposite end is absolute zero or minus 273.15 degrees Celsius (minus 459.67 degrees Fahrenheit).

All complex phenomena in the universe can be

5 One gallon of gasoline = 33.70 kWh = 28,977 kcal
28,977 kcal @2500 kcal per day = food for 11.6 days of human labor
28,977 kcal = 51 Big Macs @ 563 kcal per hamburger

characterized by a flux of energy from hot to cold, from order to disorder, as required by the Second Law of Thermodynamics. Life excels at capturing energy to drive its creative processes.

One of the most counter-intuitive insights of modern science is the scale of energy density flow. Astronomer Eric Chaisson is credited with these calculations and insights.[6] He normalizes for the mass and time frame of each energetic system (ergs per second per gram). It turns out that a single-cell eukaryote has an energy density flow 500 times that of our sun. A photosynthesizing plant has about 5,000 times the energy density flow of the sun. A mammalian body has about 20,000 times the energy density flow of the sun. The human brain, consisting of about 2 percent of our body weight but consuming about 20 percent of our food energy, has an energy density flow about 75,000 times greater than the sun.

Of course, the sun is enormous and the flux of energy from the sun to the cold of outer space—3.86 x 10^{26} watts—far exceeds that of our human bodies running on the equivalent of a 100-watt light bulb. Remember that we are normalizing for the mass of the system and its time scale. What this means is that if you could expand the human body—ca. 70 kilograms—to the mass of the sun—1.9 x 10^{30} kilograms, then it would be 20,000 times more energetic (assuming you could feed it enough hamburgers).

And if we include all of the energy consumed outside of our bodies in our global civilization, then humans today achieve energy density flows average 250,000 times that of the sun. The privileged few of us who fly around the world—the energy rich—achieve energy density flows millions of times that of the sun. As we will further explore below, this exponential growth in energy density flow extends to the evolution of our technology. Humans and our advanced machines may well achieve the greatest sustained energy density flows in the entire universe.

Remember that there are three ways to increase energy density flow: 1) you can consume more energy more efficiently, 2) you can compress the time scale of the system, and 3) you can decrease the mass of the system. With too little energy (e.g., starvation) the system will collapse. Too much energy flow and the system will also crash and burn. Each complex adaptive system has a range of optimal energy flows that are often limited within a narrow range, a kind of Goldilocks parameter.

There have been tremendous and accelerating increases in efficiency of our machines and technologies over the last 200 years. Much of this efficiency gain has been accomplished by reducing the mass of the system with new materials and technologies, for instance in the miniaturization that has occurred in electronics and computers, as well as more effective combustion processes with less heat wasted.

The increases in efficiency over the last hundred years are stunning. And yet overall consumption of energy continues to increase even as the technologies become more efficient. Our cars are more efficient, but more of us drive more miles. Our planes are more efficient, but more of us fly more miles. Our furnaces and air conditioners are more efficient, but we have bigger houses. These gains in efficiency mean increases in energy density flow. Energy efficiency is as important today as energy production, or for that matter labor productivity, because the reduced input can now be deployed elsewhere in creating additional economic growth and prosperity—more on that later. For now, we welcome energy density flow as the third axis in the Great Matrix.

The others scales below—light, sound, information-ingenuity, sentience-consciousness, evolved cultural hierarchies, and emergent complexity—are themselves forms of energy flow, a subset of energetics, each with unique properties essential to the evolution of complexity in nature and economics.

6 Eric Chaisson, *Cosmic Evolution: The Rise of Complexity in Nature* (Cambridge, MA: Harvard University Press, 2001); *Epic of Evolution: Seven Ages of the Cosmos*; "Energy Rate Density as a Complexity Metric and Evolutionary Driver," *Complexity* 16, no. 3 (2011); "Energy Rate Density. Ii. Probing Further a New Complexity Metric," *Complexity* 17, no. 1 (2011).

5) Scales of Electromagnetism

Electromagnetism is one of the four fundamental forces in physics about which there is an active search for a Grand Unified Theory (GUT) that would combine these four forces in a single equation. In this exercise, we are going to ignore gravity and the two nuclear forces, because electromagnetism governs almost all of the phenomena that we encounter in daily life here on Earth.

Normal chemistry is all about the affinities of electrons and the indestructible elements of the periodic table. Negatively charged electrons are bound by electromagnetic attraction in orbitals around positively charged atomic nuclei that make up the periodic table of elements. Atoms combine into complex molecules through electromagnetic geometries and preferences. All chemistry, and therefore all biology, is governed by electromagnetic force. The ATP molecules in your cells, the neurons in your brain, the gasoline burning in your car, the food you eat, and all the electronic devices in your life—from the light bulb to the Internet—all utilize electromagnetic properties.

Electromagnetic radiation has three properties—frequency, wavelength, and photon-energy. Frequency is measured in Hertz (Hz) in a range of 104 to 1030 oscillations per second. Wavelength is measured in meters in a range of 10-12 to 106. Energy is measured in electronvolts (eV) in a range of 106 to 10-15. The entire spectrum of electromagnetic radiation spans from long and slow oscillating radio waves at one end to short and fast oscillating gamma waves at the other end.

While the electromagnetic spectrum is a single, continuous physical phenomenon, our human eyes have evolved to perceive only a small range of visible light. We can feel infrared radiation on our skin, but the rest of the electromagnetic spectrum is invisible to our direct senses.

The entire spectrum is divided into ionizing and non-ionizing electromagnetic radiation (EMR). Short-fast EMR interacts energetically with atoms and molecules to break and remake chemical bonds. Long-slow EMR reacts minimally with most matter. The break between ionizing and non-ionizing EMR has to do with the size of atoms and the wavelength of the photons. Ultraviolet EMR gives us sunburn. We use short-high frequency EMR to heat our food (microwave), look through our bodies (x-rays, magnetic imaging), and kill cells (Gamma ray).

Visible light is tucked in a sweet spot just below the range of the ionizing radiation. A prism famously splits white light into its constituent rainbow colors, each with their own specific range of wavelengths and frequency. Beyond purple we move into ultraviolet radiation (<380 nm, <400 terahertz). As the wavelengths get shorter, the frequencies get faster. Beyond red on our rainbow, we move into the infrared (>760 nm, >790 terahertz). The wavelengths get longer. The frequency slower.

Very high frequency (VHF) radio waves—used in television, radios, and other communication devices—range from 1 to 10 meters in wavelength and vibrate at a rate of 30 to 300 megahertz (MHz) (millions of oscillations per second). Electromagnetic radiation in the VHF band is long and interacts minimally with the atoms in the walls of your building or the cells in your body. It takes a special machine—your smartphone or radio—to tune into specific channels, to decode the broadcast woven in airwaves, and to convert that signal into sound or image for your enjoyment and edification.

Your eyes are basically a radio receiver at a different frequency. The photoreceptor cells in your eyes tune into specific ranges of electromagnetic radiation. Like the cells themselves, visible light is necessarily in the nanometer range in order to interact with nanometer-sized light-sensing organelles in the cells.

Electromagnetic radiation is central to all of the prosthetic "seeing" devices of science and technology — from radio telescopes to electron microscopes, from smartphones to global communications. These tools extend our vision, our abilities to see, hear, touch, taste, smell, and understand. What we know and know how to do goes far beyond our five natural senses. All of the prosthetic devices that extend human perception utilize the electromagnetic effect.

All of the machines and motors that allow us to cross continents and move mountains utilize the electromagnetic effect.

Electromagnetic radiation, as recently harnessed by humans, is magical in how it has transformed our lives and our understanding of the universe. It is a continuous spectrum of wavelengths and frequencies, though divided into qualitatively different segments—radio waves, microwaves, infrared, visible light, ultraviolet, x-ray, and gamma ray. The spectrum of electromagnetic radiation is the fourth axis in the Great Matrix of being—a form of energy flow and a critical subset of all energetics.

6) Scales of Sound

Sound is a vibration that propagates in a medium—gas, liquid, solid, or plasma. Unlike light, sound cannot travel through a vacuum; but like light, it has wavelength-frequency, directionality, intensity, and its own distinct speeds. Sound travels at the rate of 343 meters per second in dry air at 20oC. Sound vibrations jostle molecules in a wave of kinetic energy, much like ripples on still water. Sound, like light, is also a subset of energetics.

An evolving capacity to sense pressure vibrations in the environment certainly has adaptive value for evolving organisms. Rudimentary "hearing" allowed organisms to receive directional information at a distance, thus improving survival and reproduction, while eventually leading also to forms of animal communications.

Curiously, the continuum of sound audible to humans ranges from 20 Hz to 20,000 Hz (17 mm to 17 m), in contrast to the narrow range of visible light, which has no leaps in order of magnitude (790 terahertz to 400 terahertz—760 nm to 380 nm). Humans have much greater "depth" of sound perception than we have "depth" of light perception.

In any case, sound is a physical property of the universe, it has a hierarchical scale, and thus it provides yet another axis in the Great Matrix. Sound may not travel through outer space, but it is an essential component of evolution. And for humans, in particular, sound is central to our perception, communication, cooperation, enjoyment, survival, and reproduction.

7) Hierarchies of Information-Ingenuity

We might also postulate a scale of information-ingenuity in the Great Matrix, even though we don't have a universal metric or even proper definition of information adequate to all scientific disciplines. Information—seemingly immaterial and ephemeral—can be a slippery metaphysical concept. In the broadest, most abstract sense, however, information is simply physical order in a universe that marches to the tune of disorder.

Out-of-equilibrium energetic systems spontaneously create physical order.[7] This is how the weather works—a global system for dissipating the heat from the sun—most dramatically in hurricanes and tornados. The out-of-equilibrium physical order of the weather is information in the service of entropy. In gases and fluids, however, the spontaneous physical order is fleeting. It arises. It disappears. The weather is always changing. As Heraclitus observed 2500 years ago, you can't step into the same river twice.

Solid matter, however, can store information, at least for a time. Solid matter has memory. Its physical order can persist in ways that fluids and gas cannot. Solid matter, moreover, can also process that information to rearrange matter, energy, and information to make other stuff. In his book, *Why Information Grows*, Cesar Hidalgo calls the ability of matter to compute "one of the most amazing facts of the universe." (2477) Hidalgo writes:

> Information is not tangible; it is not a solid

7 Ilya Prigogine was awarded the 1977 Nobel Prize in Chemistry "for his contributions to non-equilibrium thermodynamics, particularly the theory of dissipative structures." Ilya Prigogine and Gregoire Nicolas, eds., *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations* (New York: Wiley, 1977); I. Prigogine and Isabelle Stengers, *Order out of Chaos : Man's New Dialogue with Nature* (New York, N.Y.: Bantam Books, 1984).

> or a fluid. It does not have its own particle either, but it is as physical as movement and temperature, which also do not have particles of their own. Information is incorporeal, but it is always physically embodied. Information is not a thing; rather, it is the arrangement of physical things. It is physical order, like what distinguishes different shuffles of a deck of cards. (Hidalgo, 147)

If physical order is information, and if it takes energy to transform matter, and if it takes information to specify particular ordered states, then it takes energy to make and translate that information in, out, and back into matter. Encoding and reading information is thus also a subset of energetics, and thus governed by the Second Law of Thermodynamics. Information-ingenuity is not free, but it can be cheap. With information-ingenuity, evolution can minimize entropy and maximize creativity. Evolution offers the possibility of doing more with less through more efficiently coding.

DNA is a paradigmatic case of matter encoding and computing information. In his famous 1944 Dublin lectures and later book, *What Is Life?,* the physicist Edwin Schrödinger postulated that the code of life inside the cell needed to be an aperiodic crystal. Most crystals are periodic, meaning they form highly ordered microscopic structures. The molecules self-organize, based on their electromagnetic affinities, as they transition from liquid to solid in a tightly packed geometric lattice. These microscopic patterns then grow into macroscopic structures—diamonds, snowflakes, table salt, and metals of all kinds.

In an aperiodic crystal, however, variations in the molecular structure provide the possibility of coding information. Such is the molecule DNA—two strands of long molecules—polynucleotides—connected by variable bonds of adenine (A) to thymine (T) and cytosine (C) to guanine (G). The DNA molecule does not care about the actual order of these chemical bonds, only that A binds with T and C binds with G along the backbone of the two molecular strands. The pattern of these base pairs— A-T and C-G bonds—is the template upon which DNA codes RNA and RNA codes proteins inside the cell. In cell division, the DNA replicates itself. Solid matter can encode information-ingenuity and then compute that information-ingenuity into living, reproducing, and evolving organisms.

In nature, DNA is always in solution inside the cytoplasm of cells. In a laboratory, however, biologists can separate and concentrate the DNA and then watch as the DNA molecules self-organize into crystalline forms. Because DNA is aperiodic, the crystals vary greatly, creating a profusion of psychedelic patterns under the microscope. (Image)

As scientists have developed techniques for manipulating DNA molecules, they have successfully used DNA to encode and decode digital information. It is slow and expensive work, but the potential is enormous. DNA can store orders of magnitude more data by volume than current computer hard drives with less energy and potentially over much longer time frames. A cubic millimeter of DNA can contain 5.5 petabits (10^{15}). As of 2016, a single kilo of DNA would be sufficient to store all of the world's digital data.[8]

In this technological feat—copying digital information in and out of DNA molecule—we encounter the dual nature of information that causes a lot of confusion. On the one hand, there is information as code. On the other hand, there is information as message. The code has no meaning in itself; it is merely an "arbitrary" vehicle for transmission. Information as message, however, is all about some meaning—construct this protein with this sequence of amino acids or interpret this string of A-T and C-G bonds as binary code for storing the digital contents of all computers. So we have two distinct definitions of information to keep separate, but that are also always connected.

In both senses—information as code and information as meaning—we might well imagine

8 Andy Extance, "How DNA Could Store All the World's Data," Nature, http://www.nature.com/news/how-dna-could-store-all-the-world-s-data-1.20496.

a scale in which information grows exponentially. Information measured as DNA base pairs, for instance, grows by orders of magnitude from prokaryotes to eukaryotes. Information grows again through multicellular organisms and with the evolution of the five senses—touch, taste, smell, sight, and sound—new codes conveying new meanings. And with the rise of symbolic language—spoken and later written—information as code and information as meaning experiences another exponential leap. All the while information is specifying and computing physical and social order, a continuous flux of creativity through our bodies, brains, and global civilization.

Ingenuity is simply information that does something creative, innovative, useful, and intelligent. The Latin root of the word "ingenuity" means "inborn," as in inborn genius. Indeed, the natural world is full of inborn genius. The inborn intelligibility-intelligence of nature is the precondition for scientific discovery and human technologies. We are surrounded by what philosopher Daniel Dennett refers to as "competence without comprehension."[9] Most of our biological functions, including mental functions, occur without comprehension. Human consciousness and comprehension are recent development in evolutionary history and an ongoing story. We have only recently begun to comprehend our true place in the universe.

In computer science the smallest unit of information is a bit—0 or 1. Technically, random garbage on your hard disk contains more "information" than your photos, music, documents, and applications. The latter are encoded in more compact algorithms. No such luck with random 1s and 0s. By the information-as-code definition, disorder contains more "information" than order.

This seemingly esoteric discussion of information theory turns out to be central to economics and finance. The rise of complexity in nature and culture concerns us, not simply the amount of code required to describe a particular state of physical order or disorder.

Humans are the most amazing computers of information-ingenuity, but we also encounter limits. Cesar Hidalgo introduces the concept of personbytes to represent the maximum amount of knowledge and know-how that an individual can acquire. Firmbytes are the maximum amount of knowledge and know-how that a firm can acquire. It takes time and effort to acquire expertise and skill. There is a limit to how much knowledge and know-how one individual or one firm can acquire. And to accomplish complicated manufacturing or provide complicated services requires a lot of personbytes and firmbytes in collaboration. These limits are why complex economic activity requires networks of firms working in tandem with complex supply lines. Acquiring information-ingenuity is hard work, but it plays a central role in all disciplines and professions. Humans have taken the computing of energy, matter, and ingenuity to a whole new level.

So we imagine inductively another scale in the Great Matrix—a scale of information-ingenuity—the code and the meaning—that also grows exponentially. In evolution, it grows both as the quantity of DNA code, but also as the diversity of ingenious life forms. In human culture, it grows both as the quantity of code—spoken, printed, broadcast, and digital code—and as the quality of ingenuity—knowledge and know-how that accumulate through collective learning across generations and across geographies.

8) Hierarchies of Sentience-Consciousness

We might postulate yet another axis in the Great Matrix: a hierarchy of sentience leading to consciousness. The membranes of microbes already have rudimentary sentience, in so far as they seek out food sources and flee harm. The immune system is also sentient, goal directed in response to its environment. The evolution of the central nervous system and the five senses further increased sentience. Animals evolve feeling and emotions—an amazing intensification of sentience. With the

9 Daniel C. Dennett, *From Bacteria to Bach and Back* (New York: W.W. Norton, 2017).

advent of symbolic language and collective learning, humans brought about another exponential leap in the scale of sentience-consciousness with intense first person subjectivity and social intersubjectivity. Free will is probably not the right term. Instead, we have constrained choice, but at an extent orders of magnitude beyond any of our animal kin.

Across the variety of animal species, brain-mind is an emergent phenomenon and potentially scalable. While there is no numeric scale for feeling and perception, sentience and consciousness, we can identify the ability of single cells to "sense" the presence of food or predators in the environment as perhaps the low end of the sentience spectrum in its most rudimentary form. A roundworm in a neuroscience lab has only a few hundred nerve cells, while a human brain has billions nerve cells. Surely, there are objective differences in body-brain-mind complexity between round words and humans. Perhaps some day we will have such a scale throughout the animal kingdom.

Counting nerve cells alone, however, does not really give us an adequate measure of brain-minds. Our human brain-minds require bodies and metabolism, vocal chords and oppositional thumbs, and an enriching social and natural environment, in order to realize their potentials.

While closely related to information-ingenuity, I treat sentience-consciousness as separate phenomena dealing primarily with the subjectivity, behavior, and agency of creatures. Human subjectivity and consciousness are on a kind of continuum throughout an evolutionary hierarchy. Sentience-consciousness is also a subset of energetics; and like our other scales in the Great Matrix, I imagine exponential leaps occurring throughout the drama of Big History.

Perhaps someday we may have a robust measure of sentience-consciousness that will allow us to compare dogs with cuttlefish, elephants with birds, and smart phones with smart people.

9) Culturally Constructed Hierarchies

Humans construct social and cultural hierarchies and scales that exist nowhere else in nature. These hierarchies have structured societies for good and ill throughout history. Some of these are evolved dispositions that we share with other animal species. Primate dominance hierarchy behavior may be deep in our genes, for instance, but the organizational flow charts at your company are not. All complex societies and institutions have social hierarchies reflected in their customs and laws, in their mating and childrearing practices, in and their occupations and role models. The alpha male—and sometimes alpha female—is such an archetype, one partially grounded in our evolved human nature, but also constructed by societies.

Applied mathematics also involves human constructed scales and hierarchies. While mathematics is also somehow discovered, humans have used these insights to construct new scales and hierarchies. In his famous 1960 paper physicist Eugene Wigner pondered "the unreasonable effectiveness of mathematics in the natural sciences." Indeed, mathematics is essential to understanding all of the scales and hierarchies of the Great Matrix. Mathematics has a miraculous ability to describe reality in new and profound ways. Mathematics, however, is not simply discovered "out there" in the universe; mathematics is also evolved and invented. And once mathematics is discovered-invented, it can become completely self-referential. It need not play with any reality except itself in a Neoplatonic universe of ideas.

A lot of science and economics is arithmetic—measuring and counting. Complex modeling, however, requires the rocket scientists. And here, economic and scientific models begin to up run up against complexity and chaos theory. It is important to remember that the mathematics of finance and economics is not real. It need not refer to anything real in the way that the laws of physics do. The unit of measurement—currency—is a socially constructed fiction. No other species exchanges goods and services based on a symbolic system of value. Much like advanced mathematics, the measurements and models in economics and finance can be self-referential with no basis in reality. Still, there is an "unreasonable effectiveness" of mathematics in our economic evolution. Indeed, mathematics first took hold in human brains with the growth of trade in early agricultural civilizations.

The accelerating drama of human evolution required us to create all kinds of social and cultural hierarchies, in order to scale human cooperation from intimate hunter-gatherer tribes to millions living in mega-cities around the world.

10) Hierarchies of Emergent Complexity

Taken together—size, time, matter, energy, and all the rest—bring us to the final scale in the Great Matrix of all beings—the scale of emergent complexity. Science has no numeric scale for complexity. Here too we need to appeal to informed intuition and induction, rather than some discreet, measurable qualia in nature. The emergence of complexity, however, is certainly one of key features of Big History with important implications for economics and finance. Emergence is a bottom-up process of energy, matter, and information creating top-down constraints on lower-level processes. Humans in general, and scientists in particular, are extreme examples the top-down capacities to constrain, collect, and transform energy, matter, and information in our built physical and cultural environments.

Big History traces eight or more thresholds of emergent complexity in the evolutionary narrative from the big bang to the 21st century to create a pedagogically powerful narrative that enhances understanding, retention, excitement, and relevance—all badly needed in today's classrooms and society.

For instance, the creation of the heavy elements in the stellar foundries from which we derive the elements of the periodic table was a threshold of emergent complexity necessary for complex chemistry to later evolve. When complex chemistry catalyzed life, we saw again something new and different. And when the evolution of plants and animals gave rise to species with a central nervous system, complex brains, oppositional thumbs, vocal chords, symbolic language, tool-making, fire mastery, and collective learning, something new emerged again in the universe, at least on one small planet.

It is important to emphasize that emergent complexity requires lower levels of complexity to exist and function. Higher orders of complexity are built bottom-up, though emergent properties cannot be fully explained from below. With thresholds of emergent complexity, the Great Matrix is not simply a coordinate system of reality, but now also an epic narrative of becoming.

We can distinguish between four different kinds of emergent complexity in the new epic narrative of evolution:

1) Evolutionary Emergence: The grand arc of Big History is the story of the universe, Earth, and humanity. In the beginning, there were no stars, no galaxies, no periodic table of elements, no planets, no water, no complex chemistry, no life, and no intelligence. These emerged in a sequence—thresholds of emergent complexity—over a timescale measured in billions and millions of years.

2) Developmental Emergence: Each of us begins our journey as a single cell in our mother's womb. Over nine months, that cell replicates and differentiates into a few hundred tissue-types of our bodies. The development continues through childhood, into adult bodies composed of some 40 trillion cells. In this example, developmental emergence occurs in a timescale measured in months, years, and decades.

3) Functional Emergence: Protons, neutrons, and electrons form atoms; atoms join together to form molecules; molecules attach to form more complex chemistry. The inner life of a cell is a dense and intense collection of millions of molecular machines undergoing trillions of mind-boggling chemical cascades per second. The atoms and the molecules are not alive, but when functioning together, they can form a living cell, organ, and organism. Functional emergence in our bodies occurs in a timescale

measured in nanoseconds and hours.

4) Intelligence Emergence: With the advent of symbolic language and collective learning, we see the accelerating emergence of intelligence in humans. Science, technology, engineering, math, and medicine are the most obvious examples, as new insights and discoveries are built upon the successes of the past in a progressive and self-transcending learning process. This kind of emergence also occurs in the arts and humanities. The emergence of intelligence is a complex distributed phenomenon and is evolutionary (over the course of human history), developmental (over the course of an individual's lifetime), and functional (necessary for the workings of increasingly complex societies).[10]

While evolutionary emergence is the grand narrative of Big History, coming to recognize and understand developmental emergence, functional emergence, and intelligence emergence is no less important. The emergence of complexity is an on-going process.

To truly appreciate Big History, we must not only count in billions of light-years, but also the amazing nanoscale functions and developments manifested by quadrillions of atoms dancing in and out of our bodies. We must also appreciate the history of knowledge, creativity, and discovery that characterizes the recent, dramatic, and accelerating successes of our species.

With these different concepts of emergence in hand, we can now begin to understand how the academic disciplines and departments of the modern university map onto the structure of the universe itself. From the bottom up, we begin with fundamental theory, particle physics, nuclear physics, solid-state physics, and chemistry. The disciplines then split into inanimate and animate tracks. Inanimate matter goes from physical chemistry to material sciences, earth sciences, space sciences, astronomy, and cosmology. Cosmology, it should be noted, also takes us back to fundamental theory and particle physics, so the top and bottom are connected in the disciplines of physics and astronomy. Life, animate matter, leads to the many divisions of biology all the way to the human sciences, and indeed also all the humanities. Throughout our review of specialization and division of labor in universities, we also encounter applied sciences—for instance, engineering, medicine, and economics. The distinction between science and applied science is not easy to maintain.

A Multi-Dimensional Matrix

The ten dimensions of the Great Matrix give us ten ways to measure reality—by time, size, matter, energetics, electromagnetism, sound, information-ingenuity, sentience-consciousness, cultural constructs, and the end product, emergent complexity. Some of these are inductive scales that cannot be measured in numeric units. We might measure information, as bits for instance, but we cannot enumerate the ingenuity embedded in the code. We don't have a consciousness meter that we can put on your head. And beyond the crude measure of energy-density-flow, we don't have a scale for emergent complexity. It is a triumph in intellectual history that we have come to understand the scales of time, size, matter, and the varieties energetics. These are reliably real and profoundly practical insights about how our universe and our economy function.

All phenomena, all academic disciplines, and all economic activities can be located in reference to these hierarchical scales. It is this Great Matrix of bottom-up and top-down causality that allows complexity to grow.

Medieval Europeans once understood the universe to be a Great Chain of Being. All the entities of he world — animal, vegetable, mineral — were hierarchically organized. At the bottom were metals, precious metals, and precious stones. Then came plants and trees, followed by wild animals and domesticated animals. Humans were also hierarchically ordered from children to women

10 I am in debt to George Ellis for many of these insights. See George F.R. Ellis, *How Can Physics Underlie the Mind? Top-Down Causation in the Human Context* (New York: Springer, 2016).

to men and further into the different ranks of commoners, nobility, princes, and kings. The Great Chain of Being continued up into the celestial realm — moon, stars, angels, and archangels — to the very top where God presides over the entire creation. This *scala naturae* provided humans with a natural order, which they also understood to be a natural human order that structured their societies.

Science, or so the story goes, disrupted this view of the universe and ourselves. Copernicus, Galileo, and Kepler broke the crystalline spheres of Ptolemy and demoted Earth from the center of the universe to an insignificant periphery. Darwin understood plants and animals, including the human animal, to be evolving from common ancestors all the way back into the proverbial primordial slime. Freud showed that rational man was really an unconscious mess and hardly aware of, let alone in control of, his own thoughts and passions.

The Great Chain of Being was rendered a tangled web of happenstance in an enormous universe devoid of transcendence and meaning. God was rendered an unnecessary or incompetent creator. The new existentialists and Stoics argued that the universe was indifferent, that humans were insignificant, that our consciousness was epiphenomenal, and that our evolution merely accidental. Note simply that while there is no Great Chain of Being as the medieval Europeans understood, there is a great deal of order in the universe as discovered by contemporary science. There is a Great Matrix to which all beings belong that extends far beyond our direct senses and perception.

Humans are not at the top of the scales of size and time, but somewhere in the middle. On the energetics, information-ingenuity, sentience-consciousness, and emergent complexity scales, however, humans are off the chart. The human niche is particularly favored in the matrix for the time being—each of us a nexus of causal relationships (physical, biological, social, economic, psychological, mental), realizing extraordinary flows of energy and ingenuity, intensities of subjective experience, and accelerating transformations in the modern period.

In our drive toward specialization and division of labors, we rarely reflect on these natural hierarchies and what they might mean for our understanding of science, self, and society. "The ongoing fragmentation of knowledge and resulting chaos are not reflections of the real world but artifacts of scholarship," writes Harvard biologists E.O. Wilson in his book, *Consilience: The Unity of Knowledge*.[11] Vartan Gregorian, president of the Carnegie Corporation, similarly observes:

> The fundamental problem underlying the disjointed curriculum is the fragmentation of knowledge itself. Higher education has atomized knowledge by dividing it into disciplines, subdisciplines, and sub-subdisciplines—breaking it up into smaller and smaller unconnected fragments of academic specialization, even as the world looks to colleges for help in integrating and synthesizing the exponential increases in information brought about by technological advances. The trend has serious ramifications. Understanding the nature of knowledge, its unity, its varieties, its limitations, and its uses and abuses is necessary for the success of democracy... We must reform higher education to reconstruct the unity and value of knowledge.[12]

Understanding how the Great Matrix actually works on different scales is an exercise in zooming in, zooming out, and changing perspectives. In the process, we better understand the emergent complexity from physics to chemistry, from cell biology to human brains, from individual producer-consumers to global markets. We need to take account of the energy, matter, and ingenuity that flows through nature and our economy.

We are never outside the bio-social-physical matrix, but in this scientific and philosophical exercise we seem to stand away, looking down on the matrix from above. As far as we know, no other entity in the universe has achieved this capacity, and it is in this domain that humans are

11 Edward O. Wilson, *Consilience: The Unity of Knowledge* (New York: Knopf, 1998).

12 Vartan Gregorian, "Colleges Must Reconstruct the Unity of Knowledge," *Chronicle of Higher Education* 50, no. 39 (2004).

no longer middling creatures of the matrix. Our self-transcendence, realized especially through the progress in science, economics, art, and culture, is a super and completely natural emergent phenomenon. We come to understand the matrix from the inside out, though the matrix knows nothing of us.

It is awe-inspiring to grok any or all of these natural scales and hierarchies. Simply appreciating the scales of size and time is awesome. Try also to analyze the flows of energy embodied in the objects and activities in your immediate environment. A process analysis of energy in our built environment gives us a new understanding of the economic world. Training your eyes to "see" the energy flows embedded in the world around you is sure to wake you up in the morning. I look at my library and recall that a ton of paper embodies on average about 35 gigajoules (GJ). This is about as much energy as is needed to manufacture a ton of good-quality steel, to say nothing of the human labor that went into writing these books. I walk through my home and recollect that the average three-bedroom, wood-framed house in North America embodies about 500 GJ of energy in its construction, which is equivalent to 4,157 gallons of gasoline (or something like 48,183 days of human labor). I get into my car and consider that a midsized passenger car requires about 110 GJ to build and might consume about 50 GJ of fuel annually. Over ten years, the energy cost of that car would total about 680 GJ—more than that of the average three-bedroom house mentioned above.

It doesn't stop there. If your eyes could detect the radio waves around you, the empty space of your room would be filled with an altogether different kind of rainbow—the chaotic patterns of dozens of overlapping VHF broadcasts would be evident all around you. If your skin could feel these electromagnetic waves, your entire body would be vibrating and tingling at different frequencies. Indeed, some of the airwaves entering through the window right now have traveled 13.8 billion years from the background radiation of the big bang to be with you today. Or if you prefer the wave function of quantum phenomena, consider yourself connected across space and time in a very immediate sense to the distant reaches of the universe. Understanding the Great Matrix provides a kind of instant mystical experience in which you can deconstruct and reconstruct your sense of self and world within these multidimensional scales and hierarchies.

Science offers a sixth sense—a way of seeing beyond the walls of Plato's cave into the realm of what is real. It isn't rocket science to sort-of-know how things actually work, how the pieces form an entire puzzle. Of course, god and the devil are both in the details, but it sure helps to have the picture on the puzzle box when trying to fit the pieces together.

Bibliography

Berry, Thomas, and Brian Swimme. *Universe Story, The: From the Primordial Flaring Forth to the Ecozoic Era*. San Francisco, CA: Harper, 1992.

Chaisson, Eric. *Cosmic Evolution: The Rise of Complexity in Nature*. Cambridge, MA: Harvard University Press, 2001.

———. "Energy Rate Density as a Complexity Metric and Evolutionary Driver." *Complexity* 16, no. 3 (2011): 27-40.

———. "Energy Rate Density. Ii. Probing Further a New Complexity Metric." *Complexity* 17, no. 1 (2011): 44-63.

———. *Epic of Evolution: Seven Ages of the Cosmos*. New York: Columbia University Press, 2006.

Christian, David. *Maps of Time: An Introduction to Big History*. Berkeley: University of California Press, 2004.

Dennett, Daniel C. *From Bacteria to Bach and Back*. New York: W.W. Norton, 2017.

Dunbar, Robin. *Grooming, Gossip and the Evolution of Language*. Cambridge: Harvard University Press, 1996.

Ellis, George F.R. *How Can Physics Underlie the Mind? Top-Down Causation in the Human Context*. New York: Springer, 2016.

Extance, Andy. "How DNA Could Store All the World's Data." Nature, http://www.nature.com/news/

how-dna-could-store-all-the-world-s-data-1.20496.

Grassie, William J. *Reinventing Nature: Science Narratives as Myths for an Endangered Planet*. Philadelphia: doctoral dissertation, Temple University Department of Religion, 1994.

Gregorian, Vartan. "Colleges Must Reconstruct the Unity of Knowledge." *Chronicle of Higher Education* 50, no. 39 (6/4/2004 2004): B12.

Hidalgo, Cesar. *Why Information Grows: The Evolution of Order, from Atoms to Economies*. New York: Basic Books, 2015.

Munger, Charles. "Usc Law School Commencement Address, May 1, 2007." Genius.com, http://genius.com/Charlie-munger-usc-law-commencement-speech-annotated.

Prigogine, I., and Isabelle Stengers. *Order out of Chaos : Man's New Dialogue with Nature*. New York, N.Y.: Bantam Books, 1984.

Prigogine, Ilya, and Gregoire Nicolas, eds. *Self-Organization in Nonequilibrium Systems: From Dissipative Structures to Order through Fluctuations*. New York: Wiley, 1977.

Primack, Joel R., and Nancy Ellen Abrams. *The View from the Center of the Universe: Discovering Our Extraordinary Place in the Cosmos*. New York: Riverhead, 2006.

Smil, Vaclav. *Energy in Nature and Society*. Cambridge, MA: MIT Press, 2007.

Spier, Fred. *Big History and the Future of Humanity*. New York: Wiley-Blackwell, 2011.

———. *The Structure of Big History: From the Big Bang until Today*. Amsterdam: Amsterdam University Press, 1996.

Weschler, Lawrence. *Waves Passing in the Night: Walter Murch in the Land of the Astrophysicists* New York: Bloomsbury, 2017.

Wilson, Edward O. *Consilience: The Unity of Knowledge*. New York: Knopf, 1998.

Situating human generations in Big History

Peter D Nixon

Abstract

Big History provides the cosmology for the present work in which the disproportionate impact of inequality on today's young generation is contextualised by examination of three types of societies, past empires, present capitalism and one pre-colonisation indigenous society, over time periods of more than centuries. Generations are seen as demographically referenced social locations like gender and race and at any one time several interdependent generations exist. Other referents include the work of JS Mill, his controversial ethology project, the history of ideas, the non-identity problem and social anthropology. The article concludes by anticipating a tentative reconstruction of Mill's ethology based on advances in social science in the century and a half since his time.

Correspondence | Peter Nixon, peter.nixon@isic.ekit.com
Citation | Nixon, P. (2018) Situating human generations in Big History. *Journal of Big History*, II(2); 53 - 77.
DOI | http://dx.doi.org/10.22339/jbh.v2i2.2302

Introduction

Big History looks at history from the Big Bang, which was, as far as we know, the beginning of our universe, to the present, placing humanity in this context without privileging humanity. Big History divides the past into manageable chunks which sometimes fit established fields of knowledge and sometimes do not. Books have been written about Big History (Christian, 2011; Spier, 2015) and other fields of scholarship referred to in what follows. It is claimed here that the links between these fields of knowledge establish the need for taking a very long term view of the past and future and the place of humanity in this view.

The key concepts in Big History which make it a suitable framing reference for the present work are those of complexity, energy, and the evolutionary process that has created our universe and us. After briefly describing Big History and the problem to be addressed here, an outline of the rest of the article is presented.

The principles of thermodynamics, literally, "the dynamics of heat", and the rate at which the universe has been expanding are said to be responsible for the complexity we see around us (Spier, 2015, 53-56). The first law of thermodynamics states that energy is always conserved, though it may change in form, e.g. from electricity to heat, or from heat to motion. The second law distinguishes energy which is available to do work, i.e. free energy, from that which can no longer be harnessed and states that in a closed system the amount of free energy decreases over time: it is the free energy which creates and sustains complex entities (Christian, 2011, 506-507). The rate at which the universe is expanding has led to a description of what has been called the Goldilocks Principle (Spier, 2015, 63-65). This principle states that for the bringing into being of complex entities the circumstances must be just right: the relevant circumstances relate to gravity, temperature, pressure, radiation and other forms of energy. The circumstances that brought the stars into being are quite different from those which evolved later around the stars and allowed complex life to develop. As Hawking (1996, 164-173) explains, ordered, organized or complex states of inorganic or animate being are less likely to arise than disordered states. The development of complexity is an evolutionary process. Big History offers a view of the entire universe that is based on careful observation and contestability in various fields of science rather than the narratives about our origins that have been handed down through generations in many cultures around the world but which have been constrained by the circumstances and technologies available to those cultures. Big History gives us a new cosmology.

In lived human experience cosmology may reference various time scales and social practices. The worlds of work, family, school, domesticity, intergenerational relationships, governance, etc. can be described and analyzed as discrete, though in lived experience they intersect. The way these worlds can be understood in Big History is necessarily different from the ways they may be understood from traditional religious perspectives. This work assumes that the manner in which these worlds intersect in the lives of individuals may prove a fruitful focus for Big History consideration in the future.

Human history is a very compressed, short part of the big picture. But:

> *On the modern map of complexity, humans are as central as they were within most traditional cosmologies.* (Christian, 2003, 437-458).

Big History research projects span many disciplines. At Macquarie University Gillings, Paulson & Tetu (2015) and Gillings (2017) concern themselves with internal microbial human health and the impact of antibiotics in our environment respectively. The first Big History doctorate in the world was completed at Macquarie University by Dr David Baker (2015). His interests are complexity, Universal Darwinism and cultural evolution. The International Big History Association web page carries a passage more directly relevant to the present work:

> *Does Big History provide a narrative that can help nurture the development of the empathy and cooperation that are part of our social nature? Can humans form a more perfect human community as we continue to create a more complex society than has existed before? Or will our current levels of social complexity face inexorable demise?*

Before considering a modern social paradox relevant to this goal it is important to appreciate the relation between time and human Big History. Christian (2011, 3) links Hawking's ground-breaking view of time with Big History. For the purposes of the present work it is only necessary to acknowledge Hawking's account of the pre-human history of the universe and his multi-dimensional theories that may help us predict its post humanity future. The timescale that is relevant to the present problem is that of human generational transition. Focusing on this may produce outcomes relevant at other timescales, e.g. that of global warming, and be the subject of surrounding forces which operate on yet different timescales such as structures of social power. The generational transition timescale contains representations of the past and hopes for the future: it is part of the evolutionary process which started at the Big Bang and is embedded within the Big History of the earth and that of humanity.

Spier has recently described the attributes of academic research which are common across the fields of the natural sciences, humanities and social sciences. His background as a cultural anthropologist infuses his writing which speaks strongly to the present work:

> *… doing social science research,*

> *most notably longer term participatory observations within a cultural setting that is different from our own, places far greater demands on a wide variety of skills than any other form of science.* (Spier, 2017).

The present work provides a theoretical background for social science research concerning Big History and human generations. This sets the work apart from much of the public commentary on generations and backgrounds research to be reported in a sequel article. The present work and its sequel will take account of the four basic requirements for academic research identified by Spier: logical consistency, explanation or structuring of empirical observations, openness to falsification and the sharing of findings. The sequel will also address his additional requirement in relation to social science arising out of the participatory action research tradition: it will locate the researcher in generational and institutional terms, in relation to the research process, participants and outcomes. Without specifying them in the same terms, these points were addressed by the author in an earlier publication (Nixon, 2009).

The paradox: Rising inequality in the presence of publicly declared 'poverty lines'

A current issue which focuses our attention on intergenerational relationships is the paradoxical co-existence over decades in developed countries of publicly set 'poverty lines' on the one hand and increasing pecuniary inequality on the other. Other issues could have been chosen, but this will suit the purpose of increasing our understanding of the relevance of human generational transition to Big History.

There are voluminous and separate literatures on poverty, poverty lines and inequality: some focus on subsistence, others on the debate about relative versus absolute poverty. A full appraisal of them is beyond the scope of the present work. The three key points to be demonstrated in what follows are that:

1. poverty lines have been publicly set by government in Australia and other developed countries for decades,
2. inequality has been steadily rising,
3. the disproportionate impacts of inequality on the young.

Firstly, poverty lines were first drawn in Australia based on a survey conducted by Prof. Ronald Henderson in Melbourne in 1966 (Kewley, 1973, 391-393; Henderson, Harcourt & Harper, 1970). This led to him being appointed by the Federal government to lead the National Commission of Inquiry into Poverty which updated these lines (Henderson, 1975, 14), guided Federal budget provisions in 1972-73 and featured in parliamentary debates at the time (Kewley, 1973, 400-401). Today those poverty lines are updated quarterly by the Melbourne Institute of Applied Economic and Social Research at Melbourne University.

Similar accounts can be given concerning other developed countries. For example, in United Kingdom it began with the Beveridge Report in the 1940's (Gordon & Townsend, 1990). In Canada, Germany, Great Britain and the United States, comparisons were not possible until the 1980's when comparable data first became available, implying that data had been gathered in those countries for some time prior to the 1980's: in the United States poverty thresholds had been set in the 1960's (Valletta, 2006). Other national governments that began around this time to systematically collect and publish population level data concerning the incidence of poverty were Canada in 1961, Germany in 1963, Finland in 1966 and Sweden in 1967 (Atkinson, 1991). Governments in developed countries have been setting poverty lines for decades.

Secondly, reports prepared by the Organization for Economic Cooperation and Development (OECD) over recent years have shown increasing inequity within and between countries. The titles of the OECD reports convey concern about poverty in 2001 but increasingly *inequality* is the recurring theme (OECD, 2001; OECD, 2008; OECD, 2011; OECD, 2015). The most recent one begins by proclaiming:

> *The gap between rich and poor keeps widening.*

In keeping with the OECD report of 2001, Kanbur's (2001) focus is poverty in an international context and his work benchmarks the argument concerning inequality. He acknowledges wide areas of agreement between abstract entities he characterizes as the 'Finance Ministry' and 'Civil Society'. His Finance Ministry contains economists, government finances, banks, etc. His Civil Society contains human rights advocates, NGO's, the United Nations, etc. They are said to largely agree on the value of education and health, the role of international public goods, the relationship between market and state and the central importance of regulatory institutions over markets, governments and the household-market relationship. He then describes the following areas of disagreement. Aggregate statistics require careful explanation if they are to be accurately interpreted, e.g. the percentage of a population in poverty may fall, but if the population is increasing, the absolute number in poverty may be increasing; different poverty lines are drawn from mean compared to median income levels; and the value of public services is not always counted and varies according to whether they are working well or not. The duration being considered makes a difference to perceptions of poverty e.g. the:

> … *daily reality of poor people's lives* … (ibid.)

can be presented in the media with great immediacy; in the medium-term policy makers often refer to safety nets; and suggest that:

> … *technological change will come to the rescue as it always has in the past.* (ibid.)

The final area of disagreement concerns market structure and power. Kanbur asserts that the 'Finance Ministry' assumes that markets are competitive while in 'Civil Society' it is considered that markets are not competitive. He further asserts that when markets are distorted by either poor infrastructure or local monopoly, capital mobility will hurt labor in countries both sourcing and receiving the capital. Referring to international financial institutions Kanbur calls for a widening of debate on economic policy and agreement on definitions:

> … *'growth' is used to mean both an increase in per capita income, and to refer to a policy package.* (ibid.)

Referring to a meeting of the Governors of the World Bank he concludes:

> *When the institution whose self-stated mission it is to eradicate poverty can only hold its Annual Meetings under siege from those who believe its mission is to further the cause of the rich and powerful, there is clearly a gap to be bridged.* (ibid.)

Based on the more recent OECD reports cited above it is clear Kanbur's call has gone unheeded: inequality has been rising for at least a decade and a half, if not more.

Finally, two recent Australian reports on inequality highlight its impacts on the young. The Grattan Institute report (Daley, Wood, Weidman & Harrison, 2014) examined the impact of inequity across generations and found that:

> *Growing wealth has not benefitted the young.*

And

> *Spending policies increasingly benefit older Australians.*

These authors conclude that the financial policies of governments need to change to enable a fairer share of burdens and benefits between generations, but they note the more substantial voting power of the older generation based simply on its size, relative to the young, may be an impediment to this. Increasing longevity and the burden experienced by the young have combined to create the 'Sandwich generation' (Grundy & Henretta, 2006), those in middle age who find themselves caring for both younger and older generations simultaneously. While poverty remains a concern for the Australian Council for Social Service (ACOSS) their recent report of over 50 pages on inequality found that income inequality in Australia is higher than the OECD average and that wealth is even more unequally distributed (ACOSS, 2014; ACOSS, 2015).

In summary, despite the existence for decades of government declared poverty lines, inequality has been growing and is creating disproportionate burdens for the young. The question thus arises about whether governments have the capacity to play a role in addressing this inequality? The paradox arises because democratic governments put a lower limit on inequity through poverty lines but do not set limits on how rich it is possible to become.

It is argued below that by placing these observations about generations and inequality in the context of human Big History a new perspective on this paradox emerges.

Responding to the paradox

The rest of this article describes a theoretical approach which can be used to structure research processes which address intergenerational issues in the context of Big History.

The next section places the paradox in the human period of the earth's Big History by reference to two completely independent fields of scholarship, (Turchin, 2003; Turchin, 2006; Turchin & Nefedov, 2009; Figueroa, 2015) both of which consider various time periods in human civilisations up to and including centuries. These scholars show that there are circumstances in which some complex societies can be expected to become dangerously less egalitarian. The third section relates the classical writers on generations to Mill's notions and to unborn future generations. The fourth section considers further independent fields of scholarship and contains a sympathetic description of a non-literate civilisation, the Australian Indigenous peoples whose civilisation has lasted continuously for tens of thousands of years (Mulvaney & Kamminga, 1999). The ways Australian Indigenous peoples would have regarded the non-identity problem and handled knowledge are critical points in this section (Gammage, 2012; Kelly, 2016). The fifth section considers Steinbruner's (2011) sympathetic interpretation of Mill's controversial writings, and the ways in which the word "ethology" has been interpreted since Mill used it to describe a scholarly pursuit he considered to be highly important – the study of the formation of character.

Some scholars have suggested remedial actions aimed at problems very similar to the paradox addressed here (Figueroa, 2015, II, 189-201; James 2006, 292-314). Proposed remedies are outside the scope of the present work but worthy of consideration elsewhere. The following section discusses how the paradox identified arises at the confluence of several social patterns which occur over different time periods, from years to centuries. The central concern here is to create a theoretical framework that will facilitate timely feedback as different remedies are tried.

Placing the problem in the human period of the earth's Big History.

The paradox concerning the coexistence of government set poverty lines and increasing inequality with disproportionate impacts on the young, can be better understood by reference to various timescales in the history of human civilisations. The scholarly fields of evolutionary biology and economics work with different kinds of evidence and methodologies but one author in each field has sought to explain changes in different types of civilisations. Each author builds concepts important to the other field into their analysis, so that, for example, the evolutionary biologist considers productivity of humans and land and the economist emphasises that 'history matters'. Both see humanity in the context of the natural world, hence an appropriate subject for rigorous scientific enquiry. The evolutionary biologist, Turchin (2003), presents theories which explain much of the rise and decline of agrarian empires, while the economist, Figueroa (2015), presents a unified theory of contemporary global capitalism. Both authors test their theories against appropriate available data. Their theories consist of mathematical models and while Turchin begins and ends with the timescale of the life of an empire and treats shorter timescales such as generational ones in the decline phase, Figueroa begins with a very short time scale, and concludes with timescales of decades and centuries, including reference to generational transitions in the longer timescales. In the present context the most important thing about their work is that they reach very similar conclusions about the decline of agrarian empires (Turchin) and the future of capitalism (Figueroa) and these conclusions have a direct bearing on the paradox identified above.

The rise and fall of agrarian empires

Mathematical models describing the rise and decline of agrarian empires are at the centre of Turchin's (2003) work. He defines empire as:

> *… a large multiethnic territorial state with complex power structure.* (Turchin, 2006, 338)

However, the political structures in the states which Turchin defines as empires are not only internally complex; they may be very different:

> *The political arrangements by which a state is governed are irrelevant to the definition of empire* (2006, 339).

In summary, Turchin tested his theories on complex societies of the past with diverse forms of government.

Turchin views empires as beginning in an ethnic group, or ethnie, which, unchecked, expands by natural increase until it encounters the territory of another ethnie which it then dominates: his meta-ethnic frontier theory models this expansionary phase. The ethnic core may be a coercive elite, an elite which cooperates with the population or a:

> *Lateral empire wide integration of elites … who rule over ethnically heterogeneous commoners.* (Turchin, 2003, 50).

Turchin's second model, Ethnokinetics, aims to explain the incorporation of weak ethnies in stronger ones, a process which defines an empire but is of less interest in the present context than his other models. His Demographic-Structural theory aims to explain the decline of empires. Finally, secular cycles in population size were predicted by Turchin on the basis of the Demographic-Structural modelling:

> *… population numbers in agrarian societies are expected to go through*

> *slow oscillations with a periodicity of approximately two or three centuries.* (2003, 150).

or many human generations, and his analysis of empirical observations supported this.

In Turchin's 2003 publication he reports testing against two agrarian empires and in the later one, co-authored with Nefedov, mature models of empire formation and decline are tested against several more (Turchin & Nefedov, 2009). These authors gathered comparable demographic and archaeological data for all the agrarian empires they studied. They examined eight different empires, spread across England, France, Italy and Russia and spanning almost two millennia. In relation to a particular period in mediaeval England they state:

> *Nominal wages did not exhibit a cycle but grew fairly monotonically. Thus, a building craftsman's wage increased from 3 pennies (d.) per day in the late thirteenth century to 6 d. per day in the early sixteenth century (Phelps-Brown & Hopkins 1955). Real wages, by contrast, exhibited an oscillation, driven by the cycling movement of prices … .* (Turchin & Nefedov, 2009, 38).

This level of precision allows the authors to trace and analyze patterns discretely within different economic strata of their subject populations. The statistical methods employed by these authors allowed them to map the patterns observed among different strata of the population against aggregate indices of stability and growth or decline (Turchin & Nefedov, 2009, 310-311).

The secular cycles in agrarian societies described by Turchin & Nefedov (2009) involve alternating integrative and disintegrative phases. The integrative phase typically involves:

> *…a centralizing tendency, unified elites, and a strong state that maintains order and internal stability. Internal cohesion often results in the vigorous prosecution of external wars of conquest, which may result in the extension of the state territory*. (Turchin & Nefedov, 2009, 20)

The disintegrative phase typically involves:

> *a decentralizing tendency, divided elites, a weak state, and internal instability and political disorder that periodically flare up in civil war.* (ibid. 20)

In the disintegrative phase the population becomes more stratified and cooperation within and between social strata diminishes. The authors present different indices of social instability depending on what data was available. For the period between 1100 and 1500 in mediaeval England the evidence concerns rebellions and the finding of coin hoards (ibid. 31, 47) and for the Romanov cycle in Russia from 1795 till 1928 the evidence concerns peasant disturbances, crime incidence and named executions (ibid. 284-286).

In two cases out of the eight empires studied, Turchin & Nefedov found evidence of recurring generational cycles which had a very approximate periodicity of two generations or 40 to 60 years (ibid. 28, 79-80, 106, 203, 285): the authors called them *social-psychological* cycles. In the remaining six empires, no such evidence is presented by Turchin & Nefedov, so whether generational patterns existed in them is an open question.

The social-psychological cycles described by Turchin (2006, 243-244, 285-286, 295-298) became evident to him when studying the disintegrative phase of much longer secular cycles. Turchin argues that in these cycles men who return from war and know its cruel realities have the effect of 'immunizing' their sons against war, and so what follows is a period of peace. These sons may have a slightly similar but modified reaction, but eventually a generation of men matures without experiencing war. At this point, if the demographic and other pressures that caused the conflict in the first instance are still in operation, the likelihood is that another war will break out, and so on until the pressures change. The term *generational* is used here to refer to these cycles.

Turchin's (2008) theory asserts that in agrarian empires social cohesion, synchronous equity and the state all decline together and the empirical work he refers to shows that in some cases generational cycles were associated with the decline of the empire.

Integrated economic theory of capitalism

In global capitalism nation/states cooperate to maintain global markets in which businesses compete. Using Popperian logic, Figueroa creates a theory which applies across all capitalist countries. In his view this is necessary because he sees conventional economic theories as being limited in their scope. He reviews Neoclassical, Keynesian and Classical economics and finds that none of them can account for the existence and persistence of unemployment in the first world: in addition he finds that Classical and Neoclassical economics cannot account for the short run correlation between nominal and real economic variables in the first and third worlds (Figueroa, 2015, 1, 27-39). Figueroa describes nominal variables as those under the control of government such as money supply, exchange rate and interest rate: he defines real variables as things like total output, real wages and employment (ibid, 34). The impact of changes in nominal variables are abstracted from individuals' experience through the machinery of government, where-as real variables have more tangible impacts.

Communist or second world countries are excluded from Figueroa's work by definition and he divides third world countries into two groups having either weak or strong colonial legacies. He reviewed relevant economic data, from historical and recent international datasets. Eastern Europe was excluded because the introduction of market reforms since the 1990's meant that they were 'in transition' to capitalism (ibid. 28-30).

Using data from the individual countries he creates abstract models of capitalism as it operates in:

a) First world countries (Epsilon society)

b) Third World countries with weak colonial legacies (Omega society)

c) Third World countries with strong colonial legacies (Sigma society).

A key aspect of these models is the slightly different representations of excess labor in each model. In Epsilon it is represented as unemployment, in Omega as unemployment and under-employment with income gaps between ethnies and in Sigma it is like Omega with increased self-employment.

He defines three different timescales relevant to the development of his theory:

a) The short run which typically lasts less than half a year in which he posits a static economic model,

b) The long run which typically lasts a year or so in which he posits a dynamic economic model,

c) The very long run which can last over decades and may run to centuries in which he posits an evolutionary economic model.

Fewer economic variables need to be considered in the short-term but the longer the timescale being considered, the more relevant variables there are that might change. From these models he creates a unified theory of capitalism which explains the representations of excess labor and the degradation of the biophysical environment that occurs in the very long run. His theory culminates in the presentation of a fourth model, to be described below.

Referring to the two shorter timescales Figueroa (2015, II, 93) concludes:

> *According to the theory of capitalism, the global inequality in the capitalist system (which comprises between-country inequality and within-country inequality) will persist as long as the distribution of economic and political assets at the national and international levels remains unchanged. … According to the unified theory, therefore, there is path dependence in the process of capitalist economic growth, that is, history matters.*

It is in the very long run context that Figueroa introduces his concept of the intergenerational consumption frontier and the laws of thermodynamics:

> *In conclusion, … when the stock of nonrenewable natural resources is included in the economic process, society is faced with an intergenerational consumption frontier. … The average consumption level of future generations will necessarily be smaller than that of the current generation.* (ibid. 104)

> *We can then say that the production process only rearranges matter and energy, … The economic process is not mechanical, but* entropic. (ibid. 108) [Emphasis original]

In his penultimate chapter Figueroa builds the laws of thermodynamics into his model, moving beyond the earlier static, dynamic and evolutionary theoretical models:

> *The final entropic model of the unified theory predicts that the economic growth process supplies society with increasing quantities of goods per person, which has a positive effect on quality of society, but it is subject to decreasing marginal gains. … The model thus predicts a trajectory of an inverted-U shaped curve upon the quality of society in the economic growth process. … The available facts tend to corroborate the predictions of the entropic model. … This seems to indicate that we are entering into the downward sloping segment of the curve or maybe we are already there.* (ibid. 172)

Similarities between the conclusions of Turchin and Figueroa will now be explored.

Summary

With regard to both the quality of society and intergenerational transitions, there is agreement between Turchin and Figueroa based on the evidence they considered and the theories they generated. Their work took place in totally different fields of scholarship but both considered centuries of human existence and both portray human Big History. Though Turchin did not find generational patterns in every case, the ones he found were sufficiently well pronounced to enable him to build theories around them. However Turchin's empires and Figueroa's capitalist states might be compared, it would seem that Turchin and Figueroa have independently provided demographic and economic explanations, respectively, for an underlying phenomenon. Are we witnessing a process of entropy or could this be a kind of group level evolutionary process; must we postulate that it is related to human agency or could there be some other explanation? This question is answered in part below. A civilization in decline is shown by these authors to have deleterious impacts on the young; it is the young who pay for the excesses of the old.

The length of the cycles and trends described here creates great difficulty for assessment of potential impacts of emerging cultural trends. The impact of changes made at one point in time may not become evident for decades. Timely assessment requires a synchronous method that can be repeated at intervals to build an historical record. The method needs to be informed by classical generational theorists and more recent generational scholarship.

Generations observed

The concept of *generation* is foregrounded by the paradox being addressed here. In 1987 the concept of generations entered international discourse on governance through the Brundtland Report, 'Our Common Future'. The brief of the World Commission on Environment and Development, of which Brundtland was the Chair, was concerned with sustainable development which was seen as a path to eradicating poverty (Brundtland, 1987, 24-25). Of the 22 principles emerging from the Commission's work only the first two related to social aspects, the first one related to human rights and the second related to generations (ibid. 339). The second principle expresses the desire of the Commission to establish norms about what different generations might expect of the state in relation to the natural environment:

> *States shall conserve and use the environment and natural resources for the benefit of present and future generations.* (ibid. 339).

But no way of balancing the needs of present and future generations is offered. Hence, in that significant international report concerning sustainability, both of the principles addressing social aspects related to social justice. The first addressed synchronous justice and the second hinted at diachronous justice. However, the underlying concept of sustainability is contested.

In a sociological account of the social sustainability and social resilience of an identified population in rural Argentina over the 19th and 20th centuries Adamo (2003) begins with a critique of the Brundtland report in which she addresses the justice element of the sustainable development concept:

> *Sustainable development is becoming a normative concept difficult to articulate and put into practice.*

She asks 'what is to be sustained' and by whom, etc. She treats social, environmental and economic sustainability as separate but interdependent. Adamo mentions the importance of space and time and an historical perspective as ways of contextualising sustainability, and this is also reflected in the work of Scott, Park & Cocklin (2000). Adamo designed her study to answer questions about what makes people remain in rural areas rather than what makes them leave. Adamo draws a connection between social sustainability and social resilience. Of particular relevance here is Adamo's definition of the 'community/culture attribute' of the connection and her reflection on its value for her research. In this attribute she incorporated 'intergenerational equity and cultural capital ... social institutions, formal and informal ... services and infrastructure.'

Reflecting on her own application of this definition, Adamo comments:

> *This (Community/Culture/ Institutions/Services) attribute is more difficult to address, given the imprecision of the elements involved.*

As much as she wants to escape from the dominant normative concept of sustainable development, Adamo concedes having difficulty when trying to articulate elements of social sustainability which stand beyond polity. In light of the discussion above concerning empires and capitalism, it might be said that the state (or polity) has a generational identity which may vary depending on whether the state is rising or falling. Adamo's difficulty is understandable when considered in light of the life cycles of empires and capitalism. Brundtland's principle concerning generations is inadequate because it ignores these time scales.

Theoretical developments in the social sustainability field are tentatively moving beyond the notion of sustainable development enshrined in the 1987 Brundtland Report. The need to do this has been recognised in a recent Australian study of public sector governance (Osborne, 2010). Based on the work of Figueroa it would seem that development has been emphasised at the expense of sustainability, however that term is interpreted.

Human Resource Management consultancies (McCrindle, 2009; McQueen, 2008; Lancaster & Stillman, 2005) often consider what are perceived as the competing interests of different generations (Rayner, 2016; Goertzel, 2006; Heath, 2006; Walsh & Black, 2011; Seedsman, 2006). Sociology can help us understand generations through less biased lenses.

Mannheim, known for his contribution to the sociology of knowledge, provides a summary of much of the scholarly thought about generations, as it stood in the middle of the twentieth century (1952, 276-322). Mannheim distinguishes two definitions of generations: one based on genealogy (kin) and the other on cohorts. The cohort definition of generations is the one used in this work and it groups everyone into generations according to the years in which they were born; people born in defined sets of consecutive years are said to be separate generations. Members of a generation have, at the same stages of their lives, shared the experience of significant public events. Generation is seen as a social location in the same way as gender and race (ibid, 291). With regard to the periodicity of generational cycles:

> *It is a complete misconception to suppose, as do most investigators, that a real problem of generations exists only in so far as a rhythm of generations, recurring at unchanging intervals, can be established.* (ibid, 286)

The length of a generational cycle is not fixed.

Different groups within the same generation may respond to their social and economic circumstances in different ways. If a minority within a generation tends to respond to their circumstances in a way that differs from the response of the majority they are called a generational unit (ibid, 304). The existence of generational units adds a layer of complexity to the analysis of intergenerational relationships:

> *... there exists a uniform generational context in the sense of a shared problem community, but not a generational unity whose members could offer uniform solutions to these problems.* (Jaeger, 1985).

In making this distinction, Jaeger raises questions about whether and how points of contention between and within generations are dealt with, and by whom; sharing a time in history does not necessarily lead to cooperation.

The notion of generational identity or personality goes back to Dilthey's notion (advanced in the late 1800s) of 'the intellectual culture of a generation' and referred to by Marias (1970, 52). Generations have been seen as vehicles for transmission of culture and viewed as recurring in cycles of either two or four types (ibid, 183-184; Turchin 2003). The common feature of both the two and four type models is that influential generations alternated with generations that appear to have had less influence on the course of history. The generation cohorts were variously estimated to be between fifteen (Marias, 1970, 185) and thirty (Mannheim, 1952, 310) years long, which allowed researchers much scope for blurring the boundary years. Marias (1970, 184-185) described a pattern of four different and succeeding generations living at the same time in history and described the tensions between them thus:

> *The year 1800 is not a single date; it is four different dates that exist simultaneously and are mutually involved in an active form. Strictly speaking what we find is not movement so much as it is that which shapes movement and makes it possible – a system of tensions and working forces ... This tension which is revealed in the multiplicity of generations, is the force behind historical movement.*

This point is a guiding principle for the discussion below.

John Stuart Mill on generations

Mill was a prominent politician in Britain in the middle of the 1800s whose thousands of pages of writing including letters to individuals and to the press were published, some posthumously, between 1823 and 1898. These were assembled into 33 volumes and have been made available digitally (Mill, 2016). Mill's writings have had a significant impact on much of the political life of the 20th century and they are relevant here for two reasons, his interest in generations and something he called ethology that will be discussed below. He was born in 1806 and died in 1873. For three years from 1865 to 1868 he followed in the footsteps of his father, James, becoming an elected member of the British Parliament. In phase two of the present work I searched digital copies of his publications and, in all, found 512 references to generations. It is evident that the frequency of those references increased significantly as his maturity as a writer progressed, peaking in publications between 1868 and 1872. Some of his references to generations are not made in the same sense in which generations are referred to above; he could have used other words to describe his meaning. Setting aside those references it is clear that Mill understood intergenerational phenomena in much the same way as the later scholars discussed above understood them.

Mill's commitment to intergenerational equity was very clear from early in his parliamentary term:

> *What are we, Sir - we of this generation, or of any other generation, that we should usurp, and expend upon our particular and exclusive uses, what was meant for mankind?* (28, 143)

and also expressed in his essays:

> *The owner of capital is by no means to consider himself it's absolute proprietor.* (10, 485)

Mill was ahead of his time in many respects and lamented the parlous state of what we would now call the social sciences and in particular, knowledge about psychology (8, 223-224). The neurologist and philosopher, Antonio Damasio

(2003, 15), acknowledges Mill's grasp of the:

> *… relation between personal and collective happiness, on the one hand, and human salvation and the structure of the state, on the other, …*

It might even be said that Mill foresaw the development of human Big History:

> *The facts of each generation are looked upon as one complex phenomenon, caused by those of the generation preceding, and causing, in its turn, those of the next in order. That these states must follow one another according to some law, is considered certain: how to read that law, is deemed the fundamental problem of the science of history.* (20, 345).

Approaching his maturity as a writer, Mill became aware of a set of circumstances in Britain at that time which are similar to aspects of the paradox addressed in the present work:

> *In fact, it has now come to this, that instead of being at liberty to suppose that future generations will be more capable than we are ourselves of paying off the national debt, it is probable that the present generation and the one or two which will follow, are the only ones which will have the smallest chance … .* (28, 141)

of being able to repay it. He later qualified this by citing circumstances which he felt justified obliging future generations to partially pay the cost of benefits they did not choose, but had left to them by previous generations (Political Economy, 613).

When economists, philosophers and demographers write about overlapping (Engineer, Roth, & Welling, 2005) or non-overlapping generations (Wolf, 2003), they are making the distinction between concern for the generations alive today (overlapping) and concern for unborn generations (non-overlapping). Mill seems to have been unable to resolve this problem involving overlapping and non-overlapping generations which had burdened him from very early in his career (6, 240, 252).

Mill sees the length of a generation being approximately 25 years (3, 483) and this is in close alignment with the classical generational theorists discussed above. In harmony with the later writing of Marias, Mill believed the character of a generation was formed by the circumstances in which they grew up (24, 264). Having suggested various measures by which the younger generation could mature in improved circumstances Mill (Political Economy, 270) notes that:

> *The benefit would however arise, not from what was given them, but from what they were stimulated to acquire.*

It is clear Mill sees improvement being negotiated between older and younger overlapping generations, but not necessarily as equal negotiating partners. His recognition of generations within occupations and occupations within generations is an indication that he would have understood the modern concept of generational units and the complexity of such negotiations. The variety of occupations that he referred to include 'Political Economists' (Political Economy, 543), 'Statesman' (6, 243), 'Labourers' (sic) (4, 435), 'Painters' (14, 360), 'Irish farmers' (6, 584), 'Grecian poets' (11, 416), and 'Philosophers' (7, 231). It's clear that in some cases Mill also understood the need to locate generational units within their national contexts. Mill understood that cultural meanings are generated within countries, occupations and generations (8, 230). At one point Mill (21, 241) describes something like the generational cycles described by Turchin:

> *A new generation has grown up since the great victory of slavery abolition; composed of persons whose ardour in the cause has never been wrought upon and strung up by contest. The public of the present day think as their fathers did concerning slavery, but their feelings have not been in the same degree roused against its enormities.* (sic)

On these points Mill's thinking is in tune with that of 20th century sociology and early 21st century thinking about human Big History.

Identifying a generation

The task of identifying a generation is informed by contemporary demographic study. Not all authors agree on the boundary years between generations, but approximations are generally accepted if they fit within the descriptions above. Samples regarding Australian voting patterns and American Human Resource Management (HRM) are given for comparison with the Australian Bureau of Statistics and Australian demographers in Table 1. (ABS, 2009; Gray, Evans & Kippen, 2007; Zemke, Raines & Filipczak, 1999; Ruthven, 2004; Strauss & Howe, 1991, 84; Australian Bureau of Statistics, 2006).

In the Australian context this represents fertility cohort patterns similar to those described by

Australian Bureau of Statistics ABS	Gray, Evans & Kippen Australian Demography	Ruthven Australia Voting	Strauss & Howe USA	Zemke et al USA HRM
1891	1910	1901	1901	
1926	1925	1925	1925	1922
1946	1945	1943	1943	1943
1966	1965	1961	1961	1960
1986			1981	1980
2006				2000

Table 1 Generation boundaries of demographers and other authors.

Turchin (2003, 151-152).

The 'non-identity problem'

In the case of non-overlapping generations the 'non-identity problem' must be addressed (Page, 2006) because living generations can exploit their position (ibid.). Mill never reached a resolution of this problem and debate on it continues in fields concerned with social justice, the environment (Dobson, 1999; Woollard, 2012) and applications of technology to fertility (Weinberg, 2013; Gardner, 2015). It applies in societies where laws can only apply to people who can be clearly identified and this usually means the living. It is argued that if a law applied to the unborn, the problem thus created would be the need to determine exactly who among the people yet to be born would be harmed, and in what ways by the pursuit of the various alternatives now available to deal, for example, with climate change.

The non-identity problem was partially recognised in English law sometime prior to Mill's career:

> *It is now a considerable number of years since a London merchant having by testament directed that the bulk of his fortune should accumulate for two generations, and then devolve without restriction upon a person specified; this will, rare as such dispositions might be expected to be, excited so much disapprobation, that an Act of Parliament was passed, expressly to "enact" that nothing of the same sort should be done in future.* (Mill, 4, 259)

Some argue that because the rights of the unborn cannot be asserted in law, present generations need not concern themselves with their well-being (Gosseries, 2008; Weinberg, 2012). But examples have been given of 'historic injustice' to the descendants of slaves (Herstein, 2008) and the 'Stolen Generation' of original Australians (Dodson, 2007). In these cases governments have accepted responsibility for providing compensation many years after the original injustice. Part of the argument in favour of the historical injustice claim is that the original injustice is a separate issue from the continuation of a derivation of that injustice in the present. Therefore these instances are equivocal on the implications of the non-identity problem.

It has been argued that:

> *… we are right to keep our eye on how principles and practices affect individual persons and I will argue that, when we do, the non-identity problem does not arise.* (Weinberg, 2012)

The following section describes a civilisation in which this appears to have been done to good effect.

The pre-colonization society of the original Australians

This section considers the culture and civilisation of the original Australians, commonly called Aboriginals and Torres Strait Islanders. Based on archaeological evidence they are estimated to have settled in Australia at least 40,000 or 60,000 years ago or more (Mulvaney & Kamminga, 1999). During that time they have maintained a continuous, developing culture which would seem to have at least partially survived the devastating annexation of their lands by Europeans since 1788: this is evident in their public statements calling for a national treaty in which they have been encouraged by many, including the historian, Henry Reynolds

(Reynolds, 1996, 70, 114-116, 145, 151-153; Reynolds, 1989, 75, 87) the popularity of the music by Yothu Yindi (2016), announcements by two Australian state governments concerning treaty discussions (South Australian Treaty Discussions, 2017; Victorian Treaty Discussions, 2017) and the recent Uluru Statement from the Heart (2017).

It has been estimated that before 1788 the original Australians, in comparatively small social groups, occupied the whole continent of Australia including parts that would now be considered deserts and that there were literally hundreds of languages among them (Elkin, 1970, 17). The continuity in the civilisation of the original Australians has something to teach us in relation to the non-identity problem but it must first be distinguished from the empires studied by Turchin and Turchin & Nefedov and the capitalist economies studied by Figueroa.

The criteria Turchin used to define empires suitable for his analysis are that the society in question must be large, internally multiethnic, territorial and have a complex power structure and that a meta-ethnic frontier is present in the early stages (Turchin, 2006, 3). A meta-ethnic frontier is a place where an empire and foreign ethnie are geographic neighbours of a kind.

Figueroa defines capitalism as a set of social arrangements in which private property dominates over property held in common, the market is the dominant form of exchange and democracy, loosely defined, is the dominant form of governance (Figueroa, 2015, I, 27).

Our knowledge about the original Australians enables comparison with several defining aspects of empire and capitalism. Mulvaney tells us:

> *For longer than any other population of* Homo sapiens, *the ancestors of the Aborigines inhabiting Sahul, and later Australia, were genetically cocooned from the rest of humanity in Eurasia.* [Emphasis in original] (Mulvaney & Kamminga, 1999, 170)

Their closest neighbors were the Torres Strait Islanders to the north who:

> *resemble Papuans rather than Australian Aboriginals in physical features.* (ibid, 332)

There was trade, conflict and intermixing between the groups but the original Australians did not follow the agricultural practices of the Islanders (Gammage, 2012, 98) who may have adopted some of the original Australians' hunter-gathering practices (ibid, 300).

In relation to the power structures of original Australians Berndt & Berndt refer to them as 'stateless' (Berndt, & Berndt, 1981, 366) and Keen tells us that rules and norms were taken as given:

> *… because they were contained and transmitted in oral tradition and in the absence of centralised legislative and judicial institutions, the rules and principles of ancestral law had a large discretionary component.* (Keen, 2004, 244)

And that:

> *Networks of regional cooperation underpinned the sharing of ancestral law.* (ibid, 244)

Ancestral laws may have differed from community to community but all over Australia ancestral law was honoured. This law indicated that the original Australians belonged to the land rather than the land belonging to them (Berndt & Berndt, 1981, 135-149). The words "territory" and "property" had, therefore, a very different meaning for the original Australians than they do in empires or capitalist societies. The original Australians at the time of the early encounters between them and Europeans had no experience of a meta-ethnic frontier, they were an ethnically homogeneous population whose understanding of territory and property was completely different to ours and they were 'stateless'.

Therefore we can conclude that the society of the original Australians at the time of its early encounters with Europeans was neither an empire as Turchin would see it, nor a capitalist society as Figueroa would see it: it existed outside the mathematical models of human Big History discussed above.

Gammage has painstakingly reviewed many of the earliest records made by Europeans about the original Australians and the way they related to their landscapes. These records consist primarily of paintings and handwritten documents and he finds a substantial level of agreement between them regarding the lifestyle of the original Australians and the way they related to their landscapes. The original Australians controlled their lands and their populations so as to produce a surplus most of the time and enough to get them through the severest droughts (Gammage, 2012, 150-151). Most of his book concerns their land management practices: population management he mentions in passing indicating that a variety of laws were invoked and in extreme cases infanticide was practiced; something also noted by other authors (Marcus, 2015; Stormon, 1977, 136-137) and interpreted in a broader context as being 'essential for the survival of society' (Davies, 1981, 193). Mill was also concerned about population size and it has relevance to Big History (Christian, 2011, 311-312, 630). Gammage concludes:

> *1788's plant patterns were unnatural but universal. How people did this varied from region to region, but everywhere they made similar templates for similar purposes. Different lives, from Spinifex to rainforest, the Wet to the snow, coast to desert, obeyed a strict inheritance, followed the same Law, allied with fire and worked locally to make plants and animals abundant, convenient and predictable. They made a continent a single estate.* (Gammage, 2011, 280).

Land management practices of the original Australians had two salient features. Firstly, they used very detailed knowledge of botany, the impact of fire on different plants in different weather conditions, animal husbandry and the fire technologies that were available to them to create landscapes that were compared by the early European visitors to the estates of landed European gentry (ibid. 5-17). Secondly, given the level of technology available, to establish the land templates portrayed by Gammage the knowledge required must have been applied consistently over centuries (ibid. 41) after which only maintenance activity was required. However, even the maintenance activity in some cases required observance of plant cycles with the periodicity of approximately half a century (ibid. 27, 52). Despite being a non-literate civilization, the original Australians must have conveyed knowledge from generation to generation in ways that ensured it would be passed on to future generations so that it could be used at the appropriate times in these plant cycles, even if it was not always going to be necessary to use that information in the lifetime of some human generations.

It is noteworthy at this point to briefly consider the differences between the left and right hemispheres of the human brain and McGilchrist's (2009) argument that in the contemporary western world the dominance of the left hemisphere is responsible for many of our problems. The left hemisphere's primary focus is said to be on all that is familiar and it deals in concepts it can 'grasp'. On the other hand the right hemisphere engages with the outside world more, seeing everything in context. This is a gross oversimplification of McGilchrist's work but can serve the present purpose. Before analysing the cultural shifts in hemisphere dominance in the western world he describes the role of mimesis in cultural change and indicates that:

> *... we choose what we imitate.*
> (McGilchrist, 2009, 256)

McGilchrist describes writing as primarily a left hemisphere function (McGilchrist, 2009, 274-279): this implies that non-literate societies like that of the original Australians, compared to literate ones, lived with a different balance between the functioning of the left and right hemispheres of the brain.

Inspired by her understanding of how the original Australians handled knowledge, Kelly found she could apply their techniques profitably in her own life and that other ancient non-literate civilisations such as the Easter Islanders and the Stonehenge dwellers handled knowledge in similar ways (Kelly, 2016). Unlike the latter two which used large expensive man-made objects as their mnemonic props, the original Australians used features in the landscape. By these means they made the non-identity problem irrelevant. This kind of intergenerational relationship goes far beyond that mentioned in the Brundtland Report, discussed above, and challenges our current intergenerational practice.

A small international group from the academy and business are today concerning themselves with the well-being of future generations (Oxford Martin Commission for Future Generations, 2015). The way original Australians lived in 1788 is an example that makes it worthwhile for us to strive to meet the Big History challenge set by Spier:

> *... would we be able to tame both our biological instincts and social arrangements with the aid of culture?* (Spier 2015, 313)

Two additional fields of disciplined knowledge are relevant to this challenge and will be discussed in the following section.

Social Anthropology and Human Ethology.

The relationship between social anthropology and human ethology is complex. For historical reasons the field of human ethology will be discussed first.

An early 20th century dictionary (Porter & Harris, 1902, 513) definition of ethology indicates that the word is of Greek origin where it refers to character, custom and moral nature. It then provides two definitions:

> *1. A treatise on morality; ethics.*
>
> *2. The science of the formation of character, national and collective as well as individual.*

The source of the second definition is given as *J. S. Mill*.

An early 21st century definition given by the same authority (Websters, 2016) indicates French, Latin and Greek origins in which the relevant words are linked with character, speech and expression and provides the following two definitions:

> 1. *The scientific study of animal behavior, especially as it occurs in a natural environment.*
> 2. *The study of human ethos and its formation.*

No sources are given for these definitions.

It is evident that the link with Mill has been lost and the meaning of the word has been significantly broadened: the context of the early definition limits it only to humans while the recent one incudes the study of animals, privileging the latter by listing it first. The latter definition reflects the current existence of separate international ethology associations, the International Society for Applied Ethology (2016) for the study of animal behavior and the International Society for Human Ethology (2016).

The present work stops short of defining ethology in the Millean sense which is still seen as a work in progress to which Steinbruner (2011) has made a major contribution. Using keyword searches Steinbruner has been able to draw together and interpret Mill's publications succinctly and in a way that demonstrates their relevance to the present day. Steinbruner asserts that Mill sees:

> *… political units, patterns of human behavior and forms of social organization.* (Steinbruner, 2011, iii)

as co-constitutive of each other. Steinbruner sees in Mill's thought

> *… a framework that relates individual utility to a distinct but not independent utility of the whole.* (ibid.)

Steinbruner observes that Mill proposes

> *… a method of investigating this relationship as reflected in characteristic patterns of thought and behavior, a method - political ethology… .* (ibid.)

The ambitions of the International Big History Association (2015) announced prominently on their website, i.e. to contribute to human well-being in the present and near future can arguably be well served by pursuing concerns that Mill hoped his ethology project would address. Steinbruner's interpretation of Mill's ethology project sees it as building on our knowledge of psychology:

> *… in order to make reasonable generalizations about how specific conditions external to the individual contribute to the acquisition of specific dispositions.* [or traits] (Steinbruner, 2011, iii)

Mill himself made little progress with his ethology project but it was taken up by others. Leary (1982) describes its slow development, its demise in UK and simultaneous advance in France, and its return to the English-speaking world through the social anthropology of Radcliffe-Brown and Malinowski.

In the final footnote to his article Leary also identifies the original application of the term 'ethology' to animal behavior in 1907 and suggests that at the time of his publication, 1982, animal and human ethologies might be in the process of converging. In the same year Hinde (1982) published his book which relied heavily on the work of animal ethologists in its approach to human ethology, without reference to Mill. Seven years later, as if in response to Hinde, Eibl-Eibesfeldt (1989) published his epic work on human ethology which was also much influenced by the work of animal ethologists *and* without reference to Mill.

Based on this author's keyword search and analysis of Mill's works it appears that Mill's thoughts about generations closely approximate the thoughts of the authors referred to in the 'Generations observed' section above. However, Mill's frequent expressions of concern regarding "posterity" suggest strongly that he would not regard the "non-identity" problem as a valid reason to ignore the interests of the, as yet, unborn.

Mill's work, and particularly his ethology proposal has been controversial, like many of the other works cited here. At the end of his book which is broadly sympathetic to Mill, Miller mockingly dismisses the ethology project:

> *Mill's unfinished project, the construction of the great science of ethology, has yet to be carried out.*
> (Miller, 2010, 211)

However, Leary traces an unbroken line of ethological thought from Mill to the field of social anthropology. If Leary's conclusion is accepted, then Miller's final word on Mill's ethology can only be seen as historically correct but not a fair assessment of Mill's contribution to social science.

A comparison between the central concepts which are now applied in both animal and human ethologies and the things which are of greatest concern to 21st century social anthropologists can serve to illustrate the divergence between the two that began in the early 20th century.

Human and animal ethology today examine behavior from causal, developmental, functional and evolutionary perspectives (Hinde, 1982, 128-131; Barrett & Stulp, 2013). Time is an essential dimension in all of these perspectives except the functional and even in that one time may be involved. In ethology sentiments and tendencies were early key concepts but Leary notes that social anthropologists refer instead to interests, values, needs and satisfaction. Though all of these can change with time, time is not an essential dimension of their definitions. At the same time as Eibl-Eibesfeldt was publishing his epic work concerns were being raised about the need to rethink human ethology (Betzig, 1989). Betzig places the focus in human ethology on understanding mechanisms (For detailed discussion of social mechanisms see Hedström & Swedberg 2005) so that predictions can be made. While ethology today is imbued with a concern about the continuity of time and the making of predictions, social anthropology is much more interested in understanding the present. This understanding is pursued for its own sake in the hope that it may inform our approach to the future.

Recent work in social anthropology draws together insights from several synchronous anthropological observations to make general statements about social practices in the economy (Gudeman, 2016,183-189). The limitations of the synchronic frame are lamented by one social anthropologist (Lewis, 1981, 11) and the source of challenge for others:

> *The challenge, then, is not to do away with the synchronic ethnographic frame, but to exploit fully the historical within it.*
> (Marcus, & Fischer, 1986, 96)

The sequel to this article responds to that challenge.

Conclusion

The primary purpose of this work has been to situate human generations in Big History and to do this in a way that may lead to a more harmonious humanity. To do this the work had to address a range of controversial issues and this is not unusual in the Big History field. But without the cosmology offered by Big History, the arguments presented here would have been difficult to support.

On their own, the life cycles of empires and capitalism might be seen as reflections of an underlying evolutionary group selection process. However, the complexity of that concept and the controversy surrounding it (Okasha, 2013, 173-202) make it an unattractive proposition. After considering the example of precolonial original Australians, Popperian logic came into play and the entropy and the levels of selection hypotheses were falsified. In the absence of any other explanation for the survival of the original Australians' civilisation, and the differences between that civilisation and empires and capitalism, human agency must be considered as a potential explanation, qualified by due consideration of the balance between the left and right hemispheres of the brain. Current practice in human ethology is concerned more with behaviour than character. Bringing together the thoughts of Mill, Steinbruner and McGilchrist, one might ask: what might the young be 'stimulated to acquire' (Mill, Political Economy, 270) or disposed (Steinbruner, 2011, 144) to imitate (McGilchrist, 2012, 256)? The answer is to be found somewhere in inter-generational relationships.

Disciplined study of generations has been a comparatively narrow area relying on conventional social science disciplines. Informed by recent sympathetic scholarship on Mill and Mill's own writings this work has traced the ideas of Mill in general and his ideas concerning ethology in particular, through key moments down to the present field of social anthropology. After examining differences in emphasis with relation to time between social anthropology and the primary concepts in human and animal ethology, it may now become possible to tentatively reconstruct a model of Mill's ethology project focusing on the character of generations in a way that addresses the Big History goal of human harmony. I submit that we need to re-unite Mill's interest in the development of character with temporal references through the study of inter-generational relationships. With the benefit of advances in social science that have taken place in the century and a half since Mill, a completed attempt to do this in practice will be the subject of a sequel to this article.

It is to be hoped that this article will also lead to many other avenues of fruitful, intuitive speculation so that the best approach to the paradox which has focused this work can be found.

References

ABS. Australian Bureau of Statistics. (2009) 'From Generation to generation' in *A Picture of the Nation.* http://www.ausstats.abs.gov.au/Ausstats/subscriber.nsf/0/FCB1A3CF0893DAE4CA25754C0013D844/$File/20700_generation.pdf 19/05/17.

ACOSS. Australian Council of Social Service. (2014) Poverty in Australia. http://www.acoss.org.au/images/uploads/ACOSS_Poverty_in_Australia_2014.pdf 18/04/2017.

ACOSS. Australian Council of Social Service. (2015) Inequality in Australia. http://www.acoss.org.au/wp-content/uploads/2015/06/Inequality_in_Australia_FINAL.pdf 18/04/2017.

Adamo S. (2003) 'Social sustainability and social resilience of rural communities in drylands: the case of the Jachal (Argentina) in the 19th and 20th centuries.' (Unpublished work, quoted with permission) Mellon Foundation. Population Research Centre and Department of Sociology. University of Texas, Austin.

Atkinson A. (1991) Comparing Poverty Rates Internationally: Lessons from Recent Studies in Developed Countries: *The World Bank Economic Review* 1(1)3-21.

Baker D. (2015) Speaker Profile, Australian Institute of Health Innovation Event, Macquarie University. https://aihi.mq.edu.au/event/big-history-population-and-collective-learning-8 05/12/2015.

Barrett L & Stulp G. (2013) Four More Reasons Why Ethology Matters: Tinbergen and the Human Sciences. *Human Ethology Bulletin* 28(4)39-49. https://www.researchgate.net/profile/Louise_Barrett/publication/259070028_Four_more_reasons_why_ethology_matters_Tinbergen_and_the_human_sciences/links/00463529def3e60eb3000000.pdf 20/04/2017.

Berndt R & Berndt C. (1981) *The World of the First Australians.* Lansdowne Press, Sydney.

Betzig L. (1989) Rethinking Human Ethology: A Response to Some Recent Critiques. *Ethology and Sociobiology* (10)315-324 (0162-3095).

Brundtland GH. (1987) *Our Common Future.* Report of the World Commission on Environment and Development, United Nations, NY, USA. https://doi.org/10.1017/S0376892900016805

Christian D. (2003) World history in context. *Journal of World History* 14(4)437-458. http://www.jstor.org/stable/20079239?origin=JSTOR-pdf 03/03/2010.

Christian D. (2011) *Maps of Time: An Introduction to Big History.* University of California Press.

Daley J, Wood D, Weidman B & Harrison C. (2014) *The wealth of generations*. The Grattan Institute, Melbourne, Australia. http://grattan.edu.au/report/the -wealth-of-generations/ 22/07/2016.

Davies N. (1981) *Human Sacrifice.* William Morrow.

Damasio A. (2003) *Looking for Spinoza*. Vintage Books, London.

Dobson A. (1999) *Fairness and futurity: essays on environmental sustainability and social justice.* Oxford University Press.

Dodson M. (2007) 'Bully in the Playground: A New Stolen Generation', In *Coercive Reconciliation: Stabilise, Normalise, Exit Aboriginal Australia.* Altman J & Hinkson M, Eds, Arena, Melbourne, 85-96.

Eibl-Eibesfeldt I. (1989) *Human Ethology.* Aldine de Gruyter, New York.

Elkin A. (1970) *The Australian Aborigines: How to Understand them.* Angus & Robertson.

Engineer M, Roth E, & Welling L. (2005) *An almost exact overlapping generations model of the Rendille of Northern Kenya.* Work in progress, University of Victoria, British Columbia. http://webuvicca/~menginee/mypapers/Exactpdf 01/02/2005.

Figueroa A. (2015) *Growth, employment, inequality and the environment: Unity of knowledge in economics.* (Two volumes) Palgrave Macmillan, NY USA.

Gammage B. (2012) *The Biggest Estate on Earth: How Aborigines made Australia.* Allen & Unwin.

Gardner M. (2015) A Harm-Based Solution to the Non-Identity Problem. *Ergo: An open access journal of philosophy* (2)17. http://dx.doi.org/10.3998/ergo.12405314.0002.017 17/08/2016.

Gillings M, (2017) Class 1 integrons as invasive species. *Current Opinion in Microbiology* (38)10-15. https://doi.org/10.1016/j.mib.2017.03.002 26/02/2018.

Gillings M, Paulson I & Tetu S. (2015) Ecology and Evolution of the Human Microbiota: Fire, Farming and Antibiotics. *Genes* (6)841-857. doi:10.3390/genes6030841. http://www.mq.edu.au/newsroom/2015/08/06/antibiotic-pollution-of-waterways-may-create-superbugs-of-tomorrow/#ixzz3tNaxwpnG 05/12/2015.

Goertzel T. (2006) The World Trade Center bombing as a fourth generational turning point http://crabrutgersedu/~goertzel/fourthturninghtm 03/10/2006.

Gordon D & Townsend P. (1990) *Measuring the Poverty Line*. http://www.radstats.org.uk/no047/gordontownsend.pdf 29/04/2016.

Gosseries A. (2008) On Future Generations' Future Rights. *The Journal of Political Philosophy* 16(4)446-474. doi: 10.1111/j.1467-9760.2008.00323.x

Gray E, Evans A & Kippen R (2007) *Generational differences in Parity Progression: the role of Sex Composition of Children.* The Australian Demographic & Research Institute, Working Paper 2. http://adsrianueduau/ 03/12/2008.

Grundy E & Henretta JC. (2006) Between elderly parents and adult children: a new look at the intergenerational care provided by the 'sandwich generation'. *Ageing & Society* (26)707-722. doi 10.1007/s10680-006-9005-6.

Gudeman S. (2016) *Anthropology and Economy.* Cambridge University Press.

Hawking S. (1996) *A Brief History of Time*. Bantam Books.

Heath R. (2006) *Please just F* off: it's our turn now - holding the baby boomers to account.* Pluto Press, Australia.

Hedström P & Swedberg R. (2005) *Social mecha-*

nisms: An analytical approach to social theory. Cambridge University Press.

Henderson R. (1975, April) *Poverty in Australia: Commission of Inquiry into Poverty, First Main Report*. Australian Government Publishing Service, Canberra.

Henderson R, Harcourt A & Harper R. (1970) *People in Poverty: A Melbourne Survey*. Melbourne Institute of Applied Economic and Social Research [IAESR]. Cheshire Publishing, Melbourne.

Herstein O. (2008) *Historic Injustice, Group Membership and Harm to Individuals: Defending Claims for Historic Justice from the Non-Identity Problem.* Columbia Public Law & Legal Theory Working Papers, Paper 08150. http://lsr.nellco.org/columbia_pllt/08150 17/12/2016.

Hinde R. (1982) *Ethology: Its Nature and Relations with Other Sciences.* Oxford University Press.

International Big History Association. http://www.ibhanet.org/ 05/12/2015.

International Society for Applied Ethology. www.applied-ethology.org 06/10/2016.

International Society for Human Ethology. http://ishe.roundtablelive.org/ 06/10/2016.

Jaeger H. (1985) Generations in History: Reflections on a Controversial Concept. *History and Theory* 24(3)273-292. http://www.jstor.org/stable/2505170 26/11/2009.

James P. (2006) *Globalism, Nationalism, Tribalism: Bringing Theory Back In.* Sage.

Kanbur R. (2001) Economic Policy, Distribution and Poverty: The Nature of Disagreements. *World Development* 29(6)1083-1094. https://www.sciencedirect.com/science/article/pii/S0305750X01000171 02/03/2016.

Keen I. (2004) *Aboriginal Economy and Society: Australia at the Threshold of Colonisation.* Oxford University Press.

Kelly L. (2016) *The Memory Code: The traditional Aboriginal memory technique that unlocks the secrets of Stonehenge, Easter Island and ancient monuments the world over.* Allen & Unwin.

Kewley T. (1973) *Social Security in Australia*. Sydney University Press.

Lancaster L & Stillman D. (2005) *When generations collide. Who they are. Why they clash. How to solve the generational puzzle at work.* Collins Business, Harper Collins, New York.

Leary D. (1982) The Fate and Influence of John Stuart Mill's Proposed Science of Ethology. *Journal of the History of Ideas* (43)153-162.

Lewis I. (1981) *Social Anthropology in Perspective: The Relevance of Social Anthropology.* Penguin books.

Mannheim K. (1952) The Problem of Generations, in *Essays on the sociology of knowledge.* Oxford University Press, London. 276-322.

Marcus G & Fischer M. (1986) *Anthropology as Cultural Critique: An Experimental Moment in the Human Sciences.* University of Chicago Press.

Marcus J. (2015) 'Independent Thinking and Difficult Ideas: A Few Preliminary Thoughts on Infanticide', Conference Proceedings: Independent Scholars Association of Australia Inc. *Celebrating Independent Thought: ISAA Twenty years on*. 87-98.

Marias J. (1970) *Generations: a historical method.* University of Alabama Press, USA.

McCrindle M. (2009) McCrindle Research. http://www.mccrindlecomau/aboutushtm 19/08/2009.

McGilchrist I. (2012) *The Master and his Emissary.* Yale University Press, New Haven & London.

McQueen M. (2008) *The 'new' rules of engagement: a guide to understanding and connecting with generation Y.* Nexgen Impact, Sydney.

Melbourne Institute of Applied Economic and Social Research. (2016) https://melbourneinstitute.com/miaesr/publications/indicators/poverty-lines-australia.html 19/07/2016.

Mill JS. *Complete Works.* http://oll.libertyfund.org/people/john-stuart-mill?q=Mill# 06/10/2016. All references to Mill relate to the name or volume number of the .pdf files available at this site and their page numbers.

Miller D. (2010) *J. S. Mill.* Polity, Cambridge.
Mulvaney J & Kamminga J. (1999) *Prehistory of Australia.* Allen & Unwin, Sydney.
Nixon P. (2009) Application of engaged social research principles in inter-generational aged care workforce research. *The International Journal of Interdisciplinary Social Sciences* 4(1)77-88. https://doi.org/10.18848/1833-1882/CGP/v04i01/52791 09/03/2018.
OECD, Organisation for Economic Co-operation and Development. (2001) *Poverty reduction*. http://www.oecdbookshop.org/en/browse/title-detail/Poverty-Reduction/?K=5LMQCR2KGV7H 09/12/2016.
OECD. (2008) *Growing unequal? Income distribution and poverty in OECD countries*. http://www.oecdbookshop.org/oecd/display.asp?CID=&LANG=EN&SF1=DI&ST1=5KZRB41JLHJD 30/11/2011.
OECD. (2011) *Divided we stand: Why inequality keeps rising.* http://www.oecd-ilibrary.org/social-issues-migration-health/the-causes-of-growing-inequalities-in-oecd-countries_9789264119536-en 07/12/2016.
OECD. (2015) In It Together: Why Less Inequality Benefits All. http://www.oecd-ilibrary.org/employment/in-it-together-why-less-inequality-benefits-all_9789264235120-en 07/12/2016.
Okasha S. (2006) *Evolution and the Levels of Selection.* Oxford University Press, Oxford, UK.
Osborne SP. (2010) *The new public governance? Emerging perspectives on the theory and practice of public governance*. Routedge, New York, USA.
Oxford Martin Commission for Future Generations. http://www.oxfordmartin.ox.ac.uk/ 08/12/2015.
Page E. (2006) *Climate change, justice and future generations.* Edward Elgar, Cheltenham, UK.
Porter N & Harris W [Editors]. (1902) *Webster's International Dictionary of the English Language, Australasian Edition.* G and C Merriam Company, Springfield, Massachusetts, USA.
Rayner J. (2016) *Generation Less: How Australia is Cheating the Young.* Redback Quarterly: Schwartz Publishing, Melbourne.
Reynolds H. (1989) *Dispossession: Black Australians and White Invaders.* Allen & Unwin, Sydney.
Reynolds H. (1996) *Aboriginal sovereignty: Three nations, one Australia?* Allen & Unwin, Sydney.
Ruthven P. (2004) *Australian Families and Work: What are the Trends telling us?* Families Australia Conference Presentation, Brisbane. http://www.ibisworld.com.au/about/philruthven/ 13/05/2004.
Seedsman T. (2006) Viewing Participants as Resources for One Another, Communities and Societies: Intergenerational Solidarity Toward a Better World. *Journal of Intergenerational Relationships* 4(1)23-39. http://www.tandfonline.com/doi/abs/10.1300/J194v04n01_04 21/02/2017.
South Australian Treaty discussions. http://www.statedevelopment.sa.gov.au/aboriginal-affairs/aboriginal-affairs-and-reconciliation/initiatives/treaty-discussions 19/01/2017.
Spier F. (2015) *Big History and the Future of Humanity.* Wiley Blackwell.
Spier F. (2017) On the pursuit of academic research across all disciplines. *Journal of Big History* 1(1)20-39, https://jbh.journals.villanova.edu/article/view/2242.
Steinbruner M. (2011) *Comprehending the Polity: John Stuart Mill and the Utility of the Whole.* ProQuest, Ann Arbor, Michigan, USA. https://repository.library.georgetown.edu/bitstream/handle/10822/553108/steinbrunerMaureen.pdf?sequence=1 20/04/2017.
Stormon E. (1977) *The Salvado Memoirs.* University of Western Australia Press.
Strauss W & Howe N. (1991) *Generations: The History of America's Future, 1584 to 2069.* William Morrow, New York.
Turchin P. (2003) *Historical Dynamics: Why States Rise and Fall*. Princeton University Press, New Jersey, USA.
Turchin P. (2006) *War and Peace and War: The Rise*

and Fall of Empires. Plume, Penguin, USA.
Turchin P. (2008) Arise, 'cliodynamics'. *Nature* (454)34-35, https://www.nature.com/articles/454034a 21/02/2017.
Turchin P & Nefedov S. (2009) *Secular Cycles.* Princeton University Press, New Jersey, USA.
Uluru Statement from the Heart (2017) https://www.referendumcouncil.org.au/sites/default/files/2017-05/Uluru_Statement_From_The_Heart_0.PDF 30/05/2017.
Valletta R. (2006) The Ins and Outs of Poverty in Advanced Economies: Government Policy and Poverty Dynamics in Canada, Great Britain, Germany and the United States. *Review of Income and Wealth* 52(2)261-284. http://roiw.org/2006/2006-13.pdf 20/04/2017.
Victorian Treaty discussions. http://consult.aboriginalvictoria.vic.gov.au/Open-Meeting 19/01/2017.
Walsh L & Black R. (2011) *In Their Own Hands: Can Young People Change Australia?* ACER Press, Camberwell, Australia.
Webster's. (2016) http://www.yourdictionary.com/ethology#websters 20/09/2016.
Weinberg R. (2012) Existence: Who Needs It? The Non-Identity Problem and Merely Possible People. *Bioethics* 27(9)471-484. doi:10.1111/j.1467-8519.2012.01976.x http://hub.scrippscollege.edu/directory/wp-content/uploads/sites/3/existence-whoneedsit.pdf 20/04/2017.
Wolf C. (2003) 'Intergenerational Justice', in Frey RG & Wellman CH [Editors] *A companion to applied ethics.* Blackwell, Malden, MA, USA, 279-294.
Woollard F. (2012) Have we solved the non-identity problem? *Ethical theory and moral practice* 15(5)677-690, https://link.springer.com/article/10.1007/s10677-012-9359-2 21/02/2017.
Yothu Yindi. http://www.yothuyindi.com/music/treaty.html 20/09/2016.
Zemke R, Raines C & Filipczak B. (1999) *Generations at Work: Managing the Clash of Veterans, Boomers, Xers, Nexters in your workplace.* AMACOM, McGraw-Hill.

Acknowledgement

This article arises from a two-phase project. The first phase, from 2006-2009 was a funded workforce research project supported by The University of South Australia and the South Australian Government through Domiciliary Care SA, with reporting finalized in 2011 but not published. In the second phase, from 2013-2016 the author engaged in independent privately funded study. Sections above that draw heavily on phase one are the rise and fall of empires, Generations (excluding discussion of Mill) and the non-identity problem. Special thanks to my wife, Christina, and close friends who have seen me through this from the beginning. The author gratefully acknowledges the institutional support received in phase one and the present work owes much to support given by Paul James and Adolfo Figueroa, detailed, thoughtful and thought-provoking feedback on the manuscript from Peter Coyle and helpful suggestions offered by Lowell Gustafson. Any errors are mine alone.
Peter.nixon@isic.ekit.com

Aging and Evolution

Guibert U. Crevecoeur
Directorate General of Energy of the Belgian Federal Public Service of Economy – Brussels, Belgium (ret.)

Abstract
Similar empirical patterns of behavior are found in seemingly very different fields which once put together account for a « big history » approach of aging and evolution. It is shown that known curves for the creep of metals, the reliability of industrial systems, the mortality and survival of biological entities, learning, the general adaptation syndrome, (neo-)Darwinian evolution and the expansion of the universe can be compared and allow to conceive a general adaptation framework of complex adaptive systems. Equations are proposed for four shapes of curves found in laboratory and/or observational practice which can then be transformed into each other in function of the properties to be put into evidence. An entropy production can be linked to the phenomena under scope. The « big history » approach allows to go from global to peculiar. It thus shows the importance of using the correct orders of magnitude for the involved parameters (time, stress, temperature, energy …) when analysing a specific phenomenon.

Key words
Aging – Evolution – Creep – Reliability – Bathtub curve – Mortality curve – Survival curve – Learning – Feedback loop - General Adaptation Syndrome – Darwinian evolution – Punctuated Equilibrium - Expansion of the Universe – Complex Adaptive System – Entropy

Correspondence | Guibert U. Crevecoeur, guibert.crevecoeur@hotmail.com
Citation | Crevecoeur, G. U. (2018) Aging and Evolution. *Journal of Big History*, II(2); 79 - 102.
DOI | http://dx.doi.org/10.22339/jbh.v2i2.2303

INTRODUCTION

While Big History has only been named as such within the last few decades, [1-3] its recent analysts have recognized their indebtedness to Alexander von Humboldt, H. G. Wells and others from earlier generations. [4,5] I remember having day-dreamed when I was 14 that it would be nice to be able to narrate the evolution of my family over say 30 or 50 generations to understand my immediate, present day family. I found appealing the idea that there would be connections between the past and more recent events in this family. My goal was to see how earlier generations continued to exert their influence. I would search for a long path from what this family started with and went through to what this family had become after centuries. This was my « big history » dream!

Later, I earned an engineering degree after having studied several scientific subjects in a rather compartmentalized way. These studies allowed me to go deeply into well defined and focused scientific fields. I gained an understanding of the state-of-art methods and how science proceeds slowly, step by step.

I have more recently sought to integrate the musings of my youth with how different fields can be put together in a global approach of aging and evolution to produce a «big history» perspective.

1. FROM CREEP TO THE RELIABILITY OF SYSTEMS

After my education, I worked in a research centre studying creep of steels at very high temperatures (1473 – 1523 K); I focused on this field even though the topic usually receives brief attention in most curricula. Creep is the slow deformation of materials under stress at given temperature. It is usually taught very briefly in the field of physical metallurgy; it is often given about ½ hour in a 5 years engineering programme.

Afterwards, I went into a licensed inspection agency and was involved in the follow-up of fossil-fired thermal power plants. In those times, the issue at stake was the residual life assessment of that kind of plant. These plants had been designed for their thermal components (boiler, pipes, headers, turbine …) to operate during 100,000 hours under creep at around 813 K (or more). However, as they reached this time limit seemingly in good working order, it was not clear if they could be operated any longer. How could their residual life beyond 100,000 hours be assessed? So I started investigating creep curves for temperatures of the order 813 K and was astonished to find that although the microscopic structure of the metal was completely different in function of the temperature, the creep curves were alike for 813 K and 1473 K. For instance, the crystalline microstructure is face-centered cubic at 1473 K and body-centered cubic at 813 K, the mechanism of creep was different, the diffusion of atoms and vacancies in the microstructure varied a lot, etc. But the creep curves were alike in shape. Only the duration of the process was strongly different : of the order of 100,000 hours at 813 K and of a few seconds at 1473 K. Fig. 1 shows a typical creep curve.

This kind of curve is found in the creep of metals subjected to load or stress[1] at higher temperature (only the first part is seen at lower temperature). It can be interpreted as follows: (1) firstly, after the « shock » of being put under stress by the testing machine, the test piece shows a progressive adaptation : it is strained and the strain curve (the curve reflecting the « reaction » in Fig. 1) is concave from below (« primary creep »). This stage is

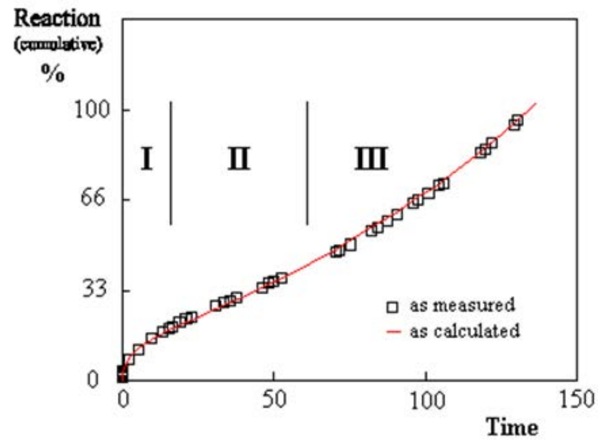


Fig. 1 Creep curve of a steel at higher temperature showing three stages : I) **a first concave part (« primary creep ») followed by II) a quasi-linear part (« secondary creep ») and III) a third convex part (« tertiary creep »), before rupture (arbitrary units).**

followed by (2) a second stage called « steady or stationary stage », where there is a balance between the defence reaction (strain) and the defects caused by the stress ; the strain curve is (quasi-)linear. If the stress is maintained, (3) a final stage develops with the exhaustion of the test piece up to rupture ; the strain curve is then convex from below.

In course of further investigations, it appeared that a good mathematical description of such creep curves was given by following equation for the strain [7-10]:

$$Æ(t) = k.e^{\alpha.t}.t^{\beta} \qquad (1)$$

With :

$k > 0$
$0 < \alpha << \beta < 1$
t : time (s)
k can be related to the applied stress « σ » by the expression: $k=k'.\sigma^b$
k', b, α and β are constants at given temperature, metallic structure and composition (they may change when temperature and/or metallic structure and

1 In the specific meaning used in the field of material testing.

composition change).

Because of the multidisciplinary approach adopted in the present article, I'll make use of the diphthong letter Æ(t) to express that it is a measured parameter reflecting an « **A**ging » or an « **E**volution » which develops in function of time (the reason for this choice of words will become obvious in the course of the text). In the case of creep, this measured parameter Æ(t) is the creep strain.

Eq. (1) is not the more frequent manner of describing creep curves. The usual way consists in focusing on the stages themselves, e.g. : $y(t)=t^{1/3}$ for the first concave part of the curve [11], $y(t)=K+a.t$ for the second linear part of the curve corresponding to the secondary creep [12], which is called the « useful life » in practical applications, and a growing exponential for the tertiary creep, etc.

But Eq. (1) has the advantage to combine all three stages. Moreover, in addition to fitting creep curves, it allows to easily derive a formula for the classical Logσ vs. Logt (logarithm of the stress vs. logarithm of the rupture time) diagrams [13,14] as widely used in design standards, what the default presentation can not. And as we shall see, Eq. (1) will also allow for comparisons with observations in other fields.

Indeed the derivative of Eq. (1) gives the strain rate in function of time:

$$\frac{dÆ(t)}{dt} = k.\beta.t^{\beta-1}.e^{\alpha.t}.\left(1+\frac{\alpha.t}{\beta}\right) \quad (2)$$

The shape of the corresponding curve is shown on Fig. 2. The first part of the curve (for small times, i.e. when $\alpha.t\approx 0$) is like the learning curve of Duane [15].

Several observations were made by Duane in the 1960's. He investigated repairable mechanical systems and found that their failure rate ($\lambda_{1,2}$) always obeyed a law of decrease with time: $\lambda_{1,2}=k.\beta.t^{\beta-1}$.

This meant that the number of failures of the system over a period of time decreased when time was going on. This behavior induced the feeling that the system had "learned" from its past experience and had found solutions to make less mistakes. Therefore, Duane called this behavior "learning" and the curve given by $\lambda_{1,2}$: the "learning curve". He found that, in general, one had: $\beta\approx 0.4\rightarrow 0.5$.

The systems considered were repairable mechanical devices such as airplane generators, submarine diesel engines, hydromechanical appliances.

Other authors after Duane reported similar behavior on a broad range of devices, e.g. loading cranes [16], army trucks [17] etc. and it is since then a widely accepted evidence that the learning curve can be observed on many operating mechanical devices, but also electr(on)ical devices, etc.

Using Eq. (1), Eq. (2) can also be written:

$$\frac{dÆ(t)}{dt} = Æ(t).\left(\alpha+\frac{\beta}{t}\right) \quad (2')$$

It is noteworthy that Fig. 2 is the derivative of the creep curve but also shows a typical « bathtub curve » as known in the reliability analysis of mechanical, electrical, … systems. This kind of curves is found for a lot of systems : engines, cars, …. During the life of these systems, there is a first phase of adaptation where the failure rate is first elevated and decreases (« infant illnesses »). Then the failure rate is around a minimum (« true life »). After a time, the failure rate starts to increase again first slowly then more steeply up to collapse of the system (« aging »).

In the frame of reliability analyses of systems, Eq. (2) can be seen as the failure rate for a modified Weibull distribution [18,19]. The modification lies in the adding of an $e^{\alpha.t}$ factor. The Weibull distribution is found back when α is put = 0. This modified Weibull

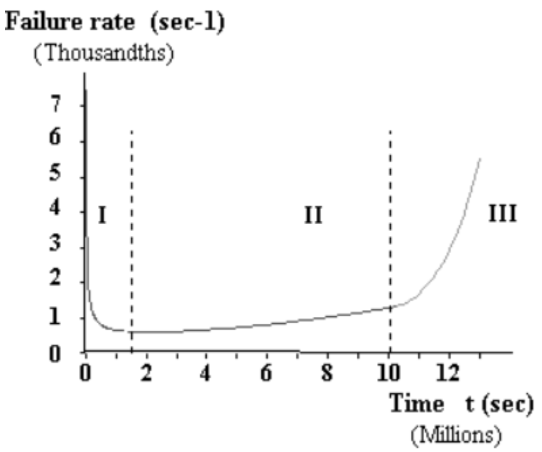


Fig. 2 Typical bathtub curve (arbitrary units)

distribution may be considered as a generalization of many other models [20]. Observe that the failure rate function of Eq. (2) can be considered as a sum of two Lee's failure rate functions [21] by noting :

$$\frac{dy(t)}{dt} = k.\beta.t^{\beta-1}.e^{\alpha.t} + k^*.\beta^*.t^{\beta^*-1}.e^{\alpha.t} \qquad (3)$$

where $k^* := \alpha.k/(\beta + 1)$, $\beta^* := \beta + 1$ [21]. The parameter β is called the *characteristic of the system*. When α = 0, the model reduces to a Weibull process; when α = 0 and β = 2, the model reduces to the Rayleigh process ; when β = 1 and α.t is very small, the model is approximately an exponential model. When β = 0, the model reduces to a type I extreme-value model, also known as a log-gamma model. As pointed out in [22], the model can be considered as a limiting process of the Beta integrated model introduced in [23].

The reliability corresponding to the bathtub curve model given by Eq. (2) is by definition given by Eq. (4) and visualized on Fig. 3.[2]

$$R(t)=e^{-Æ(t)} \qquad (4)$$

2. In [20], it is proposed to use a constant multiple of the sum of absolute residues of the cumulative failure rates (SAR) to assess how well a given model fits the intensity of a failure data set. The reliability-related decision and prediction were also studied using the bathtub curve model given by Eq. (2). The relations of the reliability characteristics during the improvement phase to those of the steady service phase are studied, which makes possible predicting the evolution behavior of a system by using the limited data observed during the improvement phase.

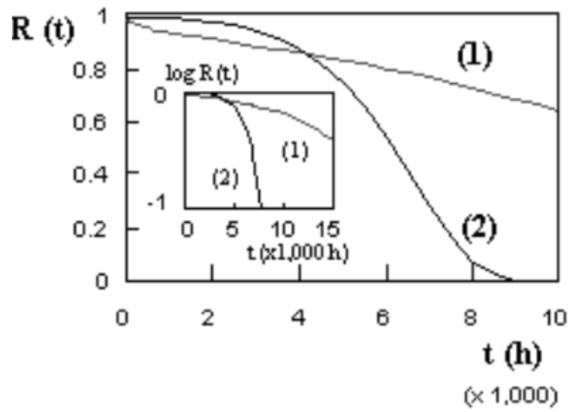


Fig. 3 - Examples of reliability curves.

The system is fully reliable in the beginning (R(t)≈100%), but with the time going its reliability decreases, first slowly then steeply down to zero.

The loads/stressors in mechanical and electr(on) ical devices are the constraints and environmental challenges due to operating.

2. GROWTH AND DECLINE CURVES

In the same way that allows the reliability to be given by the exponential of minus the « creep curve », interesting curves are obtained using the exponential of minus the « bathtub curve ». We shall call such curves « growth and decline curves » (GD-curves).[3] Following equation is used:

$$GD(t) = C'.e^{-c.\frac{dÆ(t)}{dt}} \qquad (5)$$

Where : « c » or « C' » are constants > 0 and can be = 1.[4]

Growth and decline curves (GD-curves) are found in constant strain rate tensile tests (CSRTT) of metals

3. They appear to be very general (as are bathtub and reliability curves) and could therefore also be called « capacity curves » or « resistance curves » or « force curves », when one wishes to insist on a « power » characteristic of a particular system under scope.

4. For instance, « C'» was put C'=1 to obtain the curves of Fig. 4.

at high temperature. These tests are the counterpart of creep tests. During creep, the stress/load is constant and the strain is recorded. Then the strain rate varies as shown by the bathtub curve. In CSRTT, the strain rate is maintained constant and the load is recorded. This gives GD-curves. As seen on Fig. 4, these curves show three stages : (1) first there is an increase of a « capacity », then (2) the curve reaches a maximum (peak) before to (3) progressively decline to levels encountered in the first stage. This is like climbing and descending an hill. Two typical shapes are given on Fig. 4. However, the shape can be different - with e.g. smoother growth and steeper decline - in function of the values of the constants « c » and « C' ».

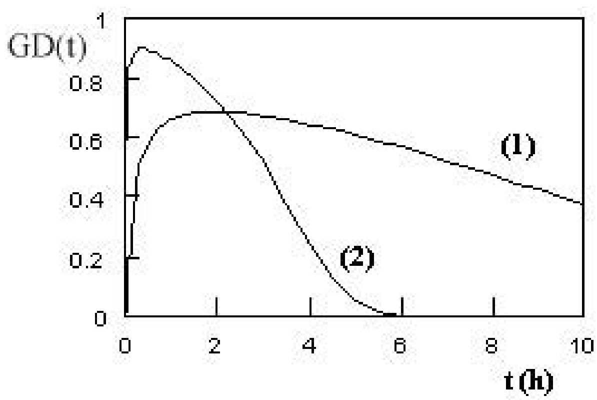


Fig. 4 - Examples of growth and decline curves.

As will be seen later on in the text, GD-curves are also found in biology and astronomy, and probably in other fields.

3. <u>FROM RELIABILITY TO BIOLOGY</u>

When one reads the descriptions by Hans Selye [24-28] and others of the General Adaptation Syndrome (G.A.S.) they are - *mutatis mutandi* - astonishingly close to the way creep curves would be described. The G.A.S. is the non specific reaction of the biological system to external attack. Let us give the word to Hans Selye :"...[the general adaptation] syndrome is...[an]...expression of **general defence divided into three stages**. During the first, or acute stage,observed in the rat ordinarily 6 to 48 hours after the **initial injury**, one notes a rapid decrease in the size of the thymus, spleen, lymph glands and liver... **After a few days, however, a certain resistance is built up against the damaging stimulus**...the animals **became resistant...If ...[the stressors]... were continued still longer the animals lost their resistance, and in a third stage died** with organ changes similar to those seen in the first stage... **We have termed...[the] 3 stages: the stage of alarm, the stage of resistance and the stage of exhaustion**..."[25] (emphasis added).

Concerning the human body, one can refer to the summary given by the Counselling Connexion (the Official Blog of the Australian Institute of Professional Counsellors) [29]. It starts with the words :

"The G.A.S. describes a three stage reaction to stress: (1) initial alarm / reaction to the stressor, (2) resistance / adaptation to coping and (3) eventual exhaustion."

In addition, it is noteworthy that Selye made a correlation between the G.A.S. and aging [27].

Reading such kind of descriptions and features pushed me decide to earn an additional degree in molecular biology.

A way to visualize the G.A.S. is shown on Fig. 5.

This is similar to a growth and decline curve. Probably with precise measurements, one would get a curve like the one shown in Fig. 6.

This induced to think that there would probably also exist curves of the kind bathtub curve, reliability curve or creep-like curve to be found in biology as far as aging and evolution are concerned. And indeed a lot of examples are found in the literature.

Just to start with mentioning all curves currently called « mortality curves » which are in fact bathtub

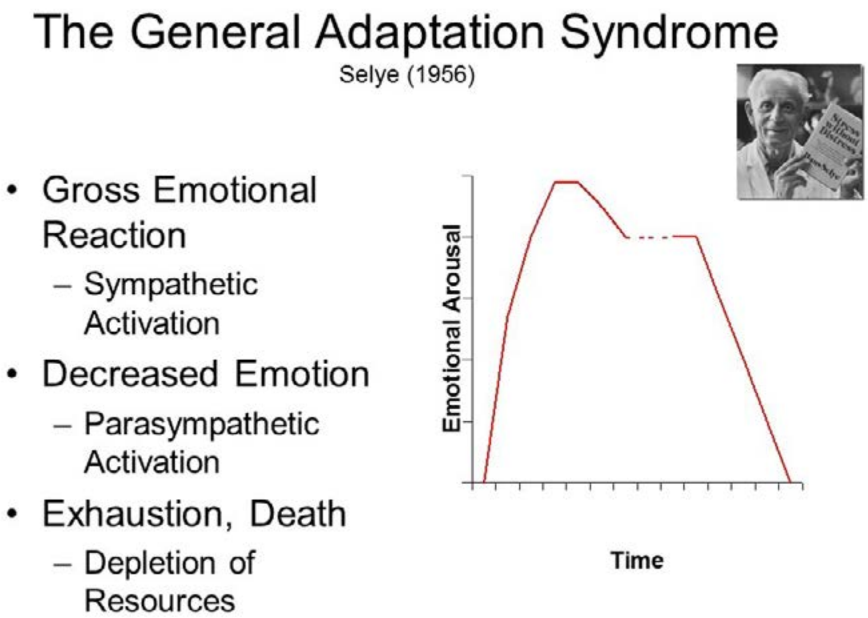


Fig. 5. Example of representation of the General Adaptation Syndrome (found on Internet).

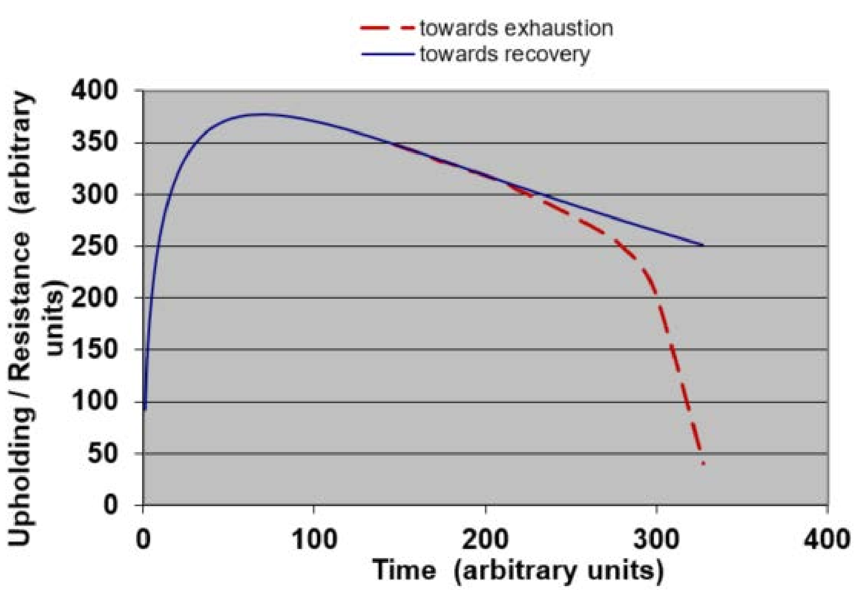


Fig. 6. Example of representation of the General Adaptation Syndrome following Eq. (5) – The blue plain curve corresponds with Eq. (5) when the stressor holds but is sufficiently low to allow the system to recover after removal of the stressor – The red interrupted curve shows what happens when the stressor is that high that the system will soon be harmed and subject to premature exhaustion.

curves, see e.g. Fig. 7. Mortality curves give the rate of mortality of a population in function of the age. They thus also reflect the probability of death at a given age of a lambda individual in this population. This is equivalent to the failure rate of systems as given by the bathtub curve. The fact that the increase of mortality rate is quicker as measured than as calculated is an often seen feature. As will be noticed later on, it is due to the fact that instability can settle after a time $t_i=1/\alpha$ (instability time) from which the growing exponential in Eq. (1) and (2) becomes higher than « e ». [5]

And all « survival curves » or « survivorship curves » found in biology are in fact reliability curves, see e.g. Fig. 8. But creep-like curves are also found, e.g. in tumour growth as shown on Fig. 9.

And finally, growth and decline curves are encountered in biology as already shown by the example of the G.A.S., but also in several other cases : changes in bone mass [33], muscular force [34], maximal oxygen uptake in healthy fit men [35], incidental memory scores [36], cumulative increase in human diploid fibroblast cells number [37], immune functions (e.g. antibody formation) in humans and animals, … Each time there is (1) an increase of a « capacity » (the bone mass, the muscular force, the maximal oxygen uptake, incidental memory scores, immune functions, average speed of sprinter on 100 meters, etc. etc.) up to a maximum where (2) it is about maintained a while

5. It is general feature in physics that instability may occur when $x \geq 1$ in a growing exponential e^x.

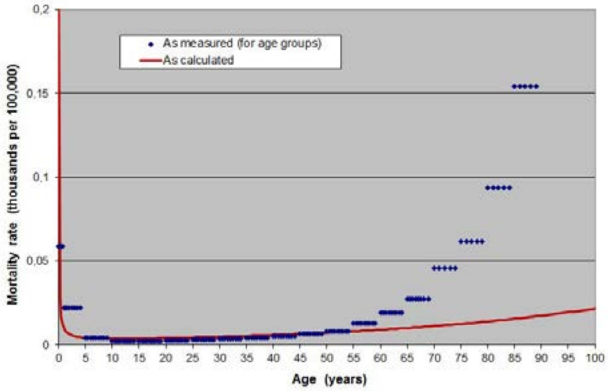


Fig. 7. Mortality rates per 100,000 for women in China by age group for the year 1957 [30]. Rhombs correspond to recordings per age group. The plain line is obtained using Eq. (2).

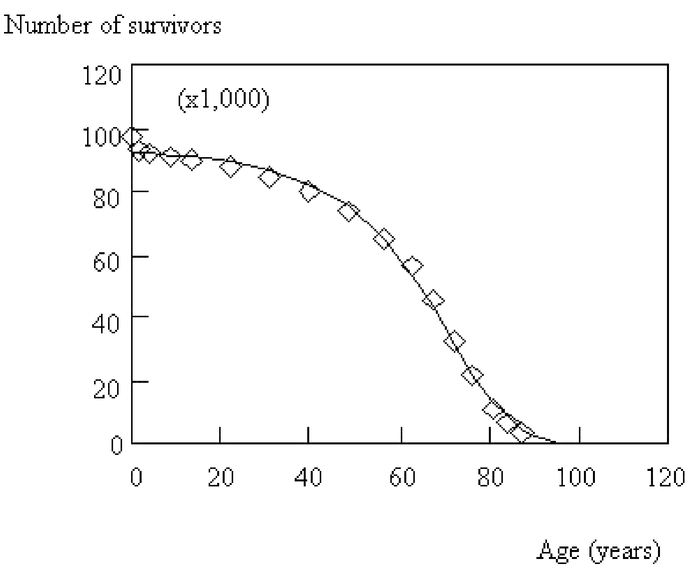


Fig. 8. Survival curve of United States whites for the period 1929-1931 [31]. The rhombs correspond to empirical data while the plain line is obtained using Eq. (4).

in a steady-state. Then (3) the « capacity » starts to diminish first slowly then more and more quickly with the time going.

Two further examples are given below:

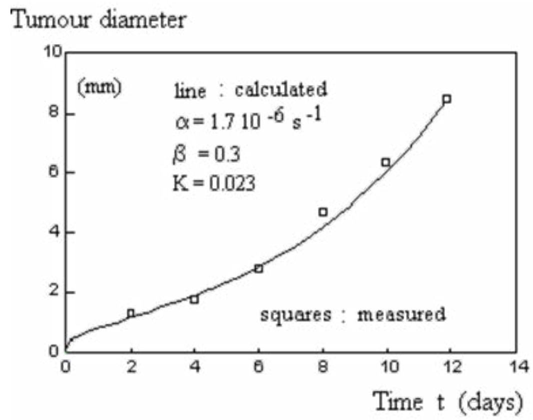


Fig. 9. Growth of Meth-A sarcoma tumours induced in mice [32]. Squares correspond to measured values. The plain line is obtained using Eq. (1) (time in seconds).

1) From the pioneer work of Baulieu [38], the dehydroepiandrosterone sulfate (DHEAS, a kidney produced hormone) concentration in human serum has been considered a good marker for the aging of humans [39] [40]. Fig. 10 gives the DHEAS levels of normal males and females plotted against age [40]. Measurements have been made on 25 children, 32 adult males and 42 adult females. In order to get a better view, only the centres of the rectangles corresponding to the zones of measured values are shown, together with the rectangles. One observes that, after a first increase, the DHEAS content in plasma reaches a peak around an age of twenty five then continuously decreases. The plain line corresponds to the computations using Eq. (5) and c=1.

2) The decline in phytohemagglutinin responsiveness of spleen cells from aged mice was measured by Hori et al. via recording the *in vitro* proliferative capacity of immunocompetent T-cells (by [^{3}H]Tdr incorporation in Δ cpm) [41]. The results are summarized in Fig. 11. Dots correspond to

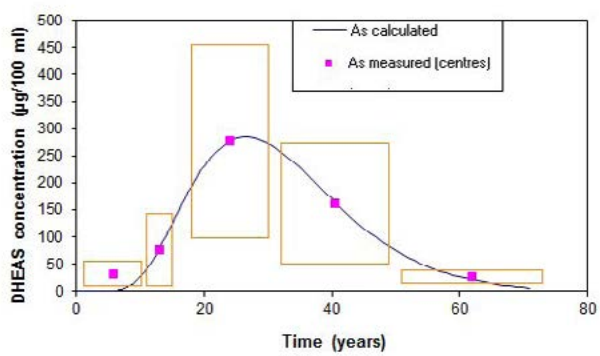
experimental averages as taken from [41]. The plain curve was obtained by fitting the data with Eq. (5) and c=1. The values in ordinate can be converted into Δ cpm by using the conversion:

$$\Delta\, cpm = 10^{\left(\frac{v\,alue}{37.1}+2\right)}.$$

In summary, Æ(t) in biology describes the cumulative aging of a given system, dÆ(t))⁄dt the mortality curve, exp (-Æ(t)) the survival curve and exp ((-dÆ(t))⁄(dt)) the growth and decline curve (of a capacity, a resistance, …).

The loads/stressors in biological systems are the constraints and environmental challenges due to living. For instance, sub-citotoxic stresses can be experimentally imposed on cells *in vitro* to get accelerated aging curves (the « stress » would then be in parameter « k » of Eq. (1) and the influence of the imposed « stress » on the value of « k » can be inferred).

Fig. 10. Dehydroepiandrosterone sulfate (DHEAS) concentration in human serum in function of age.



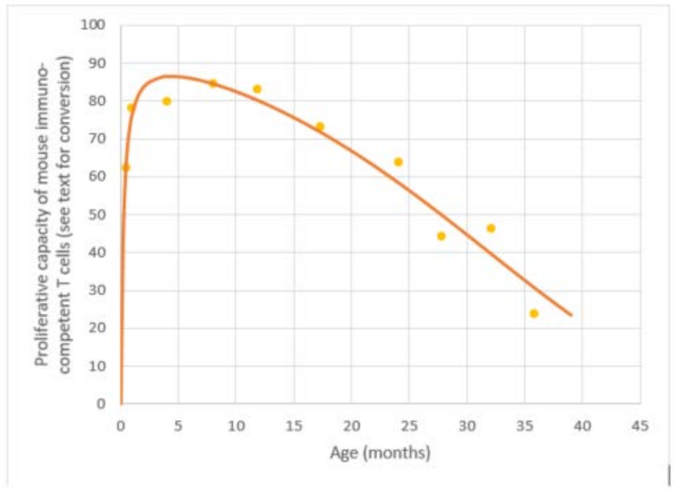


Fig. 11. *In vitro* proliferation of mouse immunocompetent T-cells in function of age.

4. FROM CREEP TO (NEO-)DARWINIAN EVOLUTION

We check Eq. (1) to (3) in the framework of (Neo-) Darwinian evolution.

One should not be astonished that it is spoken of reliability in the context of (neo-)Darwinian evolution, as it is exactly what nature is driving at : that the fittest species for a given environment will thrive, thus will prove a sufficient level of reliability. Evolution is prone to preserve the integrity of the better adapted species (e. g. for eukaryotes : stable DNA to be read many times but protected into a nucleus, double membrane of cells, …). The living entity has an organization that reveals selective advantage in its life environment. There it is armed to reliably survive. Moreover, some reliability is always needed in connection with making evolutionary attempts successful : entities and species would otherwise disappear all too quickly.

Let us now check whether curves of the kind shown in Fig. 1 can be found in connection with evolutionary data. This would suggest that evolution and aging are two faces of the same coin.

In 1975, Nei obtained a rough relationship between the number of nucleotide pairs (bp) in the DNA of different species and their reckoned time of appearance on Earth [42] :

$$n_{bp} = n_0 . e^{\alpha . t} \qquad (6)$$

with $\alpha := 2.3\ 10^{-9}$ years^{-1} for an estimated $3\ 10^9$ years elapsed from bacterium to mammal (n_0 is the initial DNA content, taken here as having the value $n_0 :=$ $4\ 10^6$ bp which corresponds to *E.coli*). Eq. (6) is equivalent to Eq. (1) where β is put β = 0.

Eq. (6) can be re-written in natural logarithms :

$$\ln(n_{bp}) = \ln(n_0) + \alpha . \mathrm{t} \qquad (7)$$

The same can be done for Eq. (1):

$$\ln(Æ(t)) = \ln(k) + \alpha . t + \beta . \ln(\mathrm{t}) \qquad (8)$$

Assuming life to exist on Earth since about $3.5\ 10^9$ years (taken as time zero) and considering the time of appearance of living entities as shown in Table 1, one may compare the results using Eq. (7) (with $\alpha := 2.3$ 10^{-9} years^{-1} and $n_0 := 4\ 10^6$ bp), and Eq. (1) (with $\alpha :=$ $2.3\ 10^{-9}$ years^{-1} , $\beta := 0.45$ and k = 5.755).

Both Eq. (7) and Eq. (8) fit the points as can be seen on Fig. 12 (ordinate in natural logarithm scale).

Case	Type of living entity	n_0 (bp)	After (years):
(1)	*E.coli*	$4\ 10^6$	$2\ 10^8$
(2)	*First unicellulars (e.g. Neurospora)*	$4\ 10^7$	$1.2\ 10^9$
(3)	*Bony fish*	$9\ 10^8$	$2.6\ 10^9$
(4)	*First mammals*	$3.2\ 10^9$	$3.1\ 10^9$

Table 1 – Time of appearance of living entities on Earth and their number of nucleotide pairs (bp) [42]

However, as can be seen from the start of the curves using Eq. (1) and Eq. (7) (see Fig. 13 where the ordinate is no longer in natural logarithm scale), extrapolating back to time zero gives a predicted value of more than $2\ 10^6$ bp (exactly 2 525 134 bp) using Nei's equation. That such a big number of nucleotide pairs would already exist in the genome of all living entities in the very first forms of life on Earth seems unrealistic. On the contrary, using Eq. (1) allows to start from small molecules as has most probably been the case (the computation gives 0 bp at time t :=0 ; 229 bp after 1 year ; etc.)

Whatever the model, the curve should start from a value close to or equal to zero.

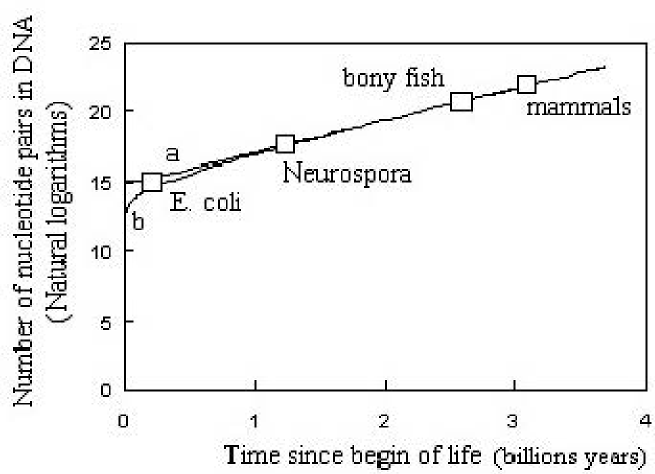


Fig. 12. Number of nucleotide pairs (bp) in genome as a function of time from the start of life on Earth : (a) As calculated using Nei's equation – here Eq. (7) ; (b) As calculated using Eq. (8). Squares correspond to 4 species as selected by Nei [42].

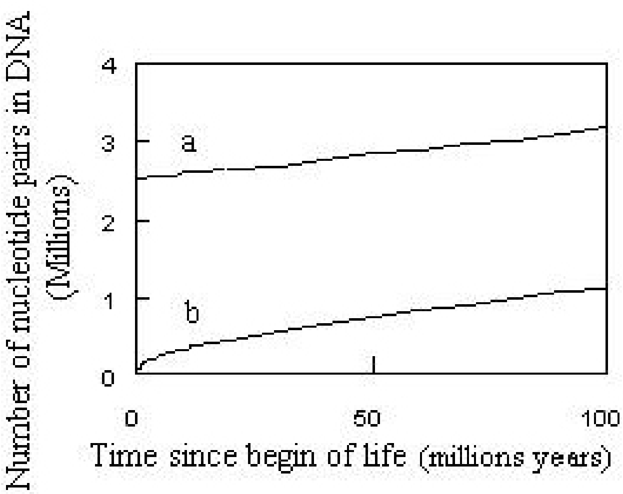


Fig. 13. Start of the curves of Fig. 12 (but with the ordinate no longer in logarithmic scale) : (a) As calculated using Nei's equation – here Eq. (6) - (b) As calculated using Eq. (1).

Using Eq. (1), one better understands why the increase was so quick in the first million years (as noticed by e.g. de Duve [43]) and slowed down afterwards : an initial quick evolution followed by a decrease of the evolution rate is a typical shape for the first stage of evolution curves as given by Eq. (1).

Now, the question rises : if the evolution curve is given by Eq. (1), has it to be smooth ?

Phyletic gradualism goes in the direction of a systematic gradual evolution. This does not necessarily mean that the evolution curve must therefore be smooth at all scales. In reality, in the cases where Eq. (1) applies, the curves should only be smooth at higher time scales as compared to the default duration of the process. When one concentrates on smaller periods of the process, smoothness would disappear. Irregular alternate periods of increase, stabilizing and sometimes even decrease would be seen. This is because adaptation to « stress » implies that there is a time component. Electro-chemical transfers of information, chemical reactions, cell divisions, DNA transcriptions, reactions to unexpected shocks, reactions after learning, …., all these adaptive and genetic actions or reactions need time even if very short: they cannot take place instantaneously. Adaptive efforts appear within the system to allow its further working. This is time (and energy) consuming. There are attempts, initial responses may miscarry or reveal inadequate, there are delays, stabilization is sometimes needed, coordination is not always perfect, … As the system is open, it can have problems of provisioning, getting energy sources, needed stuff is not (immediately) available, stressing agents may suddenly concentrate, … This all makes that the evolution curve shows signs of « hiccup ».

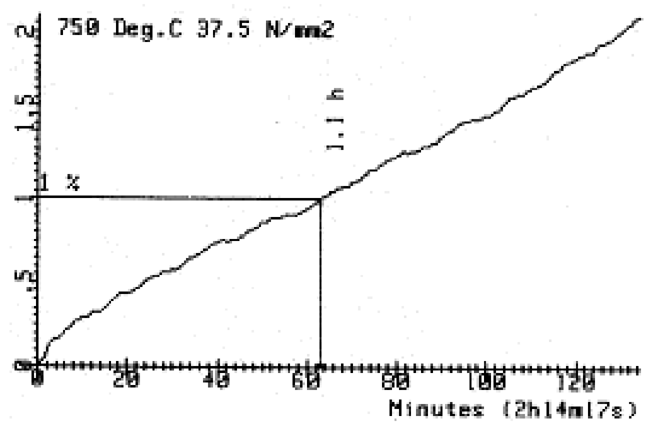


Fig. 14. First part of a creep curve (with permission of Laborelec).

Even in the case of the creep curve of metals (see chapter 1 above), the curve of Fig. 1, is not perfectly smooth when one observes it in details. An example is given in Fig. 14. Actually there are little steps, small periods of stabilization (stagnancy), asymmetries, …

Also concerning the evolution of species, favourable solutions take time to spread to a whole population. Even when occurring quickly (e.g.

resistance of bacteria to antibiotics), a finite time is needed.

In summary, one expects the process of evolution not to develop very smoothly but rather ”step by step” even if at large scales it could be seen as continuous in first approximation. This is schematized in Fig. 15.

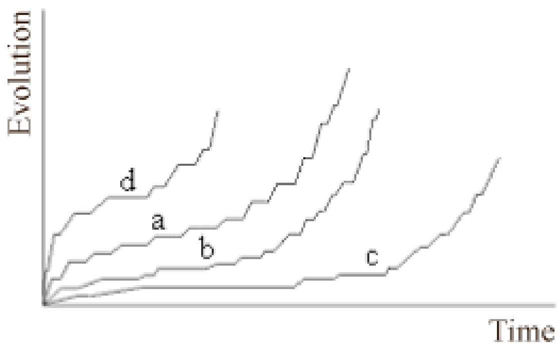


Fig. 15. Examples of « step by step » evolutions (schematic). Curve ”a” corresponds to an evolution at a rather regular pace. Curves ”b” and ”c” describe evolutions which are slow in the beginning and show periods of stagnancy (longer for ”c” than for ”b”) without apparent evolution. Curve ”d” is typical for a quick evolution in the beginning which results in a shorter duration of the process.

As shown in Fig. 16, the actual evolution curve which is observed on Earth for the period starting 600 millions years ago,[6] is rather a curve of the types « b » or « c » in Fig. 15. In spite of five mass extinctions and several more reduced extinctions, the curve of Fig. 16 giving the number of taxonomic families (taken as marker for the evolution of species) in function of time shows an evolution in three stages similar to the curve of Fig. 1. The periods of extinction resulted in « stagnancy » as not the totality of species were extinct then and new species appeared.

The idea of « step by step » evolution as reflected by « hiccups » in the evolution curve is in line with « punctuated equilibrium » [45]. The adaptations are not instantaneous. Evolution works by fits and starts at small scale. There are periods of changes and periods of consolidation, spreading of the gained selective advantages in the population and thriving. Therefore at the right scale, step-like curves would be observed.

Fig. 17 gives an example of punctuated equilibrium as re-drawn from [46] and quoted by [45]. It shows the evolution of the mean thoracic width of the antarctic radiolarian *Pseudocubus vema* over 2.5 millions of years from the advent of the first samples. Similarly to the curves of Fig. 15 and Fig. 16 one observes periods of stagnancy between the periods of increase.

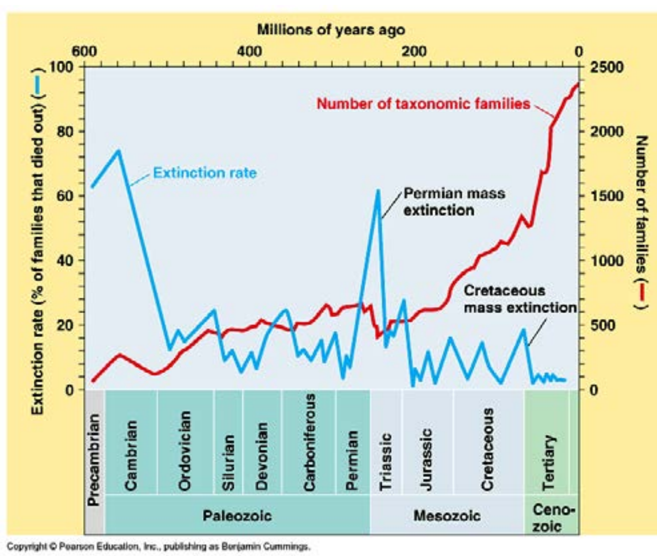


Fig. 16 Extinction rate (blue curve) and evolution of the number of families (red curve) over the last 600 millions of years [44].

6. This covers the Cambrian mass extinction and the Phanerozoic Eon, i.e. the period marking the appearance in the fossil record of abundant, shell-forming and/or trace-making organisms [44].

Finally, it is interesting to draw a parallel with the experimental work of Lenski [47] and Elena [48]. These teams recorded evolutionary changes in bacterial populations (E. coli) propagated for 10 000 generations in identical environments. They first found that the cell volume and the relative fitness (to the ancestor) grew according to a concave (from below) curve [47]. Such curves are similar to the beginning of the curve in Fig. 1. But, focusing on a population at a smaller time scale (3 000 generations), they got a better fit with a step-like model, what they considered to show evidence of punctuated equilibrium [48]. This work has since been continued up to 50 000 generations which is an inestimable source of information on *in vitro* evolution [49][50].

More details on the biological issues related to aging and evolution can be found in [51], [52] and [6].

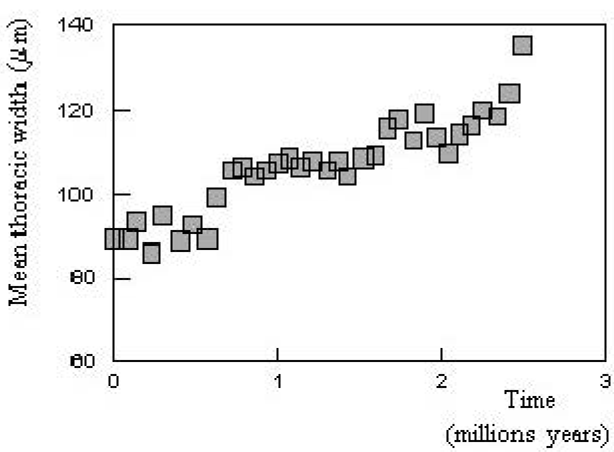


Fig. 17. Example of punctuated equilibrium. Evolution of the mean thoracic width of the antarctic radiolarian *Pseudocubus vema* over 2.5 millions of years from the advent of the first samples (re-drawn from [46] as quoted by Eldredge and Gould [45]).

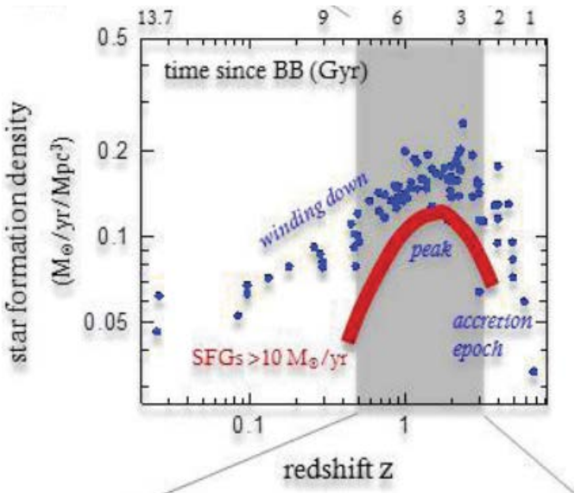
5. APPLICATIONS IN ASTRONOMY AND THE EVOLUTION OF THE UNIVERSE

Stars and galaxies have a life. The life of stars can also be divided into 3 stages (with sub-stages) : (1) a stage of formation of the star ; (2) a stationary stage of true life (where the fuel in the star is used to maintain its life); (3) a final stage after the whole fuel has been burnt, where the star's life goes to end : black hole, neutron star, supernova or white dwarf in function of its mass. The aging of stars is usually analysed with the Hertzsprung – Russell diagram. However it is not possible to extract data from this diagram to find the parameters as described in the present article.

Concerning galaxies, I have found very few information which could become integrated in the present point of view, but it is interesting to notice that galaxies also have a life with infancy, maturity and senescence. For instance, Fig. 18 taken from [53] shows the star formation density of « main sequence » star forming galaxies (> 10 solar masses / yr) in function of time since the Big Bang (BB): see upper abscissa.[7] Here the evolution is to be read from right (Big Bang) to left, corresponding to now (13.7 Gyrs after the BB). The blue points correspond to the observations. They form a global curve which shows the shape of a GD-curve as put into evidence by the schematic thick red line. In fact, reading the cloud of blue points from right to left, i.e. from the BB to present time, we see that it nicely follows a curve of the kind shown for a GD-curve in Fig. 4, 6 or 11 above. In this case, the « capacity » would be that of forming stars. The peak would be around 3 - 6 Gyrs. However, this interpretation is still speculative as the issue should first be investigated on base of the existing models and computations for the formation and evolution of galaxies.

Fig. 18. Star formation density of « main sequence » star forming galaxies in function of time since the Big Bang (BB): to be read from right (BB) to left (now). Taken from [53] page 16.

7. Or the redshift z, which is equivalent (see lower abscissa).



Now concerning the evolution of the Universe, let us first cite the cosmologist who has been at the roots of the BB idea, G. Lemaître : "...If the world has begun with a single quantum, the notions of space and time would altogether fail to have any meaning at the beginning ... Clearly the initial quantum could not conceal in itself the whole course of evolution ... The whole matter of the world must have been present at the beginning, but the story it has to tell may be written step by step."(these words are taken from a seminal text published by G. Lemaître in 1931 in Nature [54]).

As is well known, the classical approach of the evolution of the Universe is based on Einstein's field equations [55] as usually reduced using Robertson–Walker metrics and on the available measurements (I_a supernovae redshifts, cosmic microwave background radiation – CMBR, satellite missions – Cobe …). It follows an hot BB flat Universe scenario. Using Friedmann-Lemaître-Robertson-Walker (FLRW) models, it is thought that the Universe started from an hot BB and expanded into a radiation-dominated era up to the emission of the CMBR (around 380,000 years after the BB), followed by a matter-dominated era of about 13.4 billions years up to present time. A stage of inflation (around 10^{-35} to 10^{-32} s after the BB) is also assumed as it explains empirical data which are otherwise difficult to explain like the homogeneity and the flatness of the Universe. The cosmological constant Λ, initially introduced by Einstein in his equations to allow a non-evolving Universe was firstly put = 0 in front of the evidence of the expansion.

In connection with the present text, it is interesting to focus on the Hubble parameter « *H* » which gives the recession velocity of galaxies divided by their distance. This is usually expressed using the distance scale factor « *R* » by:

$$H = \frac{dR}{R.dt} \qquad (9)$$

Computations give then: H=α=constant (for the inflationary stage), H = 1/2.t (for the radiation-dominated era) and H = 2/3.t (for the matter-dominated era).[8]

This could be summarized in following equation :

$$H = \frac{dR}{R.dt} = \alpha + \frac{\beta}{t} \qquad (10)$$

With β = 1/2 during the radiation-dominated era and β = 2/3 during the matter-dominated era.

This makes sense as integrating Eq. (10) results in a distance scale factor $R(t) \sim e^{\alpha t}\, t^{\beta}$ giving $R(t) \sim e^{\alpha t}$ (typical for an inflationary stage) for β=0 and $R(t) \sim t^{\beta}$ (with β = 1/2 or 2/3 typical for the radiation - and matter-dominated stages respectively) for α=0. Provided that simple assumptions were made for the transitions between stages, the whole evolution curve could then be drawn. This was done in [56] with a reference to the evolution of systems as Eq. (10) is exactly the same as Eq. (2'). In addition, the values of the main cosmological parameters (deceleration factor, pressure / density ratio, Hubble parameter etc.) remained at values as expected in the classical model during the three stages [57].

8. The Hubble parameter is constant in time and space during the inflationary stage, but only in space during the radiation- and matter-dominated eras.

It has to be noted that when « α » is not strictly α=0, Eq. (10) points to the possibility of an accelerated expansion (because of the $e^{\alpha t}$ factor in $R(t) \sim e^{\alpha t} t^{\beta}$).

And indeed, around the end of the years 1990, an unexpected discovery made adaptations to the classical model necessary. The expansion of the Universe is accelerating. It has first been put into evidence by far I_a supernovae redshift measurements [58-60], then confirmed by the WMAP [61] and Planck [62] missions. This also coincided with increasing frustration about the classical model because the calculated density of the Universe on base of visible baryonic matter was an order of magnitude lower than the critical density. In addition, new observations showed the possible presence of big amounts of invisible matter (called « dark matter ») in galaxies. Therefore, the classical model was adapted in the beginning of the years 2000 by interpreting the acceleration of the expansion as due to « dark energy » and reintroducing the cosmological constant Λ as possible cause of it. Adding the densities of visible baryonic matter, dark matter and dark energy allows to reach the level of the critical density. This is called the Λ-CDM model (for Λ – Cold Dark Matter).

As before, the whole evolution curve has been drawn starting from Eq. (10) with $\Lambda = 0$, and compared to the results with the Λ-CDM model where $\Lambda \neq 0$ in the light of recent Planck measurements of H_0 (= H at present time t_0) [63]. The results are given in Fig. 19 (« class » is for « classical » and « alter » when Eq. (10) is used). The corresponding values of the main cosmological parameters (deceleration factor, pressure / density ratio, Hubble parameter, etc.) computed from the BB are reported in [64].

It is noteworthy that when Planck's data (H_0 = *67.8 km/s.MPc* and t_0 = *13.8 Gyrs*) are used in Eq. (10), the deduced value of « α » is:

$$\alpha = H_0 - \frac{\beta}{t_0} = 20.56 \frac{km}{MPc}/s$$

which makes nearly a third of H_0 and is thus certainly not nil.

Fig. 19 shows that the results are very close to each other using the Λ-CDM model and the present approach, at least in the past. In the future, the acceleration of the expansion would be lower with the present approach, but we shall not be there anymore to check it. The computations also show that the acceleration starts 7.58 Gyrs after the BB in the classical model and a little bit sooner in the present approach (7.12 Gyrs) [63 ,64].

The loads/stressors in astronomical systems and in the Universe are first of all the constraints related to gravitation and energy exchanges (but could also be vacuum fluctuations, …).

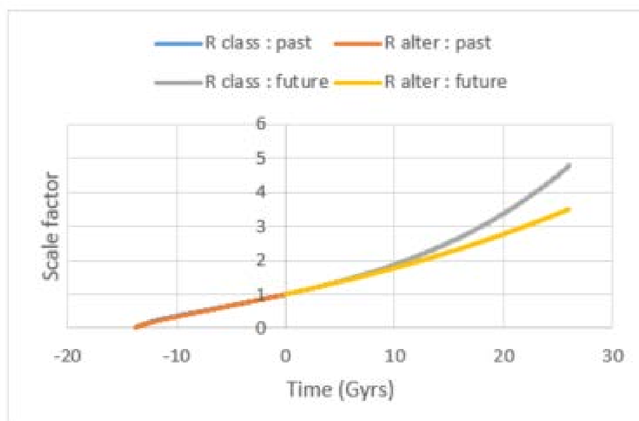


Fig. 19 : Evolution of the distance scale factor R/R_0 from the Big Bang until now and in the future.

6. <u>COMPLEX ADAPTIVE SYSTEMS AND FEEDBACK LOOPS</u>

What is common in all these systems showing similar behaviors although they come from different disciplines : creep of metals, mechanical devices, biology, (neo-)Darwinian evolution, astronomy, cosmology? Answer: they are complex systems adapting to constraints to which they are subjected with time going. The adaptations are the result of many positive and negative feedback loops of several orders which must occur in due time (this may be very short) for the system to maintain its integrity as a complex system. This is possible thanks to the free energy and material available in its environment.

We shall call such systems « complex adaptive systems » (CAS). A definition is given hereafter.

A preliminary remark is that this is not the usual view of complex adaptive systems as referred to in social science etc. where partly independent (human) operators work together or in a network and give rise to an emergent complex behavior [65]. The definition proposed here has the aim to allow for quantitative description of the global (like learning, aging, evolution …) behavior of physical and biological systems as they are encountered in nature. Such systems are also adaptive and complex. However their adaptations are commonly not voluntary but structural, while their complexity results from the high number of interrelated components (as currently considered in systems theory) not as an emergent consequence of relatively simple behaviors of many people (as in complexity theory). Although these two theories have not been duly formalized yet, the publications corresponding to them show different approaches [65]. This does not preclude that the present approach might be useful for groups of human operators too.

One may define a system as a ”*group of elements operating together with a common goal*” [66]. Here, we define a complex adaptive system (« CAS ») as: ”*a group of sub-elements operating together with a common goal and interlinked in such a way that the group shows self-organized criticality*”.

Another equivalent definition is: ”*A complex adaptive system is a collection of numerous interlocked subparts that is in a state of self-organized criticality such that it operates as a whole at an higher scale than that (those) of its subparts*”.

A CAS has to be seen from what we could call ”the point of view of the neutrino”. What the neutrino sees during its endless trip around the Universe are atoms which are sometimes locally more concentrated in a zone sometimes not. However, if it is a good watcher, it will not only see that atoms concentrate in zones but also that some individual atoms seem to be interlinked to each other and not (or to a much looser way, or at much lower frequencies) to other atoms present in their neighborhood.

What are the interlinks ? In most situations considered here, they can be reduced to links between atoms and molecules. This means that the sub-elements interact through the usual chemical bonds : ionic, covalent, metallic, weak bonds (Van der Waals, hydrogen, ...). Interactions may also involve information exchanges, e.g. by transmission of chemical or electric signals, ... The pattern of interlinks makes the organization of the CAS.

Complex adaptive systems can then themselves be interlinked to form bigger complex adaptive systems and so on. The links would then consist in a subtle combination of physics, chemistry and information.

The concept of self-organized criticality (or « SOC ») has been introduced by Bak and co-workers [67,68] on base of observations on sandpiles and computer simulations of cellular automata. According to this concept, many composite systems naturally evolve towards a ”critical” state characterized by four facts :

1) minor events can be at the origin of chain reactions which can affect any number of elements in the system : therefore, ”avalanches” of events of any size can be

produced ;

2) the sizes of the avalanches are distributed following a power law with a negative exponent of the order of ...1...2..;
3) this is true as well in space as in time ;
4) when the system is constrained to produce events, it tends, thanks to the avalanches, to return to a steady state of criticality (therefore, the expression ”self-organized criticality”).

Complex adaptive systems are usually subject to constraints : due to their external environment and/or internal structure, they are not allowed to do everything. They are restricted to evolve in a limited volume, they are subjected to external and/or internal forces or stressors, they have to operate under given conditions of temperature, pressure, humidity, etc. When subject to operating conditions, the subparts and their interlinks steadily reorganize in an adaptative process saving the integrity of the complex adaptive system. Even the Universe can be seen as constrained by internal forces (gravitation) while it expands. This corresponds to the fourth item of SOC.

The adequacy of the adaptive responses will depend on the system’s internal organization and on the raw materials and energy available to carry out solutions in due time.

When one speaks of complex adaptive systems, one expects the reorganizations and adaptive responses to be characterized by several positive and negative feedback loops of any order and level of magnitude and appearing at any time.

The resulting evolution of the CAS can be modelled by the combination of all the feedback loops occurring within the subparts and between them, when time elapses.

Positive feedback loops occur e.g. when, starting from an initial state, non-standard responses are given to the challenges, i.e. responses which do not necessarily re-establish the steady state in all details. The non-standard responses can, for instance, be due to small local errors, because of the limited time available for adaptation. I say ”small” in the sense that the errors do not impair the further working of the system or the subpart : they are integrated in the further operation. But, non-standard responses can also be to the advantage of the CAS. The operating elements must continue to react in due time to the challenges with the consequence that new non-standard responses may be given. There is a cumulative effect (”snowball effect”). This makes that the mathematical expression for a first order positive feedback loop is an increasing exponential [66] :

$$z(t) = z_0 . e^{b.t} \qquad (11)$$

With « z_0 » an initial value and « *b* » the reverse of the time constant for the response.

Positive feedback loops induce exponential growth. They are often found in the CAS under scope in this article.

Negative feedback loops correspond to re-adjustments when some characteristic of the CAS gets out of balance (e.g. because of a challenge, a shortage of material or food component, …). The balance is then re-established (e.g. homeostasis maintained in biological entities).

The mathematical expression for a first order negative feedback loop is given by [66]:

$$y(t) = K . (1 - e^{-a.t}) \qquad (12)$$

With « *K* » a constant and « *a* » the reverse of the time constant for the response.

First order negative feedback loops are often found in the CAS under scope of this article.

Successive first order negative feedback loops will give a concave (from below) curve which can be fitted by a power law (t^{β} with $0 < \beta < 1$).

Higher order feedback loops will result in similar equations except when roots of negative numbers are encountered. This is e.g. the case for a second order negative feedback loop combining a first order negative feedback loop and a first order positive feedback loop as shown on Fig. 20. Then the time constant for the response is given by $\sqrt[2]{(-1/a).(1/b)}$ which means $i.\sqrt[2]{1/(a.b)}$.

The resulting response has then the shape of a wave [66] not of the curve of Eq. (1) or part of it. This is because $e^{i.x} = \cos x + i.\sin x$. Therefore one is not astonished that wavy behaviors are found in complex adaptive systems, e.g. for the TK lymphocyte activation in the immune system or with the tryptophane operon. Also, even if a little bit outside the strict definition of CAS used here, wavy behaviors appear in economy, e.g. in the management of stocks in trade [66]. These wavy behaviors are to be developed in another article (as the present article is first focused on aging / evolution curves).

From the above, we understand that combining many positive and negative feedback loops of any order would globally result in Eq. (1), except when there is a majority of wavy responses. This can be tested by computer simulations.

This is as for several laws in physics (thermodynamics, diffusion equation, …). At microscopic level, the behavior of individual particles, atoms or molecules looks erratic, but when myriads - say 10^{23} - atoms or molecules act together it results in a simple global behavior at macroscopic level because of the great number. Attention was already drawn on this by Schrödinger in his book « What is Life ? » [69].

One can consider that the same occurs at other scales. Cells are like « atoms » at the scale of an organ or body, but the combination of huge numbers of them being interlinked in the frame of a CAS shows simple global aging / evolution. Similarly, orders of 10^{11} galaxies each with 10^{11} stars existing in the Universe will result in a simple global evolution of it.

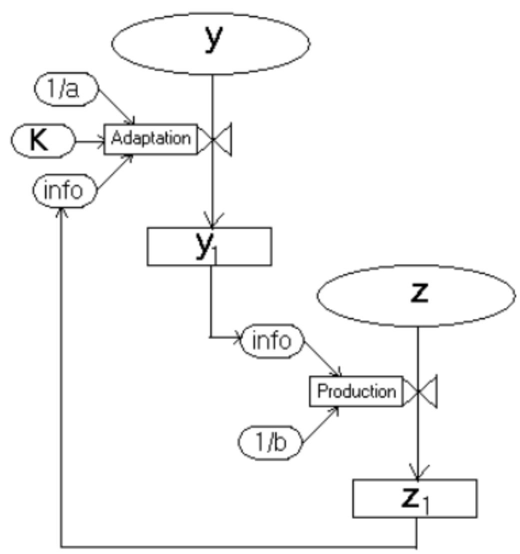


Fig. 20. Schematic view of a second order negative feedback loop obtained from a first order positive feedback loop ($z \rightarrow z_1$)) embedded into a first order negative feedback loop ($y \rightarrow y_1$).

Back to biology, each individual organism is seen as a CAS. Its components and sub-components (molecules, organelles, cells, organs, sub-systems, …) are continuously subject to negative and positive feedback loops of several orders in function of the tasks they have to perform in the frame of the everyday life of the organism.

Depending on the species, a number of tasks and the related feedback loops are genetically encoded: they result from natural selection and have grown into instinct.[9] They respond to the needs of the organism and depend on the available solutions offered by the environment. The system is open as there is a continuous interaction between the organism and its environment. In default situations, the organism is genetically fully adapted to its environment. It has also several internal adaptive features allowing it to cope with external stressing agents, environmental changes, unexpected events, … Finally, thanks to their feedback loops based organization, several species have developed emerging capacities like learning, memory, …

Considering biological entities as complex adaptive systems is consistent with (neo-)Darwinian evolution. Indeed, a first observation is that the process of natural selection itself may be seen as combining positive and negative feedback loops.

For instance, it is known that some mutations in genes may increase the adequacy of the responses to external challenges, others not. Mutations will usually appear at random in genes at a regular pace (or not) depending on the involved gene and environmental triggering factors : radiations, chemical reactions, stressing agents, ... It has in other respects been demonstrated that favourable mutations at a given locus may happen under specific stress conditions (environmental changes) much more often than in neutral conditions [70][71].

Mutations will proliferate in some genes and not or at lower pace in others (positive feedback loops). Then natural selection will favor the biological entities which are most adapted to the external challenges. Those entities whose genes allow to generate better responses (negative feedback loops) will be fitter for the challenges and thrive in their local environment.

9. Not necessarily all tasks.

7. COMPLEX ADAPTIVE SYSTEMS AND ENTROPY

As Æ(t) describes the global aging / evolution behavior under constraints, it can be seen as the statistical emergence of the aging / evolution of the system. As noticed, the statistics is given by a modified Weibull distribution. Æ(t) will reflect the number of microscopic configurations of the system compatible with its macroscopic behavior at each increment of time. In line with Boltzmann's interpretation of entropy [72], we can then define an entropy $S_{Æ}$ (t) characteristic of the aging / evolving system :

$$S_{Æ}(t) = k_B . ln Æ(t) \qquad (13)$$

The « entropy production », i.e. the rate of production of entropy of the system during its life will then be given by :

$$\frac{dS_{Æ}(t)}{dt} = k_B . \frac{1}{Æ(t)} . \frac{dÆ(t)}{dt} \qquad (14)$$

This gives using Eq. (2') :

$$\frac{dS_{Æ}(t)}{dt} = k_B . \left(\alpha + \frac{\beta}{t}\right) \qquad (15)$$

We see that the entropy production diminishes towards a minimum given by $k_B.\alpha$ at time $t = \infty$. This is in line with Prigogine's theorem of diminution of the entropy production for dynamic open systems near to equilibrium [73]. It must be clearly emphasized that for the kind of organized CAS under scope, there is no « negentropy », nor « negative entropy » : it is the rate of increase of entropy which is decreasing not the entropy itself. The CAS takes free energy of high quality from its environment and transforms it for operation, see e.g. [74]. Then it renders the resulting degraded energy back to the environment. This allows diminution of entropy production. The 2nd Principle of thermodynamics is thus holding all the time of the process. Entropy is increasing. But it is increasing at a still lower rate down to a minimum rate where it stays as long as possible. It is thanks to this dynamic process of diminishing the entropy production during life that information is gained [75] and organization is created.

But, after a while, there will be a risk that the entropy production starts increasing again to reach the level which would have been expected if there would not have been any dynamic CAS under scope but only an unorganized collection of the same atoms and molecules instead. The risk of increase of entropy production appears from the time $t_i = 1/\alpha$. This is due to the fact that instability can occur in phenomena described by growing exponentials as soon as the exponent becomes ≥1 ($\alpha.t \geq 1$ in Eq. (1)).

An example is given in Fig. 21 with the aging of humans.[10] If the time t_i from which instability may settle (« instability time »), is 70 years, the life expectancy will lie between 70 (= t_i) and 140 years (= t_m: « the longevity of the species »). With low probability of sudden death at 70 or of life until 140. Most deaths will probably be distributed between 70 and 140 in the shape of a gaussian probability curve. On Fig. 21 are then mentioned : t_r (the « lifespan ») and t_μ (the « life expectancy » for a particular sample).

The instability time gives the order of magnitude of the duration of the aging / evolution process for the CAS considered here, e.g.:

- *~48 Gyrs* for the Universe (if $\alpha = 20.56 \frac{km}{MPc}/s$)
- *500 – 700 Myrs* for the number of taxonomic families of living beings
- *60-70 yrs* for human life

10. This is just an illustrative example to fix ideas. The figures do not reflect actual measurements.

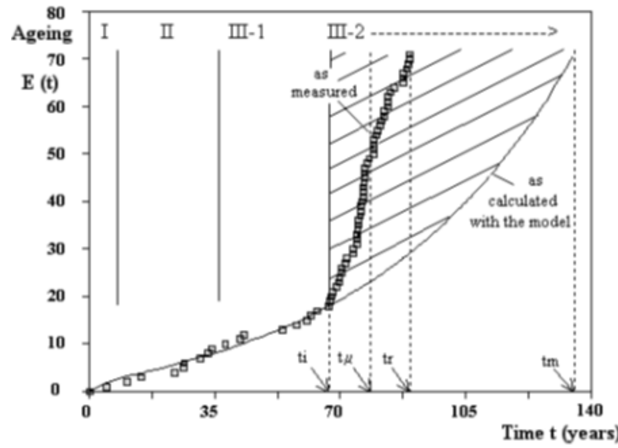


Fig. 21. Illustrative example for human aging after the instability time t_i
(t_m: « longevity », t_r: « lifespan », t_μ: « life expectancy »).

- *25-30 yrs* for the creep of low alloy Cr-Mo steels at 813 K
- *15-20 yrs* for cars
- *10 days* for flies
- *20-80 sec* for the creep of carbon steels at 1473 K
- …..

Probably, instability times for political regimes, societies, civilizations could also be inserted in such kind of table, say with values from a few dozen years to a several hundreds years?

8. CONCLUSION

Many observations point to a same global aging / evolution pattern of complex adaptive systems subject to constraints / stressors. The shapes of the curves are similar for a wide variety of such systems from cells to (Neo-)Darwinian evolution, from the creep of metals to the expansion of the Universe.

The only difference is the order of magnitude of the duration of the phenomenon: from a few seconds for the creep of steels at 1473 K to tenths of Gyrs for the Universe. The curves can be fitted using simple general equations. The recurrent similar behavior at macroscopic level seems to be the result of self-organized patterns of interactions with huge numbers of feedback loops occurring any time at sub-macroscopic and microscopic levels. The shapes of curves referred to in the article and their translations into equations could show useful for screening several behaviors in a lot of fields in the frame of a Big History approach.

REFERENCES:

[1] David Christian, *Maps of Time: An Introduction to Big History,* Berkeley, CA: University of California Press, 2nd ed., 2011

[2] Fred Spier, *Big History and the Future of Humanity,* 2nd ed., Malden, MA: Wiley/Blackwell, 2015

[3] Cynthia Stokes Brown, *Big History: From the Big Bang to the Present,* 2nd ed., New York: New Press, 2012

[4] Andrea Wulf, *The Invention of Nature: The Adventures of Alexander von Humboldt, the Lost Hero of Science*, (London: John Murray, 2015)

[5] H.G. Wells, *Outline of History: Being a Plain History of Life and Mankind*, 3rd ed., (New York: Macmillan,1921)

[6] G. U. Crevecoeur, *Quantitative General Adaptation*. JePublie, Paris, 2008 https://www.numilog.com/47344/Quantitative-general-adaptation.ebook

[7] G.U. Crevecoeur. *An Integrated Method for Remanent Life Assessment under Creep*. In Design and Analysis Methods for Plant Life Assessment, volume 112 of PVP, pages 45–56, New York, (Ed. American Society of Mechanical Engineers ASME. Pressure Vessel and Piping Conference, 1986)

[8] G.U. Crevecoeur. *Über den Einsatz zerstörungsfreier und zerstörender Prüfverfahren zur Abschätzung der Restlebensdauer von Bauteilen unter dem Einfluss von Kriechen*. Der Maschinenschaden, 60 (Heft1):31–40, 1987.

[9] G.U. Crevecoeur. *Translation into Japanese of : Über den Einsatz zerstörungsfreier und zerstörender Prüfverfahren zur Abschätzung der Restlebensdauer von Bauteilen unter dem Einfluss von Kriechen*. Kikai no songai, 33(93):15–24, 1991

[10] G.U. Crevecoeur. *Four parameters databank for full creep curve characterization*. In Life Extension and Assessment: Nuclear and Fossil Power-Plant Components, volume 138 of PVP, pages 203–210, New York, (Ed. American Society of Mechanical Engineers ASME. Pressure Vessel and Piping Conference and also reference NDE 4, 1988)

[11] E.N. da C. Andrade. *On the viscous flow in metals and allied phenomena*. Proc. R. Soc. A, 84:1–12, 1910 [DOI 10.1098/rspa.1910.0050]

[12] G. Smith. *Properties of metals at elevated temperatures*. Mc Graw Hill Book Co, Inc., New York, 3 edition, 1950

[13] F.H. Norton. *The Creep of Steel at High Temperatures*. Ed. Mc Graw-Hill, 1929

[14] Creep data sheet. N3B. National Research Institute for Metals (NRIM), Tokyo, 1986

[15] J.T. Duane. *Learning curve approach*

to reliability monitoring. IEEE Trans. Aerospace, 2(2):563–566, 1964 [DOI 10.1109/TA.1964.4319640]

[16] W.M. Bassin. *Increasing hazard functions and overhaul policy*. In Proc.1969 Ann. Symp. Reliability, 173–178, 1969

[17] R. Bell and R. Miodusky. *Extension of life of US army trucks*. In Proc.1976 Ann. Reliability & Maintainability Symp., pages 200–205, 1976

[18] G.U. Crevecoeur. *A Model for the Integrity Assessment of Ageing Repairable Systems*. IEEE Trans. Reliability, 42(1):148–155, 1993 [DOI 10.1109/24.210287]

[19] G.U. Crevecoeur. *Reliability assessment of ageing operating systems*. Eur. J. Mech. Eng., 39(4):219–228, 1994

[20] Junfeng Liu, Yi Wang. *On Crevecoeur's bathtub-shaped failure rate model*. Computational Statistics and Data Analysis 57:645–660, 2013 [DOI 10.1016/j.csda.2012.08.002]

[21] L. Lee. *Testing adequacy of the Weibull and Log linear rate models for a Poisson process*. Technometrics 22 (2) : 195–199, 1980

[22] C.D. Lai, M. Xie, D.N.P. Murthy. *A modified Weibull distribution*. IEEE Trans. Reliab. 52 (1) : 33–37, 2003 [DOI 10.1109/TR.2002.805788]

[23] C.D. Lai, T. Moore, M. Xie. *The beta integrated model*. In : Proc. Int. Workshop on Reliability Modeling and Analysis—From Theory to Practice, pp. 153–159, 1998

[24] H. Selye. *A syndrome produced by diverse nocuous agents*. Nature, 138:32, 1936

[25] H. Selye. *Studies on adaptation*. Endocrinology,21(2):169–188, 1937

[26] H. Selye. *Stress and the General Adaptation Syndrome*. British Med. J., 1383-1392, 1950

[27] H. Selye. *Stress and aging*. J. Am. Geriatr. Soc., 18:669–680, 1970 [DOI 10.1111/j.1532-5415.1970.tb02813.x]

[28] H. Selye. *Stress Without Distress*. Ed. J.B. Lippincott Co., New York, 1974

[29] Counselling Connection (the Official Blog of the Australian Institute of Professional Counsellors). http://www.counsellingconnection.com/index.php/2009/12/10/the-general-adaptation-syndrome/

[30] Z. Shao, W. Gao, Y. Yao, Y. Zhuo, and J.E. Riggs. *The dynamics of aging and mortality in the People's Republic of China, 1957-1990*. Mech. Ageing and Dev. 67:239–246, 1993 [DOI 10.1016/0047-6374(93)90002-9]

[31] A. Comfort. *Ageing : The Biology of Senescence*. Ed. Elsevier, New York, 1979

[32] T. Fotsis, Y. Zhang, M.S. Pepper, H. Adlercreutz, R. Montesano, P.P. Nawroth, and L. Schweigerer. *The endogenous, oestrogen metabolite 2-methoxyoestradiol inhibits angiogenesis and suppresses tumour growth*. Nature, 368:237–239, 1994 [DOI 10.1038/368237a0]

[33] J.E. Compston. *Osteoporosis, corticosteroids and inflammatory bowel disease*. Aliment. Pharmacol. Ther., 9:237–250, 1995 [DOI 10.1111/j.1365-2036.1995.tb00378.x]

[34] F. Verzar and H. Thoenen. *Die Wirkung von Elektrolyten auf die thermische Kontraktion von Collagenfaden*. Gerontologia, 4:112–119, 1960

[35] R.C. Young, D.L. Borden, and R.E. Rachal. *Aging of the lung : pulmonary disease in the elderly*. Age, 10:138–145, 1987 [DOI 10.1007/BF02432161]

[36] R.R. Willoughby. *Incidental learning*. J. Educ. Psychol., 20:671–682, 1929

[37] T. Matsumura, Z. Zerrudo and L. Hayflick. *Senescent human diploid cells in culture : survival, DNA synthesis and morphology*. J. Gerontol., 34:328–334, 1979 [DOI 10.1093/geronj/34.3.328]

[38] E.E. Baulieu. *Studies of conjugated 17-ketosteroids in a case of adrenal tumor*. J. Clin. Endocrinol. Metab., 22:501, 1962 [DOI 10.1210/jcem-22-5-501]

[39] N. Orentreich, J.L. Brind, R.L. Rizer and J.H. Vogelman. *Age changes and sex differences in serum dehydroepiandrosterone sulfate concentrations throughout adulthood*. J. Clin. Endocrinol. Metab., 59(3):551–554, 1984 [DOI 10.1210/jcem-59-3-551]

[40] M.R. Smith, B.T. Rudd, A. Shirley, P.H.W. Rayner, J.W. Williams, N.M. Duignan and P.V. Bertrand. *A radioimmunoassay for the estimation of serum dehydroepiandrosterone sulphate in normal and pathological sera*. Clin. Chim. Acta, 65:5–13, 1975

[41] Y. Hori, E.H. Perkins and M.K. Halsall. *Decline in phytohemagglutinin responsiveness of spleen cells from aged mice*. Proc. Soc. Exp. Biol. Med., 144:48, 1973 [DOI 10.3181/00379727-144-37524]

[42] M. Nei. *Molecular Population Genetics and Evolution*, volume 40 of Research Monographs, Frontiers of Biology. Ed. North-Holland Publishing Company, Amsterdam, 1975

[43] C. de Duve. *Vital Dust. Life as a Cosmic Imperative*. Ed. Basic Books, a division of HarperCollins Publishers Inc, 1995

[44] Wikipedia article, ref. : https://en.wikipedia.org/wiki/Timeline_of_the_evolutionary_history_of_life#Phanerozoic_Eon

[45] N. Eldredge and S.J. Gould. *Punctuated equilibria : an alternative to phyletic gradualism*. In T.J.M. Schopf, ed., *Models in Paleobiology*. San Francisco: Freeman Cooper, pp. 82-115, 1972. Reprinted in N. Eldredge. *Time frames*. Princeton: Princeton Univ. Press, pp. 193-223, 1985

[46] D.E. Kellogg. *The role of phyletic change in the evolution of Pseudocubus vema (Radiolaria)*. Paleobiology, 1:359–370, 1975 [DOI 10.1017/S0094837300002669]

[47] R.E. Lenski and M. Travisano. *Dynamics of adaptation and diversification : a 10,000-generation experiment with bacterial populations*. Proc. Natl. Acad. Sci. USA, 91:6808–6814, 1994 [DOI 10.1073/pnas.91.15.6808]

[48] S.F. Elena, V.S. Cooper and R.E. Lenski *Punctuated Evolution caused by Selection of Rare Beneficial Mutations*. Science, 272:1802–1804, 1996 [DOI 10.1126/science.272.5269.1802]

[49] Richard Lenski, E. coli Long-term Experimental Evolution Project Site, http://myxo.css.msu.edu/ecoli/index.html, last accessed February 26, 2018.

[50] Richard Lenski, Publications List, http://myxo.css.msu.edu/ecoli/publist.html#pub_2, last accessed February 26, 2018.

[51] G.U. Crevecoeur. *A mechanical system interpretation of the nonlinear kinetics observed in biological ageing*. Bull. Soc. Roy. Sci. Liège, 69(6):311–338, 2000

[52] G.U. Crevecoeur. *A system approach modelling of the three-stage nonlinear kinetics in biological ageing*. Mech. Ageing Dev., 122(3):271–290, February 2001 [DOI 10.1016/S0047-6374(00)00233-5]

[53] Max-Planck-Institut für extraterrestrische Physik – Garching (MPE-Garching). Report 2007 – 2012. Editors and Layout: W. Collmar, B. Niebisch and J. Zanker-Smith, June 2013

[54] G. Lemaître. *The Beginning of the World from the Point of View of Quantum Theory*. Nature, 127:706, 1931

[55] A. Einstein. *Kosmologische Betrachtungen zur Allgemeinen Relativitätstheorie*. Sitz. der Königlichen Preussischen Akademie der Wissenschaften zu Berlin, VI:142, 1917

[56] G.U. Crevecoeur. *A system approach to the Einstein-de Sitter model of expansion of the universe*. Bull. Soc. Roy. Sci. Liège, 66(6):417–433, 1997

[57] G.U. Crevecoeur. *A Simple Procedure to Assess the Time Dependence of Cosmological Parameters*. Phys. Essays, 12(1):115–124, March 1999 [DOI 10.4006/1.3025354]

[58] B. Schwarzschild. *Very Distant Supernovae suggest that the Cosmic Expansion is speeding up*. Physics Today, pages 17–19, June 1998 [DOI 10.1063/1.882267]

[59] A.G. Riess, A.V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P.M. Garnavich, R.L. Gilliland, C.J. Hogan, S. Jha, R.P. Kirschner, B. Leibundgut, M.M. Phillips, D. Reiss, B.P. Schmidt, R.A. Schommer, R. Chris Smith, J. Spyromilio, Ch. Stubbs, N.B. Suntzeff, and J. Tonry. *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*. Astron. J., 116:1009–1038, September 1998 [DOI 10.1086/300499]

[60] S. Perlmutter, G. Aldering, G. Goldhaber, R.A. Knop, P. Nugent, P.G. Castro, S. Deustua, S. Fabbro, A. Goobar, D.E. Groom, I.M. Hook, A.G. Kim, M.Y. Kim, J.C. Lee, N.J. Nunes, R. Pain, C.R. Pennypacker, R. Quimby, C. Lidman, R.S. Ellis, M. Irwin, R.G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B.J. Boyle, A.V. Filippenko, T. Matheson, A.S. Fruchter, N. Panagia, H.J.M. Newberg, and W.J. Couch. *Measurements of Omega and Lambda from 42 High-Redshift Supernovae*. Astrophys. J., 517:565–586, June 1999 [DOI 10.1086/307221]

[61] C.L. Bennett, L. Larson, J.L. Weiland, N. Jarosk et al. *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results.* The Astrophysical Journal Supplement ***208*** *(2): 20*,[arXiv:1212:5225], 2013 [DOI 10.1088/0067-0049/208/2/20]

[62] Planck collaboration. *Planck 2015 results. I. Overview of products and scientific results*. Astronomy & Astrophysics manuscript, [arXiv :1502.01582v2 astro-ph.CO], 10 Aug 2015

[63] G. U. Crevecoeur. *Evolution of the distance scale factor and the Hubble parameter in the light of Planck's results*, [arXiv:1603.06834v2 physics.gen-ph], 15 Feb 2017

[64] G.U. Crevecoeur. *Time evolution of dark energy and other cosmological parameters*. Phys. Essays, 30(3):255–263, September 2017 [DOI 10.4006/0836-1398-30.3.255]

[65] Phelan, S.E. *A Note on the Correspondence Between Complexity and Systems Theory*. Systemic Practice and Action Research, Vol. 12, No. 3, 237-246, 1999 [DOI 10.1023/A:1022495500485]

[66] J.W. Forrester. *Principles of Systems*. Ed. Wright-Allen Press, 1968

[67] P. Bak, C. Tang, and K. Wiesenfeld. *Self-organized criticality : an explanation of 1/f noise*. Phys. Rev. Lett., 59:381–384, 1987 [DOI 10.1103/PhysRevLett.59.381]

[68] P. Bak, C. Tang, and K. Wiesenfeld. *Self-organized criticality*. Phys. Rev. A, 38(1):364–374, 1988 [DOI 10.1103/PhysRevA.38.364]

[69] E. Schrödinger. *What is Life ?* Cambridge University Press, 1967

[70] J. Cairns, J. Overbaugh and S. Miller. *The origin of mutants*. Nature, 335:142–145, 1988 [DOI 10.1038/335142a0]

[71] B.G. Hall. *Spontaneous point mutations that occur more often when advantageous than when neutral*. Genetics, 126:5–16, 1990

[72] L. Boltzmann. *Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo*. Wiedermann Ann., 57:773–784, 1895

[73] I. Prigogine. *Introduction to Thermodynamics of Irreversible Processes*. Ed. Interscience Publishers, New York, 2 edition, 1961

[74] E. Chaisson, *Cosmic Evolution: The Rise of Complexity in Nature,* Cambridge, MA: Harvard University Press, 2001

[75] D. Layzer. *The arrow of time*. Scientific American, 12:56–69, 1975

Rocketing to Energy Sustainability

David J. LePoire

Abstract

Metaphors are an important way to facilitate understanding of new processes. This metaphor is constructed based similarity of a rocket's and civilization's transition from one stable state to another at a higher level accomplished using a limited supply of fuel. For example, fossil fuels enable the transition from sustainable pre-industrial society to another more advanced sustainable society. However, to realize this potential, society must transition to a sustainable energy supply since fossil fuels are dwindling. A major question is whether this global transition can be completed at the same time that global development continues to improve lifestyles and economic opportunities. To help understand some of the complex relationships and challenges in this transition, a metaphor is developed of the evolving technological society as a rocket, which once launched, needs to reach a critical velocity and altitude before obtaining a sustainable orbit. The basis for the metaphor is that there are two stationary locations for the rocket- the ground (pre-Industrial society) and a stable orbit (advanced technological society). The rocket transitions between the two with technology to utilize a finite amount of fuel to overcome gravity and atmospheric friction to attain a speed, altitude, and orientation for a stable orbit.

Correspondence | David J. LePoire, david.lepoire@gmail.com
Citation | LePoire, D. (2018) Rocketing to Energy Sustainability. *Journal of Big History*, II (2); 103 - 114.
DOI | http://dx.doi.org/10.22339/jbh.v2i2.2304

Introduction

Technological civilization depends on large energy flows but its current major energy sources from fossil fuels is dwindling. Much economic progress was made in the 20th century with the energy generated from burning fossil fuels improve the quality of life for many (Smil 1994, Yergin 1991, Yergin 2011). However, to continue on a successful path, civilization must transition to a sustainable energy supply. This transition involves a balance of increasing energy demand (from both increasing population and lifestyles) with the new energy sources and improved efficiency. However, the higher quality of living often results in a decreased (or negative) population growth rate (Korotayev 2015). A major question is whether this global demographic transition can be completed at the same time that the world's energy transitions from fossil fuel to renewables. If energy resources dwindle before the demographic transition is complete, economic foundations may crack.

Metaphors are an important way to facilitate understanding of new processes, such as the approach to sustainability (Alden Trust 2011, Steffan 2007, Larsen 2011, Karlsson 2015). .An evolving technological society can be viewed as a metaphoric rocket (Karlsson, 2015). An important addition to this metaphor is the consideration of the rocket reaching orbit. A rocket, once launched, needs to reach a critical velocity and height before obtaining a sustainable orbit. Once a stable orbit is attained, there are many further beneficial options such as space observations or facilitating further space exploration. The basis for the metaphor is that there are two stationary states for the rocket- the ground and a stable orbit. The ground is analogous to the historical situation of a society based on traditional solar energy for crop growth, warmth, wind, and water. The stable orbit is analogous to an improved situation of an advanced society with more freedom, comforts and fulfillment, which is also stable through technologically capturing a larger fraction of the solar energy (or supplementing it with nuclear fission or fusion).

The process of getting a rocket into orbit requires:

1) Developing the rocket technology including engines, fuel, and crew compartments;
2) Launching the rocket against gravity;
3) Climbing through the atmosphere;
4) Shifting the orientation from vertical to horizontal; and
5) Attaining enough directed speed and altitude to obtain a stable orbit

The metaphor is that society must generate enough technological progress to again generate sustainable energy (from solar or fusion sources). This transition from the sustainable preindustrial agricultural lifestyle to an advanced technological lifestyle is facilitated by the temporary use of exhaustible fossil fuels.

These stages might correspond to historical periods of:

1) Developing: the scientific & industrial revolution;
2) Launching: World War II;
3) Climbing: the Cold War and Oil Crisis;
4) Shifting: the current period of addressing entangled issues of climate change, demographic transition and energy transition; and
5) Attaining: the increased use of renewable energy to sustain the process of global convergence.

Also, it is not clear if society's transition to energy sustainability (the metaphorical stable orbit) will be completed successfully (Ausubel 1999). For example, fuel could run out or a fundamental flaw could disrupt the process.

The paper is organized starting with a brief description of relevant history with a focus on energy use, followed by the metaphorical connections, and finishing with some possible ways, based on this metaphor, to track the progress to sustainability.

From medieval energy sustainability to today

The traditional agricultural society of the Middle Ages in the early 14th century suffered through many problems. People worked hard to harvest enough resources such as wood, food, and metals to support their agricultural lifestyle. The energy from food and wood was supplemented by the energy of domesticated animals, and occasionally from water and wind for mills and transportation. Slavery which had been prominent in the Roman Empire had greatly diminished but still peasants had relatively little freedom to move or innovate in feudal society. Medicine, entertainment, and education were inadequate or ineffective. To add to the burden, fighting would often break out between the feuding feudal lords, who were obligated to protect their serfs.

However, increased interaction between geographically separated societies through trade and conflict opened new possibilities. Trade along the Silk Road increased interaction between Europe and China, leading to exchange of inventions, although many (including paper, gunpowder, and the compass) came from China. Invasions from central Asia, such as Genghis Khan's Mongols, caused more than exchange of ideas and items. Diseases, such as the Black Death in the mid 1300's, spread along with trade and Mongol invasion. This plague triggered changes in the society structure as labor became more valuable.

After the Black Death, questions arose concerning the reason for the tragedy. Society changed via shortages of labor but also in terms of questioning the teaching of the classics such as how the heavens moved. It was common knowledge that the Earth was round and with a good estimate of its size by the Greek Eratosthenes in about 200 BC by observing shadows at different latitudes. Copernicus was motivated to hypothesize that the planets went around the sun. Later measurements by Tycho Brahe led Kepler to analyze the planet data and show it followed a simple mathematical pattern not only in space but also in time. This motivated Galileo to explore the planets even further with the new invention of the telescope which he travelled to Middleburg to buy and then reproduce. With this he was able to see the moons of Jupiter and saw the phases of Venus, clearly showing that these bodies did not circle the Earth but instead that the planets orbited the sun and that many planets had their own moons. Newton then synthesized the nature of gravity on the Earth (falling apple) with the motions in the Heavens, the natural orbits. Others explored different questions. For example, alchemists wondered how one form of matter could be changed to another. While not finding the philosopher's stone to turn lead into gold, they laid the foundations of chemistry.

However, few imagined that basic science information might transform the way humans lived. Earlier innovations, such as Hero's engine in the first century, were not further pursued since such power devices could not effectively replace the relatively inexpensive human labor. By the early 17th century, Francis Bacon advocated skepticism about knowledge not gained through empirical approaches. His empirical observation method would contribute to the development of the scientific method, which also considers the importance of theory and hypotheses.

Already at this time (early 17th century), energy supply was already not sustainable. An energy challenge arose when forests diminished as wood was needed to convert into charcoal to support the higher temperatures to processes increased iron production. Coal was a potential substitute for wood to heat but it was dirtier and could not sustain the conditions for iron working without a special furnace.

The industrial revolution started in England with innovations in the interdependent technologies of iron, railroads, and coal mining. The increasing demand for rails and steam engines resulted in greater demand for iron, which required increased coal mining for operating the engines and processing the raw iron. The more intensive coal mining then required more steam engines to pump the water from the mines and railroads to transport the coal, thereby completing the cycle with an increased demand for iron.

The industrial revolution also brought major challenges to society as factory systems developed, urban areas became more densely populated, and pollution effects were experienced. To open up the potential to a wider group, human rights and principles of democracy and free markets were developed throughout the scientific revolution and industrial revolution. For example, Ben Franklin participated as a leader in both scientific areas (such as electricity, optics, and heat flow) and civil society (with participation as a founding father, diplomacy and setting up social institutions such as fire departments and libraries as well as deriving most of his money through newspaper publishing). However, life at the end of the American Revolution was much the same as both political and scientific principles needed to be expanded and applied.

The U.S. Civil War demonstrated the technology and societal changes from the coal-based industrial revolution. Technologies such as rail, telegraph, guns, and iron ships were tested along with democratic principles against slavery, regional rights, and free market industrialization transitioning from an agricultural system. Further tests occurred during World War I with the introduction of technologies such as air planes, chemical warfare, machine guns, and submarines along with the introduction of new societal processes such as the rights of military conscription and colonization. However, technology exuberance and over speculation led to an economic collapse (the Great Depression). The stresses during the economic downturn allowed groups espousing various solutions such as communism and fascism.

While oil had been known since early civilizations, it had been treated almost as a nuisance. The demand for city lighting after the Civil War encouraged exploration for substitute for candles. Whale oil from sperm whales served this purpose for a while but the supply could not keep up with the demand. Kerosene, a byproduct of oil refining, was identified. However, it could be dangerous if not prepared correctly. John Rockefeller developed a process that led to a standardized kerosene (Standard Oil) that could be safely transported and used.

The oil was supplied from Ohio and Pennsylvania by rail and then by pipeline. Other areas (such as Texas, the Middle East, Romania, the Caucasus, and Indonesia) were explored for oil resources to keep up with the increasing demand. The new machines of industrialization required access to oil so each country vied for control of oil fields, however, the U.S. maintained an early development lead.

A major factor in World War II concerned access to resources. Large feedback loops continued to grow- more technology led to better tools and larger energy demand, leading to new technological solutions. New technology was developed throughout the war including radar, jets, computers, rockets, electronics and nuclear. Pressures to use human resources to a greater potential resulted in expanded participation to women, minorities, and former colonists. Soon after the war's end, a promising technology, nuclear fission power, seemed to offer a seemingly inexhaustible inexpensive energy resource to replace the need for fossil fuel. Some claimed nuclear generated electricity would soon be "too cheap to measure" after implementing breeder reactors that would generate more fuel than used.

The environmental movement hit its peak in developed countries at about 1970 with NEPA in the U.S. as a reaction to oil spills, contaminated waterways, smog, and chemical disposal sites like Love Canal. The green revolution in India was coming to fruition recognized in the 1970 Nobel Peace Prize to Norman Borlaug. In the middle of the 1970's the U.S. hit the halfway point of oil extraction, while in many other places exploration boomed. However, the increased oil demand also enhanced power of other oil producing countries who formed OPEC to control the production and thereby also the price. Two oil price shock waves went through the economies in 1974 and 1979, quadrupling the price of oil. The transition was difficult but increased investment in energy efficient technologies eventually reduced the needed fossil fuels for tasks by about a half (LePoire 2004, 2010a, 2010b). This energy efficiency slowed the demand for oil so that by the mid 1980's the price of oil had decreased enough to hurt oil producing countries like Russia. The economic strain from the oil price drop contributed to the fall of the Soviet Union. But also about this time in the mid 1970's the economy started to shift. In the Real Progress Indicator which starts with the GDP and then subtracts economic activity that is counterproductive and adds the value of outside of market activities. This indicator peaked in the mid-1970s even though the GDP continued its exponential climb (Kubiszewsk 2013) .

- Population explosion with medicine and ag revolution India
- More oil found but U.S. hits halfway in the 70's, economy boomed
- Frictional Cold War saps resources and

hunt for new energy
- Electronics developed and miniaturized
- 70's correction: energy crisis, real GDP and wage stall, computers enter

In the early 1970's it was believed that nuclear power was to replace oil as the primary energy source, but this technology was slowed down due to security, financial and environmental issues. Another form of nuclear energy fusion seems to promise more energy with fewer complicating issues. However, it has been very difficult to control and always seems to be about 30 years in the future (the next generation). While nuclear fission energy is still being explored and new technologies exist to greatly mitigate the problems of the early forms, only Asia seems to be interested in pursuing this. It is unclear is more commitment to safe development would have allowed the developed world to quickly attain sustainability and then transfer it to the developing world. Currently nuclear energy can be viewed as insurance since it is still unclear whether renewable energy sources such as wind and solar would lead to economical sustainability (before fossil fuels are gone).

Control technology (electronics and computers) have greatly improved recently. In the early parts of WWII electromechanically devices were used for simple calculations. These were replaced with vacuum tube technology, followed by semiconductor transistors. These were then integrated to greater extents, following the Moore's law of improvement such that the capability of the integration was doubled almost every 1.5 years. While technology has often been identified as the fourth factor in economic growth with labor, land, and capital, a different analysis suggests that it is energy availability that explains this additional aspect to growth (Ayres 2008).

The rate of technological change may be slowing down. If the rate continued to increase, we would expect technology leaders to maintain their place because they would be able to innovate faster than existing technology diffuses. However, it seems like technologies are expanding throughout the world based on production with low labor costs followed by the advance into a higher quality of life. In addition, technological leaders are appearing outside the original first world, e.g., Korea, China, and India. An effect of diffusion outpacing innovation is seen in the work by Korotayev (2015) who found a correlation between the population growth rate and the gap in economic wealth. The population explosion in the 20th century is beginning to stop and in some areas such as Western Europe reverse. The world still has a long way to go if economic equality is attained. For example, the U.S. uses about 20% of the energy resources for 5% of the global population. This would mean that if everyone used energy at this rate the amount of energy would be 4 times the current demand.

We are in the process of determining whether this energy transition can be done. Learning curves for wind and solar technologies show that the cost of production reduces by a factor for every doubling of production. This cost reduction is driven by improvements (learning) in the production process. The variability of many renewable sources due to night, clouds, calm weather necessitates an ability to temporarily store energy when it is produce. The distributed nature of renewables, e.g., on each house also requires building a smart grid that can monitor the changing environment of production and consumption at each location. Environmental impacts of operation or throughout the life cycle is being monitored to ensure that unintended consequences might be identified and mitigated early.

Metaphor

The main metaphor is that just like a rocket, the society that advances with fossil fuel is initially on an unsustainable path (Figure 1, Table 1). After launching from the safety of the ground (leaving early agricultural sustainability), there is another sustainable situation – the orbit for the rocket and the use of technologically enhanced renewable energy for the society. Both systems, rocket and civilization, have to reach the second stable situation before their limited amount of fuel runs out. There are also many similar issues such as achieving stability, transitioning to new stages (phases), and the passing through the atmospheric friction (GHG release).

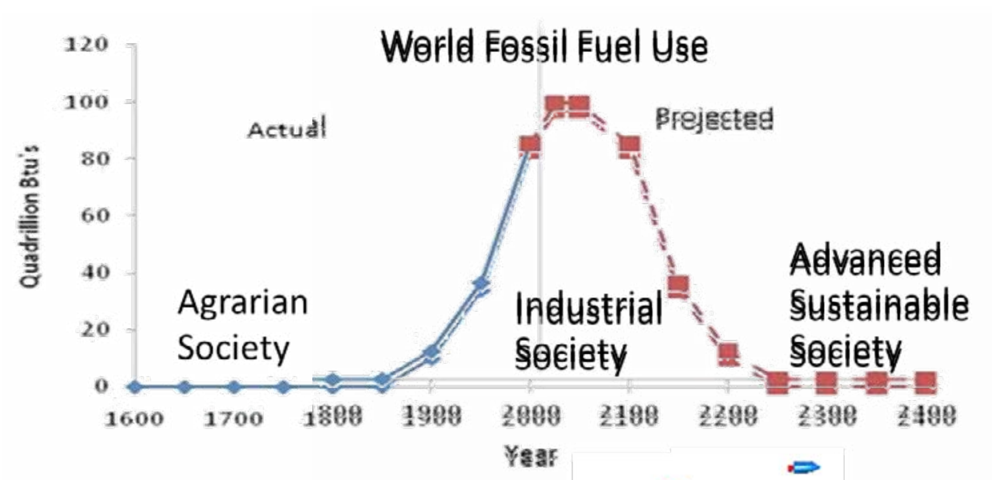
First, the situation of a successful launch into orbit includes aspects of preparation, launching, achieving stability, surviving atmospheric friction, flight planning, and obtaining orbit will be discussed. Each aspect is discussed first with the description of the rocket's process and then the aspect compared to a society transitioning to energy sustainability. Then ways that launches can be unsuccessful will also be discussed for the rocket and society.

<u>Technology Preparation</u>

Early engineers were skeptical of even being able to sustain heavier than air flight by humans. After the Wright brothers' demonstrations of manned air plane flight, others dreamed of designing rockets for space exploration. Unmanned rockets had been used in warfare and for fireworks. Some early pioneers tested principles of stability, control, and flight plans, sometimes failing spectacularly before the principles were understood. An early rocket design by Goddard for, example, had the engine on top of the fuel tanks leading to instability. Similarly, people at the start of the scientific revolution often did not foresee the potential of technology to transform society

<u>Crew compartment Preparation</u>

The counterpart to the Enlightenment is that besides the science, manned human spaceflight needs consideration on the crew compartment and conditions. The construction of the university system to not only pass along knowledge but to test and further science through research is similar to astronaut training to maintain, diagnose problems and fix the rockets components when needed.



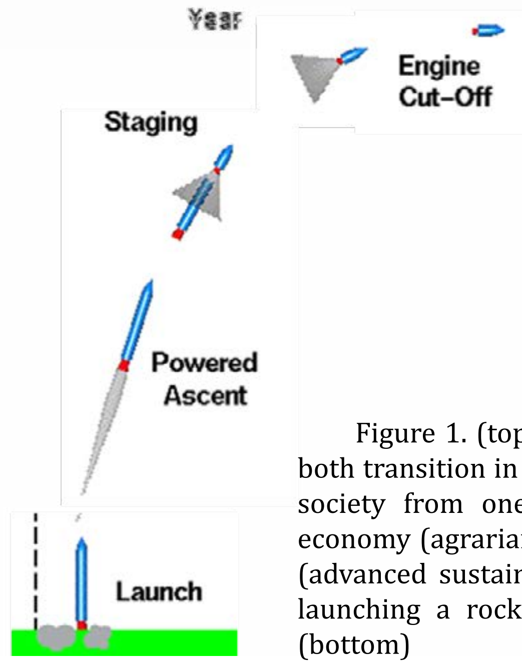


Figure 1. (top) Fuel use in both transition in transitioning society from one sustainable economy (agrarian) to another (advanced sustainable) and in launching a rocket into orbit (bottom)

Lift-off

For liftoff, the rocket engines must generate enough thrust just to overcome gravity. For example, if the rockets generated a thrust equal to that of gravity, the rocket would go nowhere. The thrust needs to be concentrated so that it is larger than gravity to be fuel efficient.

The fossil fuel society achieved "lift-off" during World War II by burning the fuel at a high enough rate to overcome a stagnant economy. The effort put into the innovations during WWII in radar, jets, nuclear, and electronics led to a positive feedback loop- the improved technology generated possibilities of further improvement. At the same time, oil resources were being explored and developed throughout the world.

Stability

Rockets are most vulnerable at lift-off because the speed is low and any small deviations from vertical up can be amplified. This has been observed in many early rocket attempts where the rockets chaotically loop around before crashing. However, once the rocket has gained some speed but still in the atmosphere, there is a natural feedback to vertical stability if the center of mass of the rocket is higher than the center of pressure. The fins at the bottom of the rocket are one way to increase the stability.

The orientation of world development was set by the winners of WWII who still argued over the ideologies of democracy and free market economy versus communism's single party and controlled economy. In a sense two rockets- one from the West and the other from the USSR raced to demonstrate their superiority. Only after about 50 years of testing did it appear that the Western approach with corrective feedback in the government and economy was stable. The rocket's velocity is then similar to the speed of economic growth, and the rocket's stability condition is similar to the automatic feedback mechanisms.

Atmospheric Friction

To avoid overheating from friction with the atmosphere, rockets launch vertically to quickly go through the atmosphere and only then align with the perpendicular orbit. The friction might be considered the counterpart of the potential overheating with greenhouse gases. Soon after the "lift off" during WWII, the release of GHG for a task was much larger. Now that we have learned how to handle energy efficiency, there is less GHG (friction) as we approach orbit.

Flight Plan

But once the lift-off is successful, how should the flight plan proceed? One space flight plan aims directly for the object of interest, e.g., moon or planet. Other plans use Earth orbit as a place to stabilize before moving on to other interesting locations. Then the transition to the flight to the moon is achieved by timing the firing of smaller engines to use the orbital speed to its advantage.

The direct flight plan is similar to those who want to realize ever increasing technological change (Kurzweil 2006) without first achieving energy sustainability (i.e., all fuel is used in one dash without achieving stability first). The second flight plan is similar to having a goal to reach energy sustainability but realizing once there, it is easier to solve other problems and move on to other interesting places. This allows some decoupling of the issues of technological change and energy sustainability.

Multistage transitions

With rockets, there are often multiple stages each with their own fuel supply and rocket engines. The first stage if often the largest since it has to lift the whole rocket off the ground. When the fuel is exhausted, the engine and the empty fuel tank are discarded.

In the metaphor, the stages are the different levels of technology. It is know that technology has proceeded in cycles such as the Kondratieff cycle of approximately 70 years. The early cycles of this were the railroad, coal and telegraph. Later oil was used with the internal combustion engine along with radio. The latest still uses fossil fuels but with much better efficiency and improved communicated through the internet.

Differences

It should be noticed that there are important differences in the transitions of society and an actual rocket. Rockets are often multi-stage meaning that the lower parts of the rocket are jettisoned after the fuel has been exhausted. This reduces the mass of the rocket to ease the work necessary to achieve the orbital velocity. However, in contrast the transition to sustainability seems to occur as more of the world's population joins a developed lifestyle with its increased demand for energy. In the metaphor, this could be viewed as adding mass to the rocket as it climbs, which then requires more work to obtain a stable orbit.

Possible Failure Mechanisms

It is not clear if society's transition to energy sustainability (the metaphorical stable orbit) will be completed successfully. In this metaphor, it is not at all clear which plan we should follow towards sustainability since we really do not know the fundamentals that any rocket engineer would know. Such information would include the weight of the rocket, the efficiency of the engines, the amount of fuel, the speed necessary to get into orbit, and the height of the orbit such that the atmosphere is negligible.

A rocket launch can crashing from loss of stability, fuel tank explosion, too slow acceleration leading to inefficient use of fuel, too much acceleration

Category	Rocket	Civilization
Stable States	Ground /Orbit	Medieval Sustainable/ Advanced sustainability
Fuel Limit	Tanks	Fossil Fuels
Self-Correcting Orientation	Center of mass above center of pressure	Free market / democracy
Concentrated Thrust	Big rockets with acceleration above 1 g	Technology feedback / acceleration
Withstand atmospheric frictional heating	Heat resistant material with proper course	Avoid large climate change
Attain orbital speed	Multistage rockets	Renewable energy with multistage of technologies
Possible Outcomes	Crash, run out of fuel, overshoot, orbit	Crash, run out of fuel, technology runaway, sustainability
Benefits of Orbit	Satellites, solar power, place to easily continue	Advanced lifestyle, ability to continue

Table 1. Comparison of various aspects in the metaphor of a rocket launched into orbit and society transitioning to energy sustainability.

damaging engines,. The rocket might also heat up too much when going through the atmosphere or if the orbit is too low. The rocket might not orient correctly for a stable orbit. Another failure would be for the rocket to enter a stable orbit but lose the capability to support humans, e.g., buildup of carbon dioxide as started on the ill-fated Apollo 13.

There are clear similarities in the transition to sustainability:

Loss of stability: A Rocket maintains stability through self-correcting as long as the stability condition is met that the center of mass is atop of the center of pressure. Similarly, if processes are not in place for self-corrective actions, progress towards sustainability could be lost or reversed. Such processes are fair markets in a democracy but with regulations to ensure market failure mechanisms are addressed. An example of a market failure might be the pricing of fossil fuel without including external costs such as pollution or defense to protect the international trade. The incorrect price signal would tend to encourage the movement towards the cheaper "subsidized" fossil fuel over renewables. This would be similar to the rocket losing its stability to continue in a vertical direction.

Fuel tank explosion: A rocket's fuel tank might leak leading to a catastrophic failure. While the oil underground could not immediately explode, there are scenarios where much of the energy might be lost, for example through resource wars at an expanded scale compared with WWII. All the oil would not have to be rendered inaccessible because the oil trade can be very sensitive to decreases in supply. The reduction in oil supply was seen during the Iran-Iraq war, the Iran embargo, and the burning of the oil fields during the Iraq war.

Acceleration too slow: A rocket launch requires a concentrated use of fuel just as the transition to sustainable energy relies on support for continuous technology innovations. If no technology advances are made, the society will not approach sustainability but instead just burn the fossil fuel until it is gone.

Acceleration too fast: The similarity to too much acceleration happens if the technological change goes too fast. This could strain the developing infrastructure. For example, if cybersecurity does not keep up, large damage might be done to the use of the smart grid. Other potential pathways include advanced technologies get in the hands of the wrong people leading to social instability. This high technological acceleration could result if the technology singularity happens. Safeguards are being proposed to mitigate these unintended consequences.

Atmospheric heating: A rocket experiences friction with the air in the atmosphere as it climbs and also in orbit if the height of the orbit is not great enough. This is similar to the transition to sustainable energy as we use fossil fuels that is generating potential GHG that leads to climate change. This suggests that it is important to reach the energy sustainability before the irreversible harmful effects occur, and that the sustainability is complete enough that fossil fuels aren't needed to fill an energy gap.

Incorrect Orientation: The orientation of a rocket needs to be change from the vertical takeoff to the horizontal when it reaches orbit. In orbit, the forces of gravity are just enough to balance the natural tendency to go straight (centripetal force). This is similar to the need to rethink the energy generation, distribution, transmission, and pricing to guide a sustainable energy transition so that the trickier balance of supply and demand satisfied.

A problem would occur if the variable solar and wind energy replaced all the electrical energy at present. During windy sunny days the lines would be overloaded and the energy not used. During quiet nights the lack of energy supply might cause brownouts. This balance can be achieved with proper storage and price incentives.

Loss of life support: Many problems in the Rocket's capsule might lead to an unsuccessful mission even if the rocket attains orbit. In the transition to sustain energy it is important to keep social and international tensions at a minimum despite the added uncertainty caused by the change. An example would be the potential conflict over water resources. Energy can be used to generate clean water from alt or polluted water through evaporation and efficient recovery of the energy after condensation. Another failure might be that technology is not shared adequately among nations.

Dashboard

To have a successful launch into orbit, many engineers, controllers, and astronauts work together to design and then monitor progress so that changes can be made to correct problems. Currently we do not have a set of equivalent monitors or controls. In many cases we really don't know the characteristics of the systems, e.g., the amount of fossil fuel, the Co2 level that causes irreversible harm, the number of people that need the energy, the level of energy efficiency that is attainable, the economic viability of renewable energy sources.

The transition to sustainable energy seems to have started (lifted off) with some self-correcting systems like market economy and democracy. It also had undergone a few changes in technologies (stage separations), for example dramatically increase energy efficiency. But only now do the possible criteria for reaching a sustainability such as total energy demands and technologies (stable orbit parameters such as height and speed).

This suggests some indicators for the dashboard:

A measure of the amount of fossil fuel remaining, similar to the rocket's fuel gauge.

A measure of strength of the self-correcting stability system in economic and political stable (rockets stability condition).

Comparison of GHG emissions compared to what the global system can sustain. (Comparison of the temperature to what the rocket can sustain.)

A measure of the global social stability and political will. (The cabin environment).

A measure of the projected demand for energy as the population increases and the lifestyle improves for many. (The mass of the payload.)

What level of renewable technology is required to supply the demand. (The required height of the rocket to escape frictional forces.)

How to encourage the implementation of the technology to generate the required demand. (The speed and reorientation of the rocket to achieve orbit.)

Some indicators have been defined and are beginning to be tracked (Henderson 2012 , U.S. Interagency Working Group on Sustainable Development Indicators 1998 , Ness 2007) including some specifically designed to help assess new potential energy technologies (Evans 2009). Others deal with environmental sustainability. For example, environmental impacts include those that affect air quality, food, water, disease, and land use. An attempt to outline basic environmental measures was made in 2009, with the nine planetary environmental boundaries which include natural resource use (land and water), atmospheric disturbances (ozone, climate change, and aerosols), and releases impacting biological activity and diversity (ocean acidity, chemical pollution, phosphorus, and nitrogen release). Estimates were made for each boundary's natural level, the current level, and the level at which impacts might rapidly rise. Some of these already have impacts much greater than their boundary action level, while a few others have greater uncertainty with undefined boundaries (Rockstrom et. al. 2009).

Conclusion

A metaphor was constructed for civilization to reach energy sustainability by comparing it with the process of getting a rocket into orbit. The basic metaphor was that the two systems transition from one stable state (subsistence farming and ground) to an elevated dynamic stable situation (enhanced technology energy sustainability and orbit). This transition is done with a finite amount of fuel.

The various phases were analyzed and compared including planning, lift-off, stability, heating, speed, and orientation. Potential negative outcomes were identified such as not enough fuel, accelerating too quickly or too slowly, and being conscious of maintaining a human livable environment. A potential dashboard of current indications was abstractly developed based on this metaphor but its realization would require much further investigation.

This metaphor might be a way to communicate the complex situation in approaching a sustainable state. The consequences can be visualized and the dashboard would provide some feedback on the critical criteria that need to be satisfied for successful attainment of sustainability. Just like the rocket in orbit however, the attainment of the second stable state is not the end of the adventure but instead a temporary transitioning point to further exploration and advances.

References

Ashden Trust, By Another Name: New metaphors for Sustainability, www.ashdendirectory.org.uk/featuresView.asp?pageIdentifier=2011410_28527468

Ausubel, J. H. (1999). "Because the human brain does not change, technology must (Reprinted from Production Efficiencies: The Engineer's April 1999 Report)." IEEE Aerospace and Electronic Systems Magazine 14(10): 3-6. https://phe.rockefeller.edu/BrainNotChange/

Ayres, R.U. (2008). Energy and economic growth. in Sustainable Energy Production and Consumption: Benefits, Strategies and Environmental Costing, F. Barbir & S. Ulgiati editors, Springer, ISBN: 978-1402084935

Evans, A., Vladimir Strezov, Tim J. Evans (2009) Assessment of sustainability indicators for renewable energy technologies. Renewable and Sustainable Energy Reviews. 13.

Henderson, H. T.J. Nash, R. Sanquiche (2012), Green Transitions Scoreboard, Ethical Markets, available at: http://www.ethicalmarkets.com/wp-content/uploads/2012/02/2012-GTS- February-Report-FINAL-032012.pdf

Karlsson, R. (2015) Three metaphors for sustainability in the Anthropocene,The Anthropocene Review 3(1) · August. 23-32. . DOI: 10.1177/2053019615599415 Available at: http://anr.sagepub.com/content/3/1/23

Korotayev, Andrey, Jack A. Goldstone, Julia Zinkina, (2015). "Phases of global demographic transition correlate with phases of the Great Divergence and Great Convergence", Technological Forecasting & Social Change 95 (2015) 163–169. DOI: 10.1016/j.techfore.2015.01.017

Kubiszewski, I., Robert Costanza, Carol Franco, Philip Lawn, John Talberth, Tim Jackson, and Camille Aylmer (2013). Beyond GDP: Measuring and achieving global Genuine Progress. Ecological Economics, 2013. 93: p. 57-68. DOI: 10.1016/j.ecolecon.2013.04.019

Kurzweil, R. (2006), The Singularity is Near, Penguin Books. ISBN: 978-0739466261

Larson, Brendon (2011). “Metaphors for Environmental Sustainability” New Haven, Connecticut: Yale University Press, ISBN: 9780300205817.

LePoire, D. J. (2004). “A ‘Perfect Storm’ of Social and Technological Transitions?” Futures Research Quarterly 20(3).

LePoire, D. (2010a). “Long-term Population, Productivity, and Energy Use Trends in the Sequence of Leading Capitalist Nations.” Technol. Forecast. Soc. Change. DOI: 10.1016/j.techfore.2010.05.003

LePoire, D. J. (2010b). Threading the Environmental Needle: Applying New Tools to Reduce Uncertainty in Environmental Foresight. Strategies and Technologies for a Sustainable Future,. C. G. Wagner. Bethesda MD, World Future Society. ISBN: 978-0-930242-67-1

Ness, B., Evelin Urbel-Piirsalu, Stefan Anderberg, Lennart Olsson (2007) Categorising tools for sustainability assessment. Ecological Economics, 60: p. 498-508. DOI: 10.1016/j.ecolecon.2006.07.023

Rockstrom, J., et. al. (2009), Planetary Boundaries: Exploring the Safe Operating Space for Humanity Ecology and Society,14(2).

Smil, V. (1994). Energy in world history. Boulder, Westview Press. ISBN: 978-0813319025

Steffen, Will, Paul J. Crutzen, and John R, McNeill (2007). The Anthropocene: Are Humans Now Overwhelming the Great Forces of Nature? Ambio; Dec; 36, 8; pg 614

Profile of John Mcneill, Georgetown University School of Foreign Service in Qatar · Department of History, www.researchgate.net/profile/John_Mcneill4

U.S. Interagency Working Group on Sustainable Development Indicators (1998), Sustainable Development in the United States: An Experimental Set of Indicators. Available at: http://teclim.ufba.br/jsf/indicadores/SDI%20US%20SUST%20DEVEL%20INDICAT.PDF

Yergin, Daniel H. (1991). “The Prize: The Epic Quest for Oil, Money, and Power, Simon and Schuster. ISBN: 978-1439110126.

Yergin, Daniel H. (2011). The Quest: Energy, Security, and the Remaking of the Modern World, Penguin Press. ISBN: 978-0143121947

Evolução do Sistema Solar Primordial em Termos de Grande História e Evolução Universal

Leonid Grinin
Centro Eurasiático de Grande História e Previsão de Sistemas,
Instituto Oriental, Academia Russa de Ciências

Traduzido por Luis Brudna and Daniel de Pinho Barreiros

Resumo

A presente contribuição é dedicada a alguns aspectos da história e evolução do Sistema Solar primordial. A origem do Sol, da Terra, de outros planetas e dos seus satélites tem sido de grande interesse para as pessoas. Ao longo das últimas décadas, astrônomos e cosmólogos avançaram consideravelmente na percepção da estrutura, história e evolução do Sistema Solar. No entanto, dificilmente seríamos capazes de estabelecer propriamente uma narrativa; mais frequentemente trabalhamos com hipóteses. O presente artigo está estruturado da seguinte forma. Primeiro, descreve a história da formação do Sistema Solar nos primeiros cem milhões de anos de sua existência, quando as mudanças mais consideráveis ocorreram. Então, ao descrever certos processos formativos, mostramos as oportunidades para defini-los em termos de leis e regras evolutivas. Claro, este artigo apresenta apenas algumas dessas leis e regras. Acreditamos que o presente estudo seja de interesse para um leitor de duas maneiras. Em primeiro lugar, existem algumas pesquisas breves e consistentes sobre a história do Sistema Solar que contabilizam os últimos avanços em astrofísica e cosmologia. Além disso, eles são muito importantes e fecundos para teorizar parte da Grande História. Em segundo lugar, a discussão que emprega as leis e regras evolutivas gerais permite definir algumas características comuns na formação do Sistema Solar e especialmente do seu sistema planetário, que são características para cada nível e estágio da Grande História. Isso nos leva à ideia da integridade da Grande História não só em termos históricos e sistêmicos, mas também em relação à sua integridade na detecção de leis, padrões e mecanismos gerais.

Palavras - chave

sistema solar, exoplanetas, protonebulosa, sub-disco de pó, planetesimais, embriões de planeta, protoplanetas, catástrofes, migração planetária, regras e leis de evolução, desencadeamento, luta por recursos, sistemas primários.

Correspondence | Leonid Grinin, Леонид Гринин <leonid.grinin@gmail.com>
Citation | Grinin, L.. (2018) Evolução do Sistema Solar Primordial em Etermos de Grande História e Evolução Universal. Traduzido por Luis Brudna <luisbrudna@gmail.com> and Daniel de Pinho Barreiros *Journal of Big History*, II(2); 115 - 127.
DOI | http://dx.doi.org/10.22339/jbh.v2i2.2305.g2175

Introdução

A presente contribuição é dedicada a alguns aspectos da história e evolução do Sistema Solar primordial, isto é, ao primeiro bilhão de anos de existência. Este período é crucial para entender como e por que o Sistema Solar se tornou o que conhecemos. Deve-se apontar que existem muitas hipóteses e teorias sobre a formação de planetas do Sistema Solar. Ainda nenhuma delas pode explicar toda a gama de problemas relacionados.

Este artigo é uma continuação do meu trabalho anterior (Grinin 2014) tanto no sentido do período abordado quanto em termos metodológicos. No meu artigo anterior (Ibid.) ~~eu~~ considerei os principais eventos da fase cósmica da Grande História a partir da Era Estrela-Galáxia, que descrevi em termos de princípios evolutivos universais. No presente trabalho, investigarei a evolução e a história do Sistema Solar primordial e, nesse contexto, vou mostrar a possibilidade de definir uma série de eventos desta história em termos de leis e regras evolutivas.

Esta abordagem foi deliberadamente escolhida, uma vez que permite amplificar a metodologia da Grande História com as realizações e os princípios da *Evolucionística*. Como escrevi em outro lugar, embora a Grande História ofereça oportunidades únicas para considerar o desenvolvimento do Universo como um processo singular, deve-se apontar que os estudos da Grande História tendem a atentar pouco para um aspecto tão importante como a unidade dos princípios, leis e mecanismos de evolução em todos os seus níveis. Eu acredito que combinar o potencial da Grande História com abordagens evolutivas pode abrir horizontes mais amplos a este respeito. Na verdade, os traços comuns em desenvolvimento, funcionamento e interação podem ser encontrados em diferentes processos e fenômenos dentro da Grande História. A este respeito, o caráter universal da evolução é expresso nas semelhanças objetivas que são detectadas em muitas manifestações em todos os seus níveis. Tal abordagem abre novas perspectivas para a nossa compreensão da evolução e da Grande História com suas forças motrizes, vetores e tendências e para criar um campo consolidado para uma pesquisa multidisciplinar (Grinin 2014, 163-164). Esta abordagem também produz um efeito sinérgico que revela novos aspectos do nosso Universo e da integridade do mundo.

Do ponto de vista evolutivo, eu dividi a história primordial do Sistema Solar em quatro grandes épocas.

A primeira época foi a formação do disco protosolar e protoplanetário a partir da nebulosa solar. Esta foi a época da formação da «ordem vinda do caos» nos termos de Prigogine e Stengers (1989) durando cerca de um milhão de anos após o colapso da nuvem protosolar.

A segunda época foi a formação de matéria sólida, embriões de planetas e planetas primários. Pode ser denotada como a época da luta pelos recursos - durando cerca de 10 a 50 milhões de anos após o colapso.

A terceira época pode ser chamada de época de migrações planetárias e catástrofes - durando cerca de 600 a 700 milhões de anos, cerca de 3,9 a 3,8 bilhões de anos atrás.

Finalmente, a arquitetura atual do Sistema Solar foi estabelecida.

A quarta época é chamada *Intenso Bombardeio Tardio* dos planetas e seus satélites por planetesimais e meteoritos que durou de 900 milhões a 3,2 bilhões de anos.

Conforme mencionado acima, a pesquisa sobre a evolução do Sistema Solar permitiu revelar um número significativo de processos e eventos que podem ser descritos em termos de leis e regras gerais de evolução. Na presente contribuição

eu tento mostrar que existem muitas semelhanças e características comuns manifestadas nos mais diferentes processos e fenômenos em vários estágios e níveis da Grande História.

Suponho que essa contribuição seja de interesse para os leitores de duas maneiras. Primeiro, existem algumas pesquisas consistentes e ainda breves sobre a história do Sistema Solar que contabilizam as últimas conquistas em astrofísica e cosmologia. No entanto, elas são muito importantes e produtivas para teorizar parte da Grande História. Em segundo lugar, a discussão que acompanha as leis e regras evolutivas gerais nos permite revelar alguns padrões na formação do Sistema Solar e especialmente do sistema planetário que são comuns para diferentes níveis e estágios da Grande História. Isso nos leva à ideia da integridade da Grande História não só em termos históricos e sistêmicos, mas também em relação à sua integridade na detecção das leis gerais, padrões e mecanismos.

Devido ao escopo do trabalho, escolhi as regras, leis e padrões evolutivos relacionados apenas a alguns eventos (e não todos) da história inicial do

Sistema Solar (para detalhes, veja Grinin 2017a).

1. A formação do sistema protosolar a partir de uma nuvem de gás

Com respeito à história do Sistema Solar, ainda há mais hipóteses do que fatos comprovados. No entanto, ano após ano, as hipóteses relativas a certos fenômenos são corroboradas por observações diretas, por exemplo, como resultado da descoberta de numerosos exoplanetas.

A idade do Sistema Solar, determinada com a técnica de datação radioativa no estudo dos meteoritos mais antigos, é de aproximadamente 4,57 bilhões de anos (Shukolyukov e Lugmair 2003; Vityazev e Pecgernikova 2010, 168; Pfalzner et al. 2015). As principais características do sistema foram formadas durante os primeiros cem milhões de anos, mas a narrativa atual desse período ainda é extremamente fragmentada e pouco confiável.

Ao longo das últimas duas a três décadas, foi elaborado um chamado cenário padrão para a formação de um sistema planetário a partir de um disco protoplanetário de pó e gás que envolve uma protoestrela, o que permite definir os contornos gerais do processo.

Apoiado por numerosas observações diretas, o modelo do nascimento das estrelas geralmente é usado para reconstruir a origem do protosol. As estrelas geralmente são formadas nas partes mais densas de nuvens de poeira de gás molecular, esta última composta principalmente de hidrogênio e hélio e com uma temperatura próxima do zero absoluto. As nuvens de gás podem preservar o equilíbrio por muitos milhões de anos. É necessário um certo impulso (um gatilho) para iniciar o processo de condensação (e subsequente colapso). Talvez, para o nascimento do Sol, tal gatilho tenha sido a onda de choque de uma supernova próxima cerca de dois milhões de anos antes do início do colapso (Adushkin et al. 2008, 276).

Aqui lidamos com uma **regra evolutiva geral** que eu defini como uma *regra de fenômenos ou eventos desencadeantes necessários para iniciar o processo evolutivo*. Por um lado, dificilmente pode funcionar sem a prontidão interna de um sistema; e, por outro lado, mesmo uma prontidão interna de alto nível em si dificilmente pode garantir o início de uma transformação, assim como a pólvora não pode ser explodida sem fogo. Sem um gatilho, um sistema pode por muito tempo permanecer potencialmente pronto para transformações e ainda assim não ocorrerá nenhuma alteração.

A regra acima mencionada funciona em todos os níveis evolutivos. Por exemplo, há uma hipótese bem fundamentada sobre o papel do resfriamento que ocorreu há 6-8 milhões de anos e levou à formação de grandes espaços abertos na África Oriental. Isso promoveu a evolução dos hominídeos denominados Dryopithecus que viviam em árvores para hominídeos de andar bípede ereto dos Australopithecus ou de outro tipo (Kessler 2017; Niemitz 2010).

Na evolução social, o fenômeno desencadeante seria necessário para a formação de um estado inicial. Além do aumento da complexidade interna de um governo e da estratificação social, um gatilho também é necessário na forma de uma mudança abrupta na sociedade. Este último pode ter sido uma guerra, um reassentamento involuntário ou a abertura da ~~tal~~ sociedade ao mundo exterior (como aconteceu com os havaianos no final do século XVIII com a descoberta das ilhas por James Cook, veja Grinin 2017b).

Juntamente com a condensação da nuvem de gás-poeira, começa uma contração, ou uma queda livre controlada pela autogravidade, que, segundo algumas hipóteses, durou dez mil anos (Marove et al. 2008, 225; Motoyama Kazutaka e Tatsuo Yoshida 2003). O colapso em curso fez o fragmento inicial da nebulosa quebrar em grupos menores, de modo que geralmente pode gerar muitas estrelas. A condensação contínua dentro do aglomerado faz com que sua matéria se concentre gradualmente, preparando assim uma transformação em uma protoestrela. A contração é acompanhada por aquecimento, enquanto a estrutura da futura estrela é formada, incluindo seu núcleo e camadas. O centro da protoestrela se aquece gradualmente.

Após a formação dos núcleos externo e interno do protosol, o resto da matéria periférica rumou parcialmente ao núcleo e adicionou-se à massa da estrela em formação. *Este processo de queda da matéria* (no caso de um protosol - de gás) *na superfície de um corpo é chamado de acreção*. Depois que a camada de acreção essencialmente vai para a protoestrela, esta se transforma em uma jovem estrela. Enquanto isso, sua temperatura interna atinge vários milhões de graus, o que inicia reações termonucleares. A formação do Sol como uma estrela deve levar cerca de um milhão de anos, mas há estimativas prolongando ou reduzindo este período de tempo.

2. A formação dos corpos protoplanetários, embriões de planetas e protoplanetas

O disco protoplanetário e sua evolução. Durante a formação de uma estrela jovem, um disco circunstelar é muitas vezes formado visível através de comprimentos de onda ópticos e curtos. A matéria restante do disco de acreção é parcialmente dispersa no espaço, bem como usada na formação de um disco protoplanetário. De acordo com as observações, esse disco em torno das estrelas existe de 5 a 25 milhões de anos.

A dificuldade na reconstrução do processo de formação planetária para o Sistema Solar é compensada por um grande número de hipóteses e teorias que foram desenvolvidas ao longo de dois séculos. Mas nenhuma das hipóteses pode explicar toda a gama de fatos relacionados aos planetas até agora.

No entanto, a grande maioria dos cosmólogos acredita que o Sol e os planetas foram formados a partir de uma única nuvem (nebulosa protosolar) cuja matéria se diferenciou no Sol e no envelope protoplanetário, o último evoluindo para um disco como resultado da rotação. A rotação e a fragmentação deste disco protoplanetário formaram os planetas ao longo de um novo ciclo de acumulação de matéria em corpos protoplanetários. A maioria dos cosmólogos procede da ideia de que os planetas foram formados a partir de material frio, que foi mais tarde aquecido por onda de choque, radioatividade e outros processos. A formação de um disco protoplanetário acredita-se durar de um a vários milhões de anos. A massa do disco protoplanetário é estimada entre 3 e 10% da massa solar. Além disso, era distribuído espacialmente e heterogeneamente. As dimensões dos discos de acreção das estrelas jovens são de 100 a 1000 unidades astronômicas.

O disco era mais aquecido em suas partes internas, enquanto suas regiões externas permaneceram relativamente frias. Ocorreram algumas contrações, o que contribuiu para o surgimento de centros gravitacionais separados de formação planetária. Ainda assim, o mecanismo desse processo era extremamente controverso.

1. A formação de um subdisco de poeira.

Aparentemente, o disco protoplanetário era composto do gás da nuvem protosolar com predomínio absoluto do hidrogênio molecular e do hélio (todas as outras substâncias equivaliam a menos de 1%). As partículas de poeira, apesar de representar 0,5 a 1,5% de massa, desempenharam um papel peculiar. Esta poeira era como formada por partículas sólidas microscópicas (gelo de água, moléculas pegajosas e átomos, em particular ferro e outras matérias sólidas). Como resultado da formação do protosol que acumulou a maior parte do gás, a concentração de poeira no disco protoplanetário aumentou na fase posterior da sua evolução. Mas, ainda mais, começou a aumentar como resultado da acreção de poeira no plano medial do disco.

Alguns cosmólogos acreditam que a forma mais provável de formação dos embriões planetários é através da acreção de partículas de poeira no plano equatorial do disco pré-planetário (Zasov e Postnov 2011, 199). Como resultado, um subdisco de gás de poeira foi formado no centro do disco, mas a razão entre poeira e gás já variou muitas vezes em comparação com o espaço circundante. Os grãos de poeira também podem aumentar de tamanho (devido a adesões e arrastes). Assim, o potencial sistema planetário passou por uma transição muito importante envolvendo a concentração de matéria sólida (até agora sob a forma de poeira), que desempenhou um papel essencial no crescimento de corpos pré-planetários e posteriores planetas. De acordo com alguns modelos, o disco protoplanetário quase solar evoluiria por um a dois milhões de anos antes da formação de um subdisco enriquecido com poeira.

Na verdade, o subdisco de poeira era comparativamente fino e sua espessura era 103 a 104 vezes menor do que o seu raio. Tinha que ser opaco para os raios do Sol e, portanto, não alcançavam a periferia do disco. Entre outras coisas, isso determinou as condições variáveis para a formação de planetas, dependendo da proximidade com o protosol.

Aqui lidamos com a regra evolutiva geral *da importância da heterogeneidade e das flutuações* . Neste contexto, o pó pode ser considerado como um elemento de heterogeneidade nas nuvens de hidrogênio molecular. E a concentração desta matéria sólida lançou o surgimento de corpos protoplanetários e, mais tarde, de planetas.

Em todos os níveis da Grande História, a mudança evolutiva requer a presença de heterogeneidade crítica que pode desencadear o reagrupamento de matéria ou elementos na reunião. E uma nova estrutura e ordem surgem nessa base. Além disso, uma homogeneidade absoluta torna impossíveis os processos evolutivos.

Por exemplo, uma mutação pode desencadear a especiação; considerando que os grupos de estrangeiros podem desempenhar um papel importante na transformação de muitos grupos étnicos e Estados iniciais.

2. A formação iniciada de corpos pré-planetários.

Como alguns cosmólogos supõem, por algum tempo, devido à gravidade e turbulência, o subdisco pode ter se contraído enquanto as condensações de poeira e gás e outros agregados podem ter sido formadas dentro dele. Mas o ponto de discussão é se os planetas foram formados a partir desses aglomerados de poeira e gás (como mantém a teoria da condensação) ou já de matéria sólida. A teoria da formação de planetas a partir de matéria sólida é chamada de *teoria da acreção sucessiva*. Muitos, se não a maioria dos cosmólogos, consideram o cenário mais provável.

De acordo com isso, as pequenas partículas de poeira se juntam, primeiro formando pequenas partículas de matéria sólida e, em seguida, objetos maiores que gradualmente cresceram em embriões planetários. As partículas de matéria sólida (de poucos a vários quilômetros ou mesmo mil quilômetros) são chamadas de planetesimais.

O estágio mais importante no processo de formação de embriões planetários é a formação de grandes (inteiros) corpos sólidos - planetesimais. Todas as teorias e hipóteses concordam neste ponto. No entanto, em relação ao número, tamanho e outras dimensões desses objetos grandes, há discrepâncias consideráveis. Existem estimativas diferentes do tamanho limite (crítico para o processo) dos planetesimais. Os proponentes da *teoria da acreção sucessiva de matéria por planetesimais* consideram hipoteticamente a formação de milhões e bilhões de corpos de tamanho de um quilômetro, que aumentam gradualmente no processo de enxameamento. De acordo com a teoria da condensação, os maiores objetos podem atingir um tamanho de mil quilômetros.

Entre muitas forças que influenciaram a concentração e acumulação de matéria, a transformação da matéria da protonuvem em objetos sólidos, determinação de órbitas e, em geral, a formação protoplanetária, reconhece-se que duas forças desempenham um papel fundamental na formação planetária: gravidade e radiação solar. E ambos dependem diretamente da distância do objeto em relação ao Sol. Entre as órbitas de Marte e Júpiter, entre 2 e 4 AU do Sol, existe um limite teórico chamado linha do gelo, ou uma linha de neve. A linha do gelo é o local onde a água transita do estado de vapor para o estado sólido, uma vez que a intensidade da radiação solar diminui com a distância das estrelas. “No local onde a temperatura é

160-170K, de modo que a água tenha uma transição do vapor para o estado sólido, as moléculas de água tendem a se acumular à medida que liberam resíduos” (Lin 2008, p. 53). *A linha do gelo se transforma em um aglomerado de gelo* que promove a criação de planetesimais.

Formação de grandes planetesimais. Quando as massas de planetesimais aumentam, sua gravidade permite que elas atraiam partículas próximas. Assim, numerosos planetesimais de quilômetros [de tamanho] captam ativamente o pó primário. Seu crescimento trouxe o surgimento dos enxames dos denominados planetesimais protoplanetários. Gradualmente, emergiu uma “elite” de poucos, composta de corpos do tamanho da Lua ou mesmo de Mercúrio. Existem muitas hipóteses quanto aos mecanismos de sua geração, bem como o número (de vários a cem). Ao longo do tempo, as órbitas dos maiores corpos tornaram-se circulares, o que os fez centros de atração para a matéria envolvente, tornando-se assim os embriões planetários. De acordo com os cálculos, a formação de planetesimais durou de dezenas a centenas de milhares de anos, enquanto a formação de corpos protoplanetários de planetesimais levou vários milhões de anos.

Hipóteses sobre os planetesimais em crescimento e a luta pelos recursos. Os planetesimais cresceriam devido à acreção de matéria, incluindo gás, bem como à atração mútua e colisões acidentais. Mas quanto maior é o planetesimal, mais forte é a sua gravidade e, de forma mais intensa varre seus vizinhos de pouca massa. Quando as massas de planetesimais individuais se tornam comparáveis à massa da Lua, a gravidade aumenta significativamente de modo que eles se tornam capazes de resvalar nos corpos circundantes, escapando das colisões. Como resultado da luta, confrontos e fusões, um pequeno número de grandes corpos cósmicos são formados, chamados de embriões planetários que dominam em suas zonas orbitais e lutam pela matéria restante.

Ao mesmo tempo, os planetesimais crescentes colidem constantemente e, às vezes, fundem-se ou pelo contrário, se separam após choques. As numerosas divisões permitiram que os corpos maiores capturassem mais e mais recursos. Os objetos já grandes o suficiente continuaram a crescer. Gradualmente, os processos de auto-organização começaram a prevalecer neste caos.

Aqui lidamos com a **regra evolutiva geral da luta por recursos e espaço vital**. A luta pelos recursos é um mecanismo comum de seleção em todos os níveis de evolução. A luta pelos recursos é um componente importante da luta darwiniana pela existência no mundo biológico e da competição econômica humana. As vantagens, inclusive as acidentais, desempenham um papel importante em todos os níveis de seleção evolutiva. Sobre a lei da luta pelos recursos, veja também abaixo.

3. Formação do sistema protoplanetário

Problemas e hipóteses da formação de grupos planetários. A maioria dos pesquisadores acredita que o período anterior à formação dos primeiros planetas tenha durado pelo menos vários milhões de anos. Mas as discrepâncias na determinação de sua duração são bastante consideráveis, dependendo se os pesquisadores consideram a formação dos planetas do Sistema Solar como um processo simultâneo ou acontecendo em momentos diferentes. No entanto, até recentemente, a ideia comum foi que todos os planetas foram formados mais ou menos ao mesmo tempo. Atualmente, mais cientistas tendem a acreditar que os planetas surgiram em momentos diferentes, e os intervalos entre suas formações poderiam chegar a milhões ou dezenas de milhões de anos.

Assim, alguns estudiosos pensam que foi Júpiter que veio primeiro, depois Saturno, e muito mais tarde os planetas terrestres foram formados (ver, por exemplo, Lin 2008; Savchenko e Smaghin 2013; Christian 2004, 60 com referência a Taylor 2002, 59-60); ainda outros acreditam que os planetas do grupo da Terra surgiram primeiro (veja, por exemplo, Marakushev et al. 2013; Vityazev et al.) Alguns estudiosos acham que, em primeiro lugar, os planetas terrestres eram semelhantes aos planetas gigantes, mas depois perderiam seus envelopes fluidos (ver, por exemplo, Marakushev e Zinovieva 2013, Yazev 2011, 357).

Há também uma ideia interessante de que existe não uma, mas duas ou mesmo mais gerações de planetas primários. Existe uma opinião de que, não sendo formados adequadamente, esses planetas primários explodiram e geraram o cinturão de asteroides. Ainda outros pensam que Júpiter e Saturno podem ter empurrado os planetas primários para o Sol ou os expelido do Sistema Solar. *Assim, teve lugar mais do que uma tentativa de formar a ordem atual dos planetas no Sistema Solar.*

Aqui lidamos com a regra do caráter arcaico dos sistemas primários. Isso se refere a planetas primários ou estrelas, bem como a espécies biológicas primárias ou digamos, a estados prístinos (sobre o último ver Grinin 2008). Os sistemas não são formados maduros e estáveis. Eles geralmente sofrem várias reconfigurações, incluindo os ciclos de destruição e recriação. É por isso que os sistemas primários costumam parecer arcaicos, enquanto os sistemas superiores emergem como secundários ou terciários e têm mais oportunidades de autorregulação. Consideremos as primeiras estrelas que surgiram no máximo ~~de~~ 200 a 400 milhões de anos após o Big Bang (por exemplo, veja European Commission 2011). É aceito que as primeiras estrelas eram gigantes, muito mais maciças do que a maioria das estrelas formadas posteriormente (May et al. 20 Devido à ausência de carbono, oxigênio e outros elementos que absorvem energia de nuvens condensantes, o processo prosseguiu mais lentamente naquela época; assim, apenas nuvens gigantes poderiam se condensar para produzir estrelas maciças centenas vezes maiores do que o Sol (Ibid.). Essas estrelas gigantes viviam apenas por alguns milhões de anos (quanto maior é uma estrela, menor é a sua vida). Além disso, as primeiras estrelas continham uma pequena quantidade de elementos pesados. Assim, mais de uma geração de estrelas poderia ter mudado até que a quantidade de elementos pesados aumentasse gradualmente. O surgimento de elementos pesados dos "restos estelares das estrelas mortas" assemelha-se à formação de solo fértil dos restos de plantas mortas. A circulação da matéria no Universo sempre é observada em todos os lugares e em todos os níveis (esta é outra lei evolutiva, a respeito da qual veja Grinin 2013, 2014).

As causas das diferenças nos modelos de formação de planetas gigantes e terrestres. Como os planetas do Sistema Solar são divididos em duas categorias (terrestres e gigantes gasosos), o problema da diferença de seus padrões de formação se torna essencial. Esta formação era fundamentalmente a mesma em ambos os grupos, enquanto as diferenças eram determinadas pela distância do Sol, ou o processo de formação de diferentes grupos de planetas era essencialmente diferente, ou ainda existiam outras combinações?

Não há dúvidas de que a distância do Sol definiu as peculiaridades dos modelos de formação do planeta. Diferentes períodos orbitais de embriões planetários (quanto mais longe o planeta está do Sol, maior a órbita) produziram oportunidades para a captura de planetesimais circundantes e, respectivamente, o raio e a massa de um protoplaneta. A linha de neve efetuou uma maior concentração de planetesimais e matéria em certas regiões do Sistema Solar, que também poderia definir o tamanho dos planetas em diferentes regiões.

Existem inúmeras hipóteses que explicam a origem das categorias de planetas observadas. Por exemplo, há argumentos de que os gigantes gasosos provavelmente foram os primeiros planetas a se formar e tomar quase todo o gás, enquanto os planetas do tipo da Terra obtiveram poucos recursos.

Aqui voltamos a lidar com a lei de luta por recursos e observamos que a distribuição de recursos no mundo cósmico é na mesma medida injusta como nos reinos sociais e biológicos. Por exemplo, a luta por recursos que, entre estrelas e galáxias, podem prosseguir sob a forma de enfraquecimento de outro objeto ou sua destruição (por exemplo, através de uma transferência direta de energia e matéria de um corpo para outro), sob a forma de 'incorporação', 'captura', que é 'anexação' de estrelas e aglomerados de estrelas por grupos maiores. Existem muitos casos de coalescência galáctica. Assim, alguns astrônomos afirmam que durante alguns bilhões de anos a nossa galáxia 'conquistou, roubou e rendeu' centenas de galáxias pequenas, pois existem alguns 'imigrantes' evidentes na nossa galáxia, incluindo a segunda estrela mais brilhante no céu do norte, Arcturus (Gibson e Ibata 2007, 30). É amplamente aceito que o surgimento e a expansão de um buraco negro pode levar à "alimentação" de matéria das estrelas e galáxias próximas. No entanto, a 'capacidade alimentar' dos buracos negros é muito exagerada na literatura popular. Em sistemas de estrelas duplas ou em sistemas estrela-planeta, também se pode observar uma forma de interação como a troca de energia e recursos (sobre a luta cósmica pelos recursos, ver também Grinin 2013, capítulo 5).

Hipóteses e teorias sobre os planetas internos. Existem três abordagens principais para a formação de planetas terrestres.

1) A massa do planeta aumenta até o tamanho do presente através do acúmulo de planetesimais (e meteoritos), o que resulta em uma separação gradual do interior do planeta em núcleo, manto e crosta (não em todos os planetas).

2) A formação dos planetas terrestres seguindo o padrão do planeta gigante. No entanto, mais tarde, os planetas terrestres perderam gases no espaço. Respectivamente, apenas seus núcleos internos de ferro-níquel e silicato permaneceram. Assim, os núcleos de silicato e ferro desses protoplanetas gigantes se transformaram em pequenos planetas independentes. A estratificação em núcleos de ferro e fortes camadas de silicatos impediu sua desintegração explosiva (Marakushev et al. 2013, 135-37).

3) O impacto de Júpiter e Saturno na formação dos planetas terrestres (ver abaixo).

Hipóteses e teorias sobre os planetas externos. A teoria da formação planetária presta especial atenção a dois planetas gasosos gigantes que representam 92% da massa de todo o sistema planetário (isto é, Júpiter e Saturno, mas especialmente Júpiter).

Existem duas hipóteses principais que descrevem os possíveis padrões de formação de Júpiter e Saturno compostos principalmente por hidrogênio e hélio. A primeira hipótese da *contração* explica a composição gasosa dos planetas gigantes pelo fato de que as condensações maciças de gás-poeira - protoplanetas - foram formadas dentro de um enorme disco protoplanetário, que mais tarde no processo de compressão gravitacional se transformaria em planetas gigantes. No entanto, esta hipótese não explica por que a composição de Júpiter e Saturno difere da do Sol, bem como alguns outros problemas.

De acordo com a *segunda hipótese de acreção*, a formação de Júpiter e Saturno passou por duas etapas. Na primeira fase, os corpos sólidos foram acumulados de forma semelhante aos processos com planetas terrestres e, após a massa dos maiores corpos atingiram um valor crítico (de duas a dez e mais massas terrestres), a segunda etapa implicaria a acreção de gás nestes já bastante enormes corpos que ocorreram em uma escala de tempo de 105-106 anos. Na primeira fase, um pouco de gás da região de Júpiter se dissipou, de modo que sua composição seria diferente da solar, e isso era ainda mais evidente na formação de Saturno.

De acordo com a concorrente hipótese de *contração*, a temperatura dos planetas gigantes também foi alta no estágio inicial, mas a dinâmica dos processos provou ser mais razoável dentro da hipótese de acreção. A formação de Urano e Netuno, que contém menos hidrogênio e hélio, também é melhor explicada pela hipótese de acreção, já que a maioria do gás já saiu do Sistema Solar depois de atingir a massa crítica.

Ainda assim, o processo de formação do planeta é bastante lento devido à acreção no núcleo. Pode demorar vários milhões de anos. Alguns pesquisadores, além do cenário de acreção no núcleo, também consideram que a instabilidade gravitacional em regiões densas e frias do disco pode levar à formação de planetas. A formação de planetas devido à instabilidade gravitacional pode levar muito menos tempo do que pode exigir quando são formados por acreção no núcleo. Há também uma hipótese que sugere que os gigantes gasosos são formados por um colapso súbito, levando à destruição da nuvem primária de gás e poeira. Mas a maioria dos cosmólogos nega a possibilidade de colapso gravitacional para planetas por causa de suas massas relativamente pequenas (reconhecendo-o apenas para as estrelas).

4. A migração planetária

Como se pensava anteriormente, os planetas permaneciam nas órbitas originais desde a sua formação. Mas, recentemente, tornou-se popular a opinião de que levou aos planetas cerca de um bilhão de anos para ocupar as órbitas atuais. Na sua história inicial, o Sistema Solar era diferente, e é bastante provável que o Sistema Solar externo fosse muito mais compacto, enquanto o cinturão de Kuiper estava localizado mais perto do Sol. Existem muitas sugestões sobre migrações de planetas; ainda, estas são apenas hipóteses.

A mudança da órbita de Júpiter e outros planetas. Há especialmente muitas sugestões sobre as migrações de Júpiter e Saturno. De acordo com uma delas, este gigante gasoso deve ter se formado dentro da parte interna do sistema planetário, perto da linha de neve, quando ainda havia uma quantidade considerável de gás no disco. Então, teve que se mudar para sua órbita atual (Lin 2008). Quando Júpiter deslocou-se para o Sol arrastando Saturno, funcionou como um trator gravitacional, "puxando" várias massas terrestres de gelo no sistema (Batygin et al. 2016). Há uma hipótese de que cerca de 600 a 700 milhões de anos após a formação do Sistema Solar, Júpiter começou a vagar e entrou em ressonância orbital com Saturno. A ressonância mudou as órbitas desses planetas, uma vez que abrandou a migração para dentro e enviou-os de volta para a parte externa do Sistema Solar. A ressonância afetou muito todo o Sistema Solar. Em particular, Netuno e Urano trocaram as órbitas uma vez que Urano costumava ocupar uma posição mais distante do Sol do que Netuno (Ibid.,veja também Batygin e Brown 2016).

Levou algum tempo para que os planetas saíssem de ressonância. Ao longo de alguns milhões de anos, a interação caótica entre gigantes instáveis "empurrou" Júpiter para o seu lugar atual, e outros planetas "se afastaram". Além disso, de acordo com uma das hipóteses exóticas no decurso dessa reconfiguração, um dos gigantes pode ter sido expulso para o espaço interestelar. Aqui queremos dizer o nono planeta hipotético que pode ter existido no passado distante. Aqui, lidamos novamente com a **regra do caráter arcaico dos sistemas primários** segundo a qual são necessárias algumas grandes mudanças (talvez, mesmo ciclos de mudanças) antes que um sistema encontre seu equilíbrio.

Além disso, a chamada era do intenso bombardeio tardio ou, mais precisamente, certa parte dessa época, provavelmente está associada a esse evento de ressonância (ver Bottke et al. 2012, Gomes et al. 2005). Uma imensa quantidade de precipitação de meteoritos caiu em planetas rochosos durante esse período tardio. Estudos relativamente recentes mostraram que esta foi uma era longa, que terminou a 3,2 bilhões de anos, ou seja, durou quase um bilhão de anos.

Colisões e catástrofes na história inicial do Sistema Solar. As mais debatidas são as duas supostas catástrofes que ocorreram durante os primeiros cem milhões de anos. A primeira foi a colisão de Vênus com Mercúrio. Vênus tem uma rotação retrógrada (contra a rotação do Sol em torno de seu eixo) enquanto a maioria dos outros grandes corpos no Sistema Solar rodam na mesma direção com o Sol. Mercúrio possui um núcleo de níquel-ferro não proporcional, uma vez que o seu núcleo metálico equivale a 60 ou mais por cento de sua massa total (Solomon 2003). Existem várias explicações possíveis aqui. Uma afirma que isso pode ser o resultado de uma colisão de Mercúrio com um grande asteroide e, como resultado desse impacto tangente, Mercúrio perdeu a maior parte de seu manto e crosta (Yazev 2011, 48). Há também uma alternativa mais exótica, de que Mercúrio estava inicialmente mais longe do Sol e, além disso, não era um planeta, mas um satélite de Vênus, do qual mais tarde "escapou". Isso explica o tamanho pequeno de Mercúrio, mais apropriado para um satélite, e a rotação retrógrada de Vênus. A teoria dominante aqui é o efeito de maré de um grande satélite (isto é, de Mercúrio) que, há muito tempo, retardava o movimento orbital do planeta e até o fazia girar na direção retrógrada (Ibid., 57-58).

Outra hipótese famosa em relação a catástrofes é a ideia de que entre 30 e 100 milhões de anos após a formação do Sol, um embrião de planeta do tamanho de Marte colidiu com a proto-Terra e gerou uma enorme quantidade de detritos que posteriormente formaram a Lua. Esta tese tem várias alternativas. Existe uma hipótese fascinante de que, por milhões de anos, um protoplaneta Theia pode ter orbitado próximo da proto-Terra e finalmente colidiu com ela. Acredita-se que a colisão tenha ocorrido quase tangencialmente e em uma velocidade relativamente lenta. É por isso que parte do manto da Terra e de Theia foram ejetados para a órbita terrestre baixa e, a partir desses detritos, formou-se a Lua que começou a girar ao longo da órbita circular.

Mais hipóteses sobre colisões. Mencionamos acima que cerca de 600 a 700 milhões de anos após o colapso da nebulosa protosolar Netuno migrou para uma nova órbita. Recentemente, uma hipótese foi apresentada de que não haviam quatro, mas cinco planetas gigantes no Sistema Solar, e que o quinto planeta colidiu com o Netuno migratório e o puxou para a atual órbita enquanto o quinto planeta gigante entrara em colapso em um aglomerado de detritos que Netuno lançou no cinturão de Kuiper, isto é, nos arredores do Sistema Solar (Taylor Redd 2015; Nesvorný 2011).

Aqui lidamos com **um padrão evolutivo amplamente difundido - o de catástrofes.** Pode-se apontar que o drama é característico da Grande História em todas as suas etapas. Em particular, uma hipótese famosa afirma que a extinção do Cretáceo-Paleogeno foi causada pelo impacto de um asteroide em Yucatã há cerca de 65 milhões de anos (Harmon 2010). Além disso, as catástrofes também afetaram consideravelmente o curso da evolução social. Deixe-nos dar o exemplo da Peste Negra na Europa do século XIV. ***As catástrofes são um dos principais mecanismos de seleção em todos os níveis da Grande História.*** Elas podem servir como disparadores, lançando alguns processos, além de destruir os sistemas defeituosos e expandir as oportunidades evolutivas para aumentar a variabilidade.

Cerca de 3,8 bilhões de anos, os gigantes estabeleceram suas órbitas atuais. Considera-se que, após o estabelecimento da ordem atual dos planetas e satélites, não houve mudanças consideráveis no Sistema Solar. Grandes mudanças ocorreram com os próprios planetas e em sua estrutura geológica, clima, composição da atmosfera e outras características.

Conclusão

Agora, podemos resumir as regras, leis e padrões de evolução acima descritos:

- a regra da necessidade de desencadear fenômenos ou eventos para iniciar o processo evolutivo; - a regra da importante heterogeneidade e flutuações; - a lei da luta pelos recursos e pelo espaço vital; - a regra do caráter arcaico dos sistemas primários; - catástrofes como mecanismo essencial de seleção.

Mas estas são apenas algumas regras e leis evolutivas. No entanto, muito do que sabemos sobre tendências, padrões e mecanismos que influenciaram as transformações dentro da Grande História, bem como leis evolutivas, as regras podem ser rastreadas já em sua fase cósmica. Às vezes, de forma incipiente e não sistêmica, ou pelo contrário, as manifestações mais vivas podem ser encontradas apenas na fase cósmica. Assim, quando várias características e padrões típicos da evolução biológica e social (por exemplo, como a luta pelos recursos) são observadas inesperadamente em fases anteriores da Grande História, começamos a perceber que o caráter universal da evolução é uma realidade com numerosas manifestações.

Referências

Adushkin,Vitaly V., Andrey V. Vityazev, and Galina Pechernikova, "To the Elaboration of the

Theory of the Origin and Early Evolution of the Earth", in *Issues of the Origin and Evolution of the Biosphere*, ed. Eduard Galimov (Moscow: LIBROKOM, 2008), 275-296. In Russian (Адушкин, В. В., Витязев, А. В., Печерникова, Г. В. В развитие теории происхождения и ранней эволюции Земли. Проблемы зарождения и эволюции биосферы: сб. науч. работ / под ред. Э. М. Галимова, с. 275–96. М.: ЛИБРОКОМ).

Batygin, Konstantin and Michael E. Brown, «Evidence for a Distant Giant Planet in the Solar System,» The Astronomical Journal 151(2) (2016): 22. URL: http://stacks.iop. org/1538-3881/151/i=2/a=22.

Batygin, Konstantin, Gregory Laughlin, and Alessandro Morbidelli, «Born of Chaos,» *Scientific American* 5 (314) (2016): 28-37. doi:10.1038/scientificamerican0516-28.

Bizzarro, Martin et al. «Evidence for a Late Supernova Injection of 60Fe into the Protoplanetary Disk,» *Science* 316(5828) (2007): 1178-1181. doi:10.1126/science. 1141040.

Bottke, William F et al. «An Archaean Heavy Bombardment from a Destabilized Extension of the Asteroid Belt,» *Nature* 485 (7396) (2012): 78-81. doi:10.1038/nature10967.

Christian, David. *Maps of Time: An Introduction to Big History* (Berkeley: University of California Press, 2004).

Christian, David. «Swimming Upstream: Universal Darwinism and Human History» in *Teaching and Researching Big History: Exploring a New Scholarly Field*, ed. Leonid Grinin et al. (Volgograd, Russia: Uchitel Publishing House, 2014), 19-40.

European Commission, «EU Marie Curie researcher discovers galaxy 13 billion light years away,» *MEMO* 237 (2011). URL: http://europa.eu/rapid/press-release_MEMO-11-237_en.htm.

Gibson, Brad, and Rodrigo Ibata, «The Phantom of Dead Galaxies,» *V mire nauki*, June (2007): 29-35. In Russian (Гибсон Б., Ибата Р. Призраки погибших галактик. В мире науки, июнь: 29–35).

Gomes, Rodney, and others, «Origin of the Cataclysmic Late Heavy Bombardment Period of the Terrestrial Planets,» *Nature* 435 (7041) (2005): 466-69. doi:10.1038/nature03676.

Grinin, Leonid «Early State, Developed State, Mature State: Statehood Evolutionary Sequence,» Social Evolution & History 7 (1) (2008): 67-81.

Grinin, Leonid. The Development of the World in terms of Big History: the Cosmic Evolution (Volgograd: Uchitel, 2013). In Russian (Гринин, Л. Е. Больш.ая история развития мира: космическая эволюция. Волгоград: Учитель).

Grinin, Leonid. «The Star-Galaxy Era of Big History in the Light of Universal Evolutionary Principles,» in *Teaching & Researching Big History: Exploring a New Scholarly Field*, ed. Leonid Grinin, Leonid et al. (Volgograd: 'Uchitel' Publishing House, 2014), 163-187.

Grinin, Leonid. The Development of the World in terms of Big History: the History and Evolution of the Solar System (Volgograd: Uchitel, 2017a). In Russian (Гринин, Л. Е. Большая история развития мира: история и эволюция Солнечной системы. Волгоград: Учитель).

Grinin, Leonid. «Complex Chiefdoms: Precursor of the State or Its Analogue?» in *Chiefdoms: Yesterday and Today*, ed. Robert L. Carneiro, Leonid Grinin, and Andrey Korotayev (Clinton Corners – New York: Eliot Werner Publications, Inc., 2017b), 195-232.

Grinin, Leonid, and Andrey Korotayev, «Introduction. At the Junction of Theories and Paradigms,» in *Evolution: Development within Big History, Evolutionary and World-System Paradigms*, ed. Leonid Grinin and Andrey Korotayev (Volgograd: 'Uchitel' Publishing House, 2013), 5-17.

Gromov, Alexander. *The Amazing Solar System* (Moscow: EKSMO, 2012). In Russian (Громов А. Н. Удивительная Солнечная система. М.: ЭКСМО).

Harmon, Katherine. «A Theory Set in Stone: An Asteroid Killed the Dinosaurs, After All,» *Scientific American*, March 4 (2010). URL: https://www.scientificamerican.com/article/asteroid-killed-dinosaurs/.

Kessler, Peter. «Prehistoric World: Hominid Chronology,» *History Files*, 26 (2017). URL: http://www.historyfiles.co.uk/FeaturesAfrica/HominidChronology1.htm.

Lin, Douglas N. C.. «The Genesis of Planets,» *Scientific American* 298(5) (2008): 50-59. doi:10.1038/scientificamerican0508-50.

Marakushev, Alexey A. and Nina G. Zinovyeva, «Meteorites and Planets of the Solar System,» Prostranstvo i vremya 4(1) (2013). URL: http://j-spacetime.com/actual%20content/t4v1/2227-9490e-aprovr_e-ast4-1.2013.11.php. In Russian (Маракушев, А. А., Зиновьева, Н. Г. Метеориты и планеты Солнечной системы. Пространство и Время. Т. 4. Вып. 1).

Marakushev, Alexey A. et al.»Origin and Evolution of the Solar System,» *Prostranstvo i vremya* 2 (12) (2013): 132-141. In Russian (Маракушев, А. А., Зиновьева, Н. Г., Панеях, Маракушев, Н. А., С. А. Зарождение и эволюция Солнечной системы. Пространство и Время. Вып. 2 (12), с. 132–41).

Marov, Mikhail et al. «From a Proto-Solar Nebula toa Planatary Systam: The Model of the Emergence of the Gas-Dust Disk,» in *Issues of the Origin and Evolution of the Biosphere*, ed. E. Galimov, M. (ed.), (Moscow: Librokom, 2008), 223-274. In Russian (Маров, М. Я., Колесниченко, А. В., Макалкин, А. Б., Дорофеева, В. А., Зиглина, И. Н., Чернов, А. В. От протосолнечного облака к планетной системе: Модель эволюции газопылевого диска. Проблемы зарождения и эволюции биосферы: сб. науч. работ / под ред. Э. М. Галимова, с. 223–74. М.: ЛИБРОКОМ).

May,‎ Brian, Patrick Moore,‎ and Chris Lintott, *Bang: The Complete History of the Universe* (Baltimore, MD: Johns Hopkins University Press, 2008).

Motoyama, Kazutaka and Tatsuo Yoshida, «High Accretion Rate during Class 0 Phase due to External Trigger,» *Monthly Notices of the Royal Astronomical Society* 344 (2) (2003): 461–67. doi:10.1046/j.1365-8711.2003.06833.x.

Nesvorný, David. Young Solar System's Filth Giant Planet? *The Astrophysical Journal* 742(2) (2011): L22. doi:10.1088/2041-8205/742/2/L22.

Niemitz, Carsten. «The Evolution of the Upright Posture and Gait – a Review and a New Synthesis,» Naturwissenschaften. 97 (2010): 241-263. doi:10.1007/s00114-009-0637-3. PMC 2819487 Freely accessible. PMID 20127307 www.ncbi.nlm.nih.gov/pmc/articles/PMC2819487/

Prigogine, Ilya and Isabelle Stengers, *Order Out of Chaos. Man's New Dialogue with Nature* (Bantam Books, 1989).

Savchenko, Valery N. and Vladimir P. Smaghin. *Conception of Modern Natural Studies*. 2 vols. Vol. 2. *Planetary, Chemical, Biological, Evolutionary, Philosophy and Methods, and Mega History of the Universe* (Vladivostok: VGUES, 2013). In Russian (Савченко, В. Н., Смагин, В. П. Концепции современного естествознания: в 2 т. Т. 2. Планетное, химическое, биологическое, эволюционное, философия и инструменты, мегаистория Вселенной. Владивосток: Изд-во ВГУЭС).

Shklovsky, Igor S. Universe, Life, Mind (Moscow: Nauka, 1987). In Russian (Шкловский И. С. Вселенная, жизнь, разум. М.: Наука).

Solomon, Sean. «Mercury: the Enigmatic Innermost Planet,» *Earth and Planetary Science Letters* 216 (4) (2003): 441-455. doi:10.1016/S0012-821X(03)00546-6.

Taylor, Ross. «The Solar System: An Environment for Life?» in *To Mars and Beyond: Search for the Origins of Life*, ed. Malcolm Walter (Canberra: Museu Nacional da Austrália, 2002), 59-60.

Taylor Redd, Nola . Our Early Solar System may have been Home to a Fifth Giant Planet. *Science* (Aug. 11,2015). doi:10.1126/science.aad1604. http://www.sciencemag.org/news/2015/08/our-early-solar-system-may-have-been-home-fifth-giant-planet.

Vityazev, Andrey V., Galina V. Pechernikova, and Viktor Safronov, *The Terrestrial Planets: Origin* *and Early Evolution* (Moscow: Nauka, 1990). In Russian (Витязев, А. В., Печерникова, Г. В., Сафронов, В. С. Планеты земной группы: Происхождение и ранняя эволюция. М .: Наука).

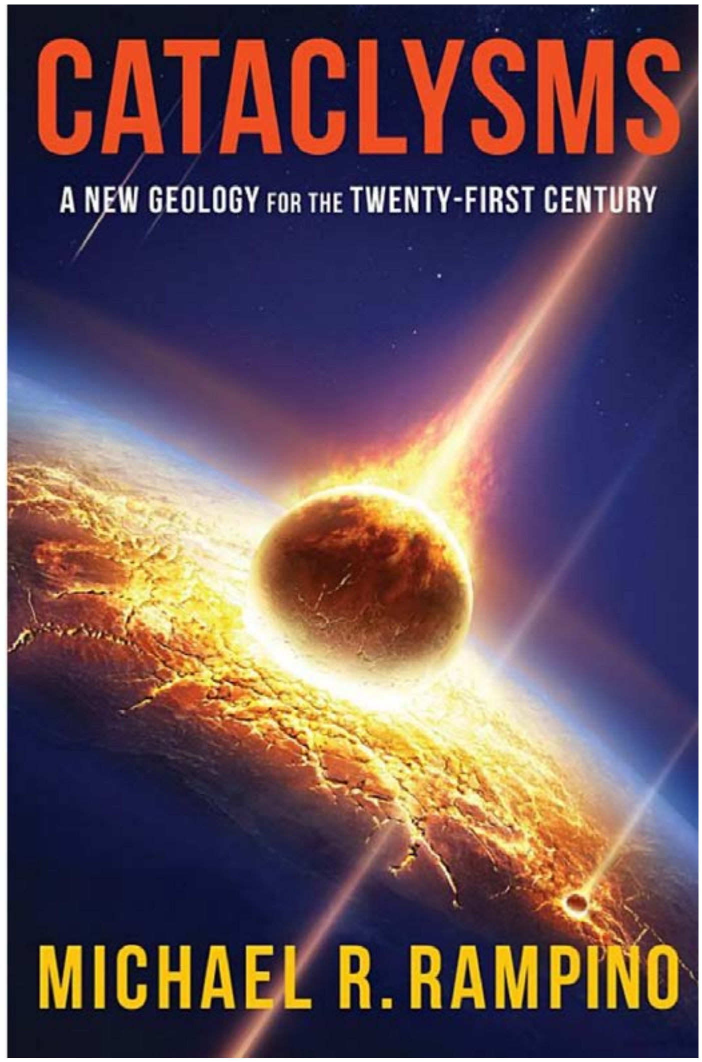
Yazev, Sergey. *Lectures on the Solar System: A Textbook* (St. Petersburg: Lan', 2011). In Russian (Язев С. А. Лекции о Солнечной системе: уч. пособ. / Ред. В. Г. Сурдин. СПб.: Лань).

Zasov, Anatoly and Konstantin Postnov, *The Course of General Astrophisics* (Fryazino: Vek 2, 2011). In Russian (Засов, А. В., Постнов, К. А. Курс общей астрофизики. Фрязино: Век 2).

Book Review
Michael R. Rampino
Cataclysms: A New Geology for the Twenty-first Century
New York: Columbia University Press, 2017
211 pages

Earth's Catastrophic Past: Asteroid Impacts and Mass Extinctions

Asteroid impacts and periodic mass extinctions dramatically revise the Lyellian gradualist paradigm



For a century geologists have adhered to the uniformitarian model presented in James Hutton's *Theory of the Earth* (1794) and documented by Charles Lyell, author of the monumental three-volume *Principles of Geology,* which went through twelve editions (1830-1875). Hutton sought evidence for a vast geological cycle "without beginning, with no end in sight." Lyell abhorred the catastrophist theories of Abraham Gottlieb Werner, the charismatic German lecturer who traced sedimentary rocks to chemical precipitation following the Noachian flood. Lyell focused on uniform structures and gradual change as the key to the geological past, present, and future. He examined deep layers in sedimentary rock in French and English canyon walls, visited the vast carboniferous deposits at the Joggins Fossil Cliffs in the Bay of Fundy, and identified some ninety distinct eruptions at Mount Aetna—all of which testified to gradual changes over vast periods of time. Additionally, he rejected not only mythical astrological influences but also calamitous asteroid impacts as explanations for geological phenomenon.

Given the power of the gradualist model, even through the revolutionary discoveries of plate tectonics in the 1960s, Michael Rampino's *Cataclysms: A New Geology for the Twenty-First Century* at first looks like a throwback to a pre-Lyellian era. But this evocative book is rather a solid geological study founded on a wealth of observations gathered over the past forty years and an invitation to reexamine the history of asteroid impacts from a new perspective. While a full theory of cataclysmic geology is still in the making, the outlines are clear. In place of the fabricated catastrophism of William Whiston and myth-based deluge geology of George

Cuvier and Werner, Rampino has compiled the emerging evidence for regularity in astronomical events affecting the Earth and life, with coherent scientific theory as explanation.

Rampino's starting point is "the Alvarez hypothesis" that emerged with the 1979 discovery by Walter Alvarez of an iridium-rich deposit at the Cretaceous/Paleogene boundary dating to 65 million years ago (mya). Discovered first in the Apennine Mountains in Italy, the same layer was subsequently located at 350 locations worldwide. The tell-tale iridium, an element almost non-existent on Earth, is a certain marker for an extraterrestrial origin. Calculating its source as a 6 to 10 mile diameter asteroid, Alvarez published his theory in *Science* (1980), linking this cataclysmic impact to the already well-established and dated extinction of the dinosaurs. A decade later the 90-mile diameter crater was located under the Yucatan Peninsula, its center near the town of Chicxulub after which it is named. Dated also at 65 mya, this crater has now been accepted as the smoking gun verifying the Alvarez theory.

The effects of the Chicxulub asteroid are several: besides plunging the world into darkness and destroying the food source of dinosaur predators, it jump-started mammalian and primate ascendancy and ultimately cleared the ground for human emergence. It was thus a chance event to which we probably owe our present existence. But the record of earlier mass extinctions raises the question whether asteroid impacts may explain cataclysmic extinction events in the more extensive history of life.

Rampino has been studying mass extinctions and asteroid impacts for more than thirty years, punctuated with exploratory articles in *Science*, *Nature*, *Geology*, and various astronomy journals. Impact craters more than a few tens of millions years old are often hidden from easy discovery by erosion, sediment accumulation, and forest overgrowth, but the larger ones leave behind well-defined evidence. The Chicxulub impact is marked by shocked quartz scatterings, microtektites on the island of Haiti, boundary sediments on Cuba, and widespread deposits from a 100+ meter tsunami in northeast Mexico and the Brazos River Valley in Texas. Paleobiology has established extended records of mass extinctions, presumably with additional evidence yet to be discovered. Five are well known: at the end of the Ordovician (444 mya), the late Devonian (372 mya), end of the Permian (252 mya), end of the Triassic (201 mya), and the end of the Cretaceous (66 mya). But noticeable disruptions in fossil inventories have established numerous intermediate extinctions, some contemporaneous with large-impact craters.

The correlation of asteroid impacts with mass extinctions demands multidisciplinary explanation spanning three hitherto discrete disciplines: astronomy, geology, and biology. During the decades of Rampino's researches, the linkage between impacts and extinctions has achieved a measure of clarity. Initially, asteroid arrivals seemed random, but careful correlation indicates impacts clustering at intervals of 26 million years—a repetition that begs for an explanation. It was assumed that asteroids most often originated in the Solar System Oort Cloud beyond the planets, but what periodic event might account for disruption of the Cloud that could send asteroids careening into the inner Solar System? Given the thousands of Near Earth Objects presently orbiting the Sun alongside the relative rarity of life-threatening impacts, it could be assumed that multitudes of dislodged asteroids would be necessary to produce a measurable uptick of collisions.

Various theories have been suggested: a star on an erratic 26-million year orbit or the passage of the Solar System through the galactic arms of the Milky Way; but decisive evidence is lacking. Rampino gives credit to Richard Strothers for pointing out what is called z-oscillation: the Solar System's periodic crossing of the galactic plane. As it circles the galaxy on its 225-million-year orbit, the Solar System crosses the galactic midplane at an oblique angle, describing a long arc as it soars away and then is pulled back toward the midplane, then crosses it again. Adopting the highly visual metaphor of a carousel up-and-down motion to illustrate the 26-million-year periodic rhythm of our galactic orbit, Rampino and Strothers tendered the theory that mass extinctions were the result of this orbital cycle. The crossing of the galactic plane necessarily exposes the Solar System for several million years to an increased density of galactic material—other stellar systems, nebulae, possibly concentrations of dark matter. The obvious result would be galactic debris swept into the Solar System and periodic disruptions of the Oort Cloud, leading to increasing numbers of asteroid impacts, with species extinction the result. The theory is set forth in a chapter called "The Shiva Hypothesis," a name adopted in an article by Rampino and his student, Bruce Haggerty, more than twenty years ago: "We decided to call it the Shiva hypothesis, after the Hindu god of destruction and renewal. This matched the idea that extinctions destroy the old world and give rise to a subsequent radiation of new species in the postextinction world."

A geology of cataclysm does not replace the earlier gradualism championed by Lyell and a century of followers. Gradualism and concomitant geological preservation is evident virtually everywhere: stable 3.5 billion-year-old cratonic outcrops are visible in Greenland, Northern Canada, South Africa, and Australia; a 4,000-mile line of seamounts northwest of Hawaii testifies to a slow-moving tectonic plate moving across a hot spot now under the Big Island; layered sedimentary rocks are found on sea bottoms, canyon walls, road cuts, and mountain sides—their folds, tilts, and unconformities providing a historical backdrop for Earth history. Cataclysmic impacts punctuate this underlying gradualist narrative without destroying it; an either/or geology is out of date. As Rampino puts it, "The old battle between gradualist and catastrophist views may have only been shadow boxing. Both types of change seem to have been important in creating the geologic record we see today."

Barry Wood
University of Houston
bwood@uh.edu