



**Universidade de Brasília – UnB**  
**Departamento de Economia – FACE**

# Sistemas Complexos Adaptáveis: a importância da Heterogeneidade na performance de equipes

Orientador: Bernardo Pinheiro Machado Mueller  
Aluno: Luísa Jaccoud de Oliveira  
Matrícula: 13/0123048

## Sumário

1. Introdução.....	3
2. Revisão de literatura .....	5
2.1. Premissas.....	7
2.1.1. Reduccionismo.....	7
2.1.2. Agente representativo.....	9
2.1.3. Linearidade.....	9
2.1.4. Regras simples de comportamento.....	11
2.1.5. Evolução.....	12
2.1.6. Heterogeneidade .....	15
3. Performance de grupos em resoluções de problemas .....	17
3.1. Are Two Heads Better than One? .....	17
3.2. Netflix Prize Competition.....	24
4. Diversity Prediction Theorem .....	27
5. Teste.....	31
5.1. Metodologia .....	31
5.2. Teste empírico .....	32
6. Resultados e discussão .....	35
7. Conclusões .....	40
8. Anexo imagens referentes ao teste empírico .....	41
9. Bibliografia.....	41

# 1. Introdução

A Teoria Econômica nos moldes atuais passou por diversas contribuições até a sua formalização contemporânea. Diversos teóricos foram instigados a elaborar modelos e pensamentos de acordo com as respectivas Eras que viveram, contribuindo com a Ciência Econômica da maneira como viam o mundo em seus tempos. A Economia é definida por alguns como uma ciência que estuda a alocação de recursos escassos e por outros como a ciência do bem-estar, mas, em termos da Ortodoxia Econômica atual, todos concordam que a construção de modelos matemáticos capazes de prever a tomada de decisão dos agentes é um dos objetivos centrais da área.

A intuição por trás da Teoria Econômica tradicional dialoga sobre agentes econômicos apresentarem comportamentos embasados na sua racionalidade. O conceito de racionalidade vem sendo lapidado em discussões acadêmicas entre economistas clássicos e economistas comportamentais, caracterizando o momento atual como um período de questionamento das bases da nossa ciência.

Eric Beinhocker (2007), um importante acadêmico da Teoria da Complexidade – área que tem pioneiramente dialogado com a Economia – cita um momento histórico similar ocorrido em 1987 em seu livro *Origins of Wealth*.

Assim, em 1987 quando o Santa Fé Institute, que é hoje um dos principais centros de estudo de complexidade, realizou a lendária conferência entre cientistas físicos e economistas (presentes Kenneth Arrow, Larry Summers, Mário Henrique Simonsen e José Scheinkman, entre outros), os cientistas físicos ficaram impressionados com o intelecto dos economistas, mas ficaram também chocados como as teorias e métodos usados pelos economistas pareciam relíquias de outras eras. A sensação, descreveu um deles, era como uma viagem a Cuba, ou seja, um lugar parado no tempo e fechado do mundo, com carros velhos da década de 50 nas ruas. Era ao mesmo tempo valoroso e patético o esforço e a ingenuidade dos nativos para manter aqueles carros funcionando por tanto tempo usando peças resgatadas de velhos tratores soviéticos. (Beinhocker, 2007, página 47)

A citação anterior mostra a visão que a área da Teoria da Complexidade tem sobre a forma com que a Teoria Econômica é construída até então. Subentende-se que a Economia está em um certo embargo intelectual, distante de um progresso científico coerente à contemporaneidade. A Teoria da Complexidade aborda temas como Teoria do Caos, Teoria de Sistemas, Teoria dos Jogos, Comportamento de Massas, Evolução e Adaptação, entre outros.

O objetivo geral da presente monografia é trazer conceitos oriundos de Sistemas Complexos – mais especificamente Teoria de Sistemas – e fazer análises sobre o funcionamento metodológico da Economia *Mainstream*. O intuito é observar se é possível aplicar com robusteza alguns dos principais conceitos complexos sobre a Economia e tentar examinar a Ciência Econômica sobre uma perspectiva completamente pioneira.

Dos objetivos específicos, busca-se analisar a importância da heterogeneidade na funcionalidade sistêmica ou, em outros termos, se o nível de diversidade presente dentro de um sistema traz benefícios, malefícios ou imparcialidade à resolução de problemas. No contexto descrito até então, não podemos negar que a Teoria Econômica – assim como qualquer outra ciência – é resultado de uma grande heterogeneidade de contribuições de teóricos ao longo dos séculos. Contribuições essas que foram oriundas de diferentes pessoas com diferentes nacionalidades, opiniões, crenças, talentos e limitações. Em outras palavras, contribuições essas que estão vinculadas à capacidade cognitiva de cada indivíduo no momento em que foi formulada sua contribuição. Será que equipes superam indivíduos na performance de resolução de problemas?

O presente documento será subdividido nas seguintes sessões: o capítulo 2 apresentará uma revisão de literatura da Teoria da Complexidade, apresentando as principais premissas desta área do conhecimento e os principais argumentos divergentes com a Economia *Mainstream*. A última premissa abordada será heterogeneidade, que abrirá espaço para o capítulo 3, momento em que serão expostas algumas reflexões sobre a importância de grupos diversos na resolução de problemas – a partir da apresentação do artigo *Are Two Heads Better than One? Monetary Policy by Committee* de Blinder e Morgan (2005) e de um estudo de caso estudado por Scott Page, o *Netflix Prize Competition*.

Em seguida, no capítulo 4, será apresentado o *Diversity Prediction Theorem*, teorema que embasará o teste empírico realizado para a construção desta monografia, devidamente descrito no capítulo 5. O capítulo 6 abordará os resultados do teste com algumas discussões e o capítulo 7, a conclusão da monografia. Por fim, temos um anexo com as imagens usadas na aplicação do teste e a Bibliografia.

## 2. Revisão de literatura

De acordo com Melanie Mitchell (2009), a Complexidade é um campo interdisciplinar de pesquisa que procura explicar como um grande número de entidades relativamente simples se organiza em um cenário coletivo, criando padrões e usando informações que formam um sistema capaz de absorver aprendizado e evoluir ao longo das interações individuais. É importante definir esse cenário coletivo como livre de um controlador central, onde as entidades agem por si só de forma interdependente.

É possível identificar na natureza uma série de exemplos de sistemas complexos e o primeiro exemplo dado por Mitchell (2009) é a organização de um formigueiro. As formigas agem de maneira relativamente simples quando analisadas individualmente, a partir de seus instintos pela busca de alimentos e sua comunicação por sinais químicos. Entretanto, uma vez marchando em conjunto com seus similares, são capazes de criar uma massa complexa e coerente de movimento que mata e devora suas presas com eficácia. Além de formularem em conjunto um bom método de caça, existem, ainda, espécies específicas que são capazes de mostrar comportamentos com um nível de complexidade ainda maior. Um exemplo são as espécies que dispõem seus corpos de maneira a unir-se como um grande bloco complexo e estável para a servir de pontes e balsas. Este sistema é estruturado quando uma formiga percebe sinais ou informações de outras formigas que, por conseguinte, faz com que ela aja de uma certa maneira e produza um novo sinal adicional – que será percebido por uma terceira formiga. Essa troca de informações faz com que todo o conjunto de formigas se comportem de maneira complexa e funcional, caracterizada por um Sistema Complexo Adaptável (SCA).

Outros exemplos de SCA são o funcionamento do sistema imunológico para o combate a doenças diversas, a organização de um grupo de células para formar olhos e cérebros ou, afunilando para a área da presente monografia, como agentes independentes de uma Economia, agindo de forma a maximizar seus ganhos individualmente, formam mercados globais bem estruturados. A natureza mostra uma série de sistemas que seguem esse padrão e, uma das aplicabilidades possíveis para o estudo de Sistemas Complexos é a Economia.

Analisando a Economia como um SCA, podemos identificar os indivíduos antes demonstrados pelas formigas como agora sendo as próprias pessoas ou organizações (grupo de pessoas). Os agentes econômicos – as partes do sistema – são orientados a agir de maneira a maximizar sua própria Utilidade o que, com a interação conjunta, leva à realização de transações comandadas por regras simples de previsão de ganho de *payoff* (ótica microeconômica). A partir disso, são formados padrões de comportamento coletivo que são complexos e de difícil predição, visto que cada agente pode decidir seu comportamento de duas maneiras: ou ele analisa o comportamento de agentes locais – ou seja, os mais relativamente próximos, sem uma perspectiva global – ou ele analisa dados globais, como indicadores de preços e taxas macroeconômicas.

Adam Smith definiu o termo “mão invisível” como o estado de equilíbrio em que os mercados se encontram após todas as interações microeconômicas devidamente realizadas (de interesse próprio, Max U). De acordo com Smith, uma vez que todos os indivíduos interagiram em prol do próprio bem-estar – e em livre mercado –, os preços dos bens serão estabilizados de modo que não há como alterar os padrões de produção e consumo para que todos fiquem em melhor situação. A literatura Econômica atual, partindo de Smith, define que um equilíbrio competitivo se dá no momento em que não há como beneficiar alguém sem deixar outro indivíduo em pior situação. Cenário este de eficiência de mercado ou eficiência de Pareto.

Paralelamente, o estudo da Complexidade abarca a imprevisibilidade e o não equilíbrio sobre o sistema estudado, sendo uma visão oposta à de Smith. A ideia é que as interações microeconômicas geram informação que fazem diferença na interação seguinte. Um preço estabelecido em uma transação seria uma informação resultante de uma interação inicial que, em um segundo momento, influencia uma transação seguinte, ainda não ocorrida dentro do mesmo mercado – que influencia uma terceira transação e o ciclo continua. Um conceito que exemplifica muito bem é a heurística da ancoragem, contribuição essa da Economia Comportamental, mais especificamente Tversky e Kahneman (1974) em *Judgment under uncertainty: heuristics and biases*.

In many situations, people make estimates by starting from an initial value that is adjusted to yield the final answer. The initial value, or starting point, may be suggested by the formulation of the problem, or it may be the result of a partial computation (...) That is, different starting points yield different

estimates, which are biased toward the initial values. We call this phenomenon anchoring. (Página 1128 – “Adjustment and Anchoring”)

Este conceito mostra que a tomada de decisão é feita a partir de estimativas que estão sujeitas a influências de valores inicialmente apresentados aos agentes. As informações geradas em uma transação já ocorrida têm a capacidade de servir de ancoragem para transações futuras. Se cada transação gera uma informação e as transações de um mercado são realizadas a todo instante, o equilíbrio não há de existir na perspectiva da Complexidade.

A aplicação de SCA à Economia fornece contribuições sobre o processamento de informações que o mercado como um todo apresenta e, por consequência, a capacidade do mercado de ajustar preços de forma realmente eficiente a partir de adaptação e aprendizado.

O estudo de Sistemas Complexos Adaptáveis é considerado por muitos como uma área pioneira capaz de trazer contribuições até então não trabalhadas nas Ciências Econômicas. Algumas das premissas iniciais traçadas historicamente para o estudo da Economia *Mainstream* vão em contrassenso com as premissas usadas no estudo de Sistemas Complexos. A seção 2.1 apresentará algumas das premissas mais duais entre essas duas áreas do conhecimento e, em seguida, no capítulo 3, serão apresentadas algumas reflexões específicas sobre a premissa de heterogeneidade e o tema diversidade.

## 2.1. Premissas

### 2.1.1. Reduccionismo

Um sistema reducionista é aquele que assume que, para o estudo de um sistema de forma completa, é possível subdividi-lo em pequenas partes e fazer análises individuais sobre as divisões feitas. Este conceito assume que para o entendimento de um sistema por completo, as partes estudadas separadamente não interferem entre si quando agregadas, mas se completam. Em outros termos, isso significa que o reduccionismo ratifica que a soma das partes é necessariamente igual ao total agregado.

Um exemplo simples que permite uma visualização concreta deste conceito é o funcionamento de relógios de pulso. Para produzir um relógio basta ter o conhecimento sobre o grupo de peças e sobre a ordem da montagem que fará o

objeto funcionar. O ato de construir e desmontar relógios pode ser feito infinitamente e de forma atemporal visto que podemos subdividi-lo em pequenas partes que somadas, resultarão no objeto inteiro e funcional. A analogia que fazemos para o estudo de ciências pelo reducionismo é que podemos estudar cada peça do relógio de forma individual e isolada sem grandes limitações.

Este é o caso da Economia Tradicional. A construção de modelos macroeconômicos é feita a partir de resultados microeconômicos, como por exemplo o próprio conceito de demanda de mercado ou demanda agregada. A partir da premissa individual de que todos os agentes agem de maneira a maximizar sua Utilidade, a microeconomia é capaz de derivar as demandas individuais pelo método Lagrange, permitindo que a macroeconomia derive a curva de demanda de mercado pelo somatório das demandas dos agentes deste mercado. Essa premissa é importante para o estudo da Economia Tradicional pois traz uma rigidez imprescindível para a construção de modelos concretos e deterministas.

Já um SCA apresenta críticas ao reducionismo ao defender que a interação entre as partes pode criar um padrão que influencie outras partes, mostrando comportamentos diferentes se aglomerado o comportamento individual. Isso significa que o contato entre as partes gera informação adicional ao sistema como um todo, fazendo do reducionismo uma estratégia limitante por ter dificuldades de identificar estes padrões adicionais.

Um contraexemplo simples que também permite uma boa visualização da negação deste conceito é o funcionamento de um bando de pássaros quando caçados por predadores (como gavião caçando pardais). A dinâmica de voo de pardais neste cenário de caça se remete a uma dinâmica caótica em um primeiro momento. Com a funcionalidade de confundir o gavião e impedir que o predador pegue qualquer um dos pássaros do bando, os pardais usam da sua proximidade com outros pardais para se aproximar e afastar muito rapidamente do gavião. Isso faz com que ele não seja capaz de alcançar nenhum pássaro justamente por ter muitos pássaros próximos a ele em uma rotatividade altíssima. Se analisarmos individualmente a estratégia de caça para um único pardal, o gavião teria sucesso por ser mais veloz e por se perder o padrão de defesa do bando.

### 2.1.2. Agente representativo

O agente representativo é uma premissa que unifica todos os agentes do sistema com base em suas características universais – ou seja, um agente que representa todos os seus similares. Essa premissa é frequentemente usada em Economia, sendo a população representada por um agente racional que segue axiomas comportamentais.

Essa premissa permite auxiliar na formulação de modelos macroeconômicos a partir de micro fundamentos, o que favorece a robustez de modelos determinísticos. Os economistas assumem que as escolhas de agentes heterogêneos em um mesmo setor podem ser assimiladas a um padrão representativo de maximização de utilidade individual cujas escolhas coincidem com as escolhas agregadas dos indivíduos heterogêneos.

Um SCA critica fortemente todas as simplificações feitas para modelar os agentes como um único ser racional e representativo. Este sistema defende a existência de heterogeneidade entre os agentes, inclusive contando com a diferenciação existente entre comportamentos locais e globais. Um agente tem a possibilidade de analisar o sistema e tomar decisões com base no que ele visualiza do sistema. Entretanto, muitas vezes a visualização dele é somente local, não possuindo uma visão holística do mundo (global).

O mesmo exemplo anunciado sobre o bando de pássaros pode representar também a negação dessa premissa, visto que se formos analisar um dos pardais sendo caçado pelo gavião, a característica universal é voar sempre de maneira a se afastar do predador – para sobreviver. Entretanto, uma vez no coletivo, existe uma dinâmica de afastamento e também de aproximação para confundir o gavião, impossibilitando o uso da premissa de agentes representativos para entender o sistema como um todo – visto que a ideia de aproximação não é observada em um cenário individual.

### 2.1.3. Linearidade

Um sistema com linearidade é aquele que assume relações de proporcionalidade entre suas variáveis de causa e efeito. Isso significa que, grosso modo, pequenas causas geram pequenos efeitos e grandes causas geram grandes efeitos no sistema. Afirmar que o aumento da oferta de moeda pelo Bacen gera, por

consequência, uma queda proporcional na taxa de juros (no sentido de que uma mudança com alta/baixa magnitude na oferta de moeda gera uma mudança com alta/baixa magnitude nos juros, respectivamente) é uma afirmativa comum no estudo da Economia, por exemplo.

Entretanto, a linearidade pressupõe a inexistência de pontos de inflexão (ou *tipping points*) em um sistema. Ponto de inflexão é um termo originado na epidemiologia, que explora a ideia de que muitos sistemas seguem o mesmo padrão de uma epidemia. Aplicado em outras áreas do conhecimento, Malcolm Gladwell (2000) defende a ideia de que comportamentos podem ser “contagiosos” e que, dessa maneira, pequenas causas podem gerar grandes efeitos em um curto período de tempo. De acordo com seu livro “*The Tipping Point*”, o autor define este conceito como:

These three characteristics – one, contagiousness; two, the fact that little causes can have big effects; and three, that change happens not gradually but at one dramatic moment – are the same three principles that define how measles moves through a grade-school classroom or the flu attacks every winter. Of the three, the third trait – the idea that epidemics can rise or fall in one dramatic moment – is the most important, because it is the principle that makes sense of the first two and that permits the greatest insight into why modern change happens the way it does. The name given to that one dramatic moment in an epidemic when everything can change all at once is the Tipping Point. (Page 9, Introduction)

O estudo de Sistemas Complexos surge em contraposição ao conceito de linearidade, defendendo a perspectiva de que existem variáveis diversas, caóticas, relacionadas por rede, sem equilíbrio real e com difícil previsibilidade. Dessa forma, um aumento da variável A não gera uma consequência linear e proporcional em B, necessariamente.

Um exemplo de sistema não linear e com a presença de pontos de inflexão, além do funcionamento de epidemias, é a propagação do ato de bocejar. Uma pessoa que visualiza outra pessoa bocejando – em outros termos, capta informação do ambiente –, muito possivelmente repetirá o ato instantaneamente ou em poucos segundos. Terceiros que visualizarem alguém bocejando, mesmo não tendo visto a primeira pessoa que o fez, repetirá o ato e propagará o ciclo. Isso significa que a informação gerada por uma única parte do sistema pode ter como resultado todo um ambiente “infectado”. Uma causa de magnitude pequena gerou um efeito de grande magnitude. A Complexidade expande estas ideias para entender o funcionamento de

tendências de moda, ondas e fluxos de crimes, aumento do tabagismo na adolescência, a ascensão de livros desconhecidos para *bestsellers* e fenômenos vinculados a comportamentos em massa – presentes por exemplo no mercado financeiro.

#### 2.1.4. Regras simples de comportamento

Um SCA toma como base que os agentes tomam suas decisões a partir de regras simples de comportamento, contrapondo-se à Economia Tradicional, que defende que o agente toma decisões ponderadas e rebuscadas a partir de expectativas racionais ou adaptativas.

O ferramental de Sistemas Complexos assume que o conjunto de regras simples de comportamento – tomadas a partir da interação entre os agentes – é que rege o andamento de sistemas e criam padrões, não as preferências individuais. Um ponto importante a ser ressaltado é que os SCA defendem que as preferências individuais dos agentes (que geram uma tomada de decisão) podem gerar consequências diferentes quando eles interagem (gerando decisões não previstas anteriormente). Um ferramental que une a lógica de Sistemas Complexos na Economia tradicional neste sentido é o estudo da Economia Comportamental, área que define vieses cognitivos e dialoga sobre processos decisórios discrepantes do que seria previsto para agentes racionais.

Um exemplo de modelo que aborda diretamente a premissa de regras simples de comportamento é o *Standing Ovation Problem* (SOP). De acordo com John Miller e Scott Page (2004), o problema baseia-se no estudo da dinâmica de aplausos de uma audiência. Após o término de uma palestra, é natural que o aplauso ocorra e, em alguns casos, algumas pessoas se levantam para propor um maior entusiasmo sobre o palestrante. Uma análise tradicional seria que os agentes se levantam ou não em função estrita das suas impressões individuais sobre a performance da palestra. A diferença é que o SOP pressupõe que o ato de levantar influencia outras pessoas a levantarem – pessoas que não se levantariam em um primeiro momento. Esse comportamento de mimetismo pode ser estratégico (os agentes querem emitir um certo sinal ao palestrante), informacional (talvez a palestra foi melhor do que o agente imaginou) ou ainda conformado (o agente se levanta para não se sentir deslocado).

### 2.1.5. Evolução

De acordo com Darwin, em *A Origem das Espécies*, a seleção natural é definida como uma dinâmica que permite a perpetuação e eliminação de variações favoráveis e desfavoráveis ocorridas inicialmente ao acaso na reprodução de espécies. Essa dinâmica ocorre em prol da competição pela sobrevivência de espécies e em função da dinâmica populacional pré-existente. Ou seja, a variação de comportamentos e habilidades são testadas pela seleção natural, resultando na sua perpetuação ou eliminação que será assim propagada pela hereditariedade da espécie – o que é denominado como algoritmo darwiniano.

(...) if variation useful to any organic being do occur, assuredly, individuals thus characterized will have the best chance of being preserved in the struggle for life and from the strong principle of inheritance they will tend to produce offspring similarly characterized. This principle of preservation I have called, for the sake of brevity, natural selection. (Chapter IV, page 127)

De acordo com Matt Ridley (2015), a evolução, em sua perspectiva biológica, é resultado de recombinação genética – via reprodução sexual – que será testado pela seleção natural e, se tal combinação apresentar vantagens, definirá a evolução em si. O autor defende que é esta a maneira pela qual organismos complexos são construídos do zero, sem um plano final previamente definido. Em seu livro, o autor defende que sociedades humanas apresentam o mesmo mecanismo evolucionário em tecnologia, cultura, crenças, instituições, entre outros. A recombinação de informações é testada pela sociedade com a única diferença de que são escolhidas diretamente quais recombinações serão aceitas – diferentemente da perspectiva biológica – e a sociedade é moldada pelas recombinações que se sobressaem.

Um exemplo dado por Ridley é a estrutura complexa que os olhos humanos apresentam. A visão veio a existir sem um planejamento final da sua estrutura atual. É possível apontar uma série de espécies que apresentam estruturas menos complexas de visão e que, de acordo com mudanças genéticas e a aplicação da seleção natural, gerou um processo evolucionário capaz de construir mecanismos de visão mais rebuscados e com melhor funcionalidade. O ponto é que as estruturas complexas não são desenvolvidas a partir da visualização de um objetivo final e a partir disso traçar os caminhos para tal. Estruturas complexas ocorrem por interação de informações heterogêneas que são combinadas da forma que a seleção natural permite.

Adam Smith, quase 100 anos antes, abordou argumentos similares em Riqueza das Nações, defendendo que o governo não controla a Economia a partir de uma formulação hierárquica de políticas públicas, sendo o governo na verdade uma instituição que somente mostra o reflexo do que é aceito ou não aceito pelas interações da população. A reação das pessoas sobre combinações de informações que a própria população faz define o que a sociedade considera bom ou ruim e, dessa maneira, não importa quem está encarregado no governo, pois seu papel é somente refletir estes padrões de forma clara. De acordo com esse raciocínio, podemos afirmar, ainda, que a Economia não é controlada em prol de um objetivo futuro já traçado, mas em prol de uma dinâmica evolucionista. Esse é um raciocínio que une ambas áreas: Sistemas Complexos e Economia.

Entretanto, sabemos que a Economia Tradicional nos mostra outra perspectiva sobre este aspecto. De fato, a formulação de políticas públicas ocorre com objetivos deterministas e linearizados para fomentar alguma mudança na Economia a partir de incentivos e desincentivos. São visões opostas sobre governança.

Além disso, a aplicação da premissa evolucionista em Economia cria um contraponto com relação à equilíbrios estáticos. Uma simulação feita por Kristian Lindgren, descrita no artigo *Evolutionary Phenomena in Simple Dynamics* (1991), mostrou como a evolução ocorre de forma real na Economia a partir da construção de um jogo decorrente da união do Dilema do Prisioneiro<sup>1</sup> com o Jogo da Vida<sup>2</sup>.

<sup>1</sup> O Dilema do Prisioneiro é uma situação de jogo em que existem 2 agentes suspeitos de cometer um crime. Existe a possibilidade de delatar ou não delatar o outro agente e, a partir dessa premissa, temos 4 estratégias de jogo: I) ambos delatam e ambos recebem 2 anos de prisão; II) Ambos permanecem quietos e nenhum é preso; III e IV) Um agente delata e outro permanece quieto, o que leva ao delator a não ser preso e o delatado a cumprir 5 anos de prisão. O estudo desse jogo é feito com o objetivo de traçar a melhor estratégia de cada agente, visto que um agente não tem conhecimento sobre a estratégia do outro. A melhor estratégia é que ambos os agentes permaneçam calados, mas ao decidir por essa estratégia cria-se uma incerteza pela possibilidade de receber um *payoff* de -5 no lugar de -2. A importância desse estudo é entender a tendência de cooperação e competição na interação de agentes econômicos.

		Prisioneiro 2	
		Permanecer calado	Delatar
Prisioneiro 1	Permanecer calado	0,0	-5,1
	Delatar	1,-5	-2,-2

<sup>2</sup> O Jogo da Vida foi desenvolvido pelo matemático John Horton Conway e é jogado a partir de um conjunto de células agrupadas de forma similar a um tabuleiro de damas. Cada célula pode estar ligada (com cor preta) ou desligada (com cor branca) e tem contato com outras 8 células. A regra simples do jogo é que o *status* das células vizinhas definirão a cor da célula central na próxima rodada. Se existem exatamente 2 vizinhos pretos, a célula central permanece em sua cor original na próxima rodada. Se existem exatamente 3 vizinhos pretos, a célula central se tornará preta na próxima rodada. Tirando

No jogo em questão, cada célula representa um agente que joga o Dilema do Prisioneiro com as 4 células vizinhas (cima, baixo, esquerda e direita) e o jogo ocorre em rodadas que se repetem infinitamente. Foram definidas 4 estratégias de jogo para cada agente, sendo elas: A) Sempre delatar (não cooperar); B) Sempre permanecer calado (cooperar); C) Tit for Tat – o agente sempre escolhe a estratégia escolhida pelo seu oponente na rodada anterior – e D) Anti-Tit for Tat.

Na rodada inicial, as estratégias são alocadas de forma aleatória e, após o fim de cada rodada, cada célula vizinha recebe a pontuação referente à média dos 4 jogos realizados com seus respectivos vizinhos. A célula vizinha com maior média tem a sua estratégia propagada para a célula central. Isso significa dizer que a célula central receberá, na rodada seguinte, a estratégia que teve maior sucesso entre seus vizinhos, como representado na imagem a seguir.

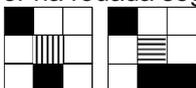
	1	
4 →		2
	-1	

Além disso, foram inseridas algumas estratégias com “mutações”, sendo elas I) *Point mutations*, onde a estratégia é simplesmente invertida; II) *Gene duplication*, onde o agente passa e ter memória histórica das duas rodadas passadas e III) *Split mutations*, onde a memória histórica é reduzida em números de rodadas passadas analisadas.

Essa simulação gerou uma dinâmica em que as estratégias Tit for Tat e Sempre Cooperar geravam muita pontuação quando interagidas e Nunca Cooperar ganhavam força às custas do Anti-Tit for Tat. Enquanto isso, sempre cooperar e nunca cooperar permaneciam com suas estratégias intactas. Com a continuidade de interações, a estratégia Nunca Cooperar formou um mar de não cooperação, enquanto algumas ilhas de cooperação também surgiram. Nestas ilhas, o centro era formado pela estratégia Sempre Cooperar – gerando muita pontuação – e era cercada por um anel de Tit for Tat, que mantinham a não cooperação na fronteira externa.

---

todas essas condições, a célula se tornará branca na próxima rodada. O jogo se repete infinitamente. As imagens a seguir têm o listrado vertical referente às situações em que a célula mantém a mesma cor na rodada seguinte e o listrado horizontal, a célula que se torna preta na rodada seguinte.



Entretanto, foi analisado um padrão instável sobre essas ilhas de cooperação, visto que em alguns momentos as ilhas expandiam-se e em outros eram invadidas por estratégias de não cooperação. Os padrões visuais identificados foram diversos, como quadrados, ziguezagues ou redemoinhos. Em determinado momento, surgiram inovações e as mutações fizeram novas estratégias aparecerem, como o *Simpleton* – a célula coopera se ambos jogadores tiverem escolhido o mesmo movimento na rodada anterior e não coopera caso contrário.

Em *Origins of Wealth*, Beinhocker (2007, capítulo 10, página 225) analisa esta simulação e traz contribuições evolucionistas sobre o estudo de estratégias de escolhas entre agentes. O autor definiu 4 premissas básicas: I) O *trade-off* entre cooperar e competir cria uma tensão central que mantém o sistema em constante desequilíbrio, formando padrões mutáveis de estruturas com auto-organização, crescimento e fragmentação; II) Os agentes que interpretam o Dilema do Prisioneiro não têm como otimizar globalmente seu comportamento, quando na verdade eles analisam a própria situação e história para que assim tracem suas estratégias; III) As interações dinâmicas dos agentes podem levar a padrões de comportamento imprevisíveis e IV) O modelo se inova e a evolução é usada para identificar estratégias bem sucedidas – mostrando o grande *design space* de todas as estratégias do Dilema do Prisioneiro. Por fim, fica clara a visualização de não equilíbrio que os SCA defendem por meio da premissa de evolução – o que se contrapõe à Economia *Mainstream*.

#### 2.1.6. Heterogeneidade

A heterogeneidade nada mais é do que a diversidade das partes do sistema em questão. Isso significa que dentro de um sistema existem unidades individuais ou grupos de unidades com características diferentes e não necessariamente generalizáveis para todas as partes do sistema.

Na Economia Tradicional existe certo nível de heterogeneidade quando diferenciamos comportamentos de consumidores, firmas e Estado; mas tradicionalmente vemos a construção de modelos embasada com o uso de agentes representativos. Já os SCA ratificam a premissa de heterogeneidade ao defender que cada parte do sistema possui um comportamento não unificado.

Um exemplo desta premissa em um Sistema Complexo é o funcionamento de um esporte de quadra, como o futebol. Um jogador atacante, por mais que ele seja considerado o melhor atacante da história do futebol, quando colocado na posição de defesa ou até como goleiro, pode apresentar habilidades questionáveis. Neste sentido, a dinâmica de um jogo de futebol pode ser vista como um Sistema Complexo e apresenta clara heterogeneidade ao mostrar-se com uma divisão de partes diversa (jogadores) e com habilidades não únicas. Outro exemplo claro da premissa de heterogeneidade é o estudo de Lindgren (1991) apresentado anteriormente, sobre evolução, cujo conjunto de estratégias e mutações é caracterizado por ser diverso e heterogêneo.

De acordo com *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies* de Scott Page, a diversidade pode ser definida como:

Diversity as characterized in the book means difference in how people see, categorize, understand, and go about improving the world. (Page 1)

É importante ressaltar que Page dá ênfase, de acordo com o próprio autor, na diversidade cognitiva, ou seja, a diversidade advinda da percepção, associação e raciocínio. Entretanto, é fato que a diversidade cognitiva não nega a diversidade de identidade, visto que pessoas pertencentes à diferentes grupos ou experiências de vida tendem a desenvolver habilidades cognitivas de forma diversa.

A diversidade cognitiva, em um primeiro momento, pode ser visualizada como óbvia para a construção de um benefício coletivo, visto que a diversidade de informações tende a reduzir o viés na resolução de um problema. Mas quando analisada com cautela, podemos ver que muitas vezes o progresso e a inovação são impedidos justamente pelas diferenças identificadas em terceiros. Em muitos momentos a diversidade causa estranhamento no indivíduo, levando-o a recuar e ter uma resposta negativa a ela.

Essa é a conclusão de um estudo feito por Robert Putnam (2007) sobre a confiança entre agentes de uma mesma vizinhança. Ele mostrou que o engajamento cívico é menor em uma vizinhança mais diversa. Quanto maior o nível de diversidade em uma vizinhança, menos confiança existe entre os vizinhos. Entretanto, quando levado em conta a variável controle, como crime, renda, educação ou tamanho da cidade, os efeitos negativos da diversidade diminuem consideravelmente.

Esses resultados inicialmente tratados como antagônicos fez Page contra argumentar que existe uma diferença básica entre sua teoria e a pesquisa de Putnam: Page aborda especificamente a produtividade de um grupo diverso e Putnam trata de diferenças étnicas. Quando abordamos diversidade no quesito habilidades e ferramentas, é de se esperar que ela tenha a capacidade de alavancar os resultados de qualquer grupo que tenha alguma funcionalidade.

Para Page, ainda, a diversidade supera a capacidade do grupo homogêneo e especializado. Isso significa que um grupo com baixos níveis de diversidade, por mais que apresentem ferramentas e habilidades ótimas para a concretização da atividade em questão, tem muitas vezes resultados piores do que um grupo diverso pode alcançar. Especialmente para habilidades de predição, talentos individuais e diversidade coletiva têm a mesma importância. Isso significa que mesmo se a opinião de terceiros for menos precisa do que a nossa, mas havendo diferenças e diversidade, a combinação das informações faz com que as previsões sejam mais acuradas, precisas.

A seção a seguir apresentará reflexões sobre o tema a partir de um artigo e um estudo de caso. Primeiramente o artigo de Alan Blinder e John Morgan que analisa se há reais diferenças entre a performance de grupos e de indivíduos e, em seguida, um estudo de caso apresentado por Scott Page sobre a empresa Netflix, que aborda a objetivamente a performance de equipes em função da diversidade de habilidades.

### 3. Performance de grupos em resoluções de problemas

#### 3.1. Are Two Heads Better than One?

O artigo *Are Two Heads Better than One? Monetary Policy by Committee* de Blinder e Morgan traça uma abordagem experimental a partir de duas simulações laboratoriais em que se foi analisada, *coeteris paribus*, a natureza do corpo decisório. A questão central baseou-se em analisar como os vieses individuais são refletidos nas decisões em grupo. A importância econômica desse questionamento se dá, por exemplo, em um contexto de formulações de Políticas Públicas, visto que são feitas por comitês na maioria dos países. As atividades empenhadas por Bancos Centrais

no geral são rodeadas por tomadas de decisões importantes e a formulação de políticas monetárias é feita de forma coletiva pelos comitês responsáveis. Este artigo de Blinder e Morgan visa simular o funcionamento dessas decisões sobre políticas monetárias a partir de dois experimentos que serão descritos a seguir.

Primeiramente, foram definidas funções objetivas para cada experimento – sendo elas conhecidas pelos respectivos participantes –, com o objetivo de distinguir claramente para todos o que foi tratado como uma decisão boa e uma ruim. Cada grupo experimental consistia de 5 pessoas que alternaram jogadas individuais e grupais. Isso significa que tanto os grupos como os indivíduos operaram com quantidade semelhante de informações, pois foram compostos pelas mesmas cinco pessoas. Esse experimento foi realizado com 100 pessoas e foi formulado com 20 sessões, cada uma com 20 rodadas/períodos (representando trimestres).

O primeiro experimento consistia em simular o processo decisório de formuladores de políticas monetárias a partir da exposição de duas equações da literatura teórica. Os valores de parâmetros e coeficientes foram escolhidos pelos autores – justificada a escolha, no artigo, por semelhança à economia dos EUA na época.

$$(1) U_t - 5 = 0,6 (U_{t-1} - 5) + 0,3(i_{t-1} - \pi_{t-1} - 5) - G_t + e_t$$

$$(2) \pi_t = 0,4\pi_{t-1} + 0,3\pi_{t-2} + 0,2\pi_{t-3} + 0,1\pi_{t-4} - 0,5(U_{t-1} - 5) + w_t$$

A equação (1) trata da combinação da curva IS com a Lei de Okun e a equação (2) representa a Curva de Phillips aceleracionista. As variáveis são  $U$  (taxa de desemprego),  $i$  (taxa de juros nominal),  $\pi$  (taxa de inflação),  $G$  (efeito das ações fiscais sobre o desemprego),  $e_t$  e  $w_t$  (choques estocásticos).

No modelo escolhido pelos pesquisadores, foi definida uma taxa natural de desemprego de 5% e taxa de juros real de equilíbrio também em 5%. O desvio da taxa de juros real de seu valor de equilíbrio é representado por  $\Delta i_t^r = i_t - \pi_t - 5$ . A taxa alvo de inflação é 2% e taxa alvo de desemprego é 5%. Os participantes devem focar em manter esses indicadores sempre o mais próximo possível desses valores-alvo, tendo controle somente da taxa nominal de juros.

O equilíbrio do estado estacionário do modelo em equilíbrio é  $U = 5$  e  $i - \pi = 5$  (juros real). Como a curva de Phillips de longo prazo é vertical, qualquer taxa de inflação constante pode caracterizar o estado estacionário.

A dinâmica do jogo foi complementada com uma característica estatística no choque fiscal  $G_t$ , definido no experimento como um evento aleatório que começa o jogo nulo e muda repentina e aleatoriamente para +0,3 ou -0,3 em algum momento nos primeiros 10 períodos. Fica evidente que  $G_t$  influencia em proporção 1:1 no desemprego (mas em sentido contrário). O objetivo dos participantes é reagir às mudanças em  $G_t$  mexendo na taxa de juros nominal (a reação sempre ocorre com atraso temporal e a duração desse atraso foi uma das observações captadas pelos pesquisadores). Além disso,  $e_t$  e  $w_t$  são choques estocásticos e têm distribuição uniforme no intervalo [+0,25; -0,25], com desvio padrão aproximadamente 0,14 (metade do tamanho de G).

Por conta da equação (2), a política monetária afeta somente indiretamente a inflação, com atraso de dois períodos (posições de  $U_t$  e  $U_{t-1}$ ). Todos os participantes estavam cientes dessa defasagem temporal, mas não tinham o conhecimento dos coeficientes ou da estrutura de defasagem do modelo em pauta. Seus conhecimentos se limitavam à probabilidade do choque G, que mudaria em distribuição uniforme entre os períodos 1 e 10. A equação (2) se baseia em uma raiz unitária, fazendo com que o modelo divirja do equilíbrio quando ocorre um choque  $G_t$ . A menos é claro, que seja estabilizado pela política monetária.

Todas os 20 jogos, com a participação dos 100 jogadores, foram iniciados sempre com  $U_0 = 5\%$ ,  $\pi_0 = 2\%$  (atuais e defasadas),  $G_0 = 0$ ,  $i^r = 5\%$  e, por consequência,  $i_0 = 7\%$  (atuais e defasadas), visto que  $i_t^{nominal} = i_t^{real} + \pi_t$ . Como proposto em todas as rodadas,  $e_t$  e  $w_t$  foram iniciados com valores sorteados pelo computador, assim como em todas as rodadas.

Depois de todas as taxas definidas inicialmente, o computador calcula os valores de  $U_1$  e  $\pi_1$  do período 1 e expõe para os participantes verem. Normalmente, muito próximos dos valores ótimos  $U = 5\%$  e  $\pi = 2\%$ . Após isso, novos valores eram sorteados para os choques estocásticos e as variáveis defasadas que apareceram nas equações foram atualizadas. A cada rodada, o computador exibia as taxas de desemprego e inflação da rodada que acabara de ocorrer, assim como as antigas – criando, assim, uma série histórica. Os jogadores deveriam sempre escolher a taxa de juros a ser implementada no próximo período com o objetivo de tentar manter o desemprego e a inflação em seus valores ótimos.

Com o intuito de avaliar a qualidade das decisões, foi definida uma função perda a cargo da diferença entre as taxas alvos e as taxas alcançadas por cada participante a partir de sua performance. A pontuação do jogo seria a média de  $s_t$  ao longo dos 20 trimestres apresentada a seguir.

$$(3) s_t = 100 - 10|U_t - 5| - 10|\pi_t - 2|$$

Para cada alteração na taxa de juros nominal, foi cobrada uma taxa de 10 pontos com o intuito de diminuir a frequência de mudanças pelos participantes para deixar o experimento mais realista, visto que o FED não altera a taxa de juros nominal todo trimestre – pois provavelmente apresenta algum custo sobre essa ação. No final de toda seção, as pontuações das pessoas foram convertidas em incentivos monetários pela participação e performance na pesquisa.

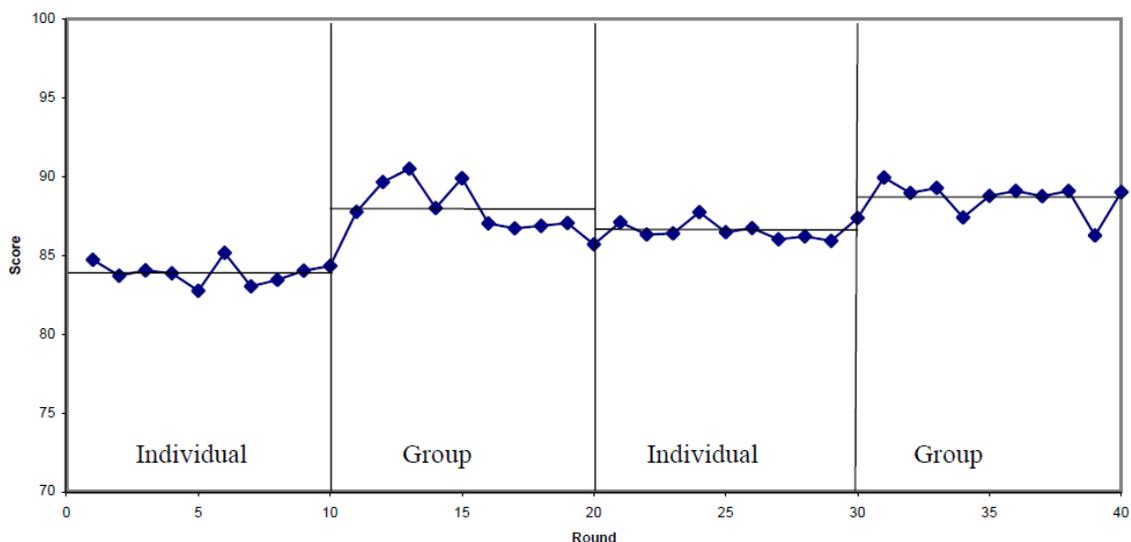
Na Parte Um do experimento, cada um dos 5 participantes jogou somente 10 rodadas do jogo, com toda e qualquer comunicação entre si vedada. Na Parte Dois do experimento, os 5 participantes se uniram e jogaram coletivamente mais 10 rodadas, recebendo pontuações comuns do grupo. Na Parte Três os participantes jogaram novamente individualmente e, por fim, na Parte Quatro, coletivamente mais uma vez. No total, cada sessão é formada por 40 rodadas, alternando-se jogos individuais e coletivos.

Essa modelagem permitiu que se analisasse qualitativamente as respostas dos participantes em termos de boas e más decisões, além de permitir uma comparativa direta entre as respostas individuais e coletivas. Um ponto importantíssimo é que o fato do coletivo ser formado pelas mesmas pessoas que atuaram individualmente, faz com que não haja assimetria de informação no sistema analisado. Além disso, pela pontuação coletiva ser equivalente para todos os membros do grupo e a pontuação individual servir de incentivo para a otimização de suas escolhas (por ser convertida posteriormente em ganhos monetários), não há incentivo à competição, mas à cooperação.

Uma das hipóteses analisadas nesta pesquisa foi se grupos tomam decisões melhores do que os indivíduos. Se os nossos participantes se comportassem como *homo economicus*, todos tomariam as mesmas decisões. Mas não foi o comportamento visto no experimento: a pontuação média sobre o grupo foi de 88,3% (com desvio padrão de 4,7%) e a pontuação média individual foi de 85,3% (com desvio padrão de 10,1%). Essa diferença é economicamente e estatisticamente

significativa ( $t=5,9$  se tratarmos a rodada como a unidade de observação). Os grupos apresentaram estatísticas  $t = 2,2$  e os indivíduos,  $t = 3,4$ ; o que mostra, também com significância estatística, que grupos superaram os indivíduos.

Acentuando ainda mais este resultado, os pesquisadores definiram uma variável  $C$  equivalente à 1 se a primeira alteração da taxa de juros nominal for feita na mesma direção que a alteração  $G_t$  e equivalente à zero se na direção oposta. O valor médio de  $C$  foi 0,843 para indivíduos e 0,905 para grupos (diferença estatisticamente significativa,  $t = 3,6$  quando a decisão individual é tratada como uma observação). Ao jogar como indivíduos, os criadores de políticas monetárias movimentaram a taxa de juros nominal na direção errada em 15,7% do tempo. Sob a perspectiva coletiva, o mesmo erro foi de 9,5% do tempo. A taxa de erro foi reduzida em 40% quando tomadas as decisões coletivamente. A figura a seguir foi retirada do artigo e mostra a pontuação média por rodada.



BLINDER, Alan and MORGAN, John (2005)

Um fator que pode fazer significativa diferença neste experimento é a existência de aprendizado pelos participantes ao longo das rodadas. Os pesquisadores precaveram-se explicando e simulando o jogo antes de iniciar oficialmente cada seção. De acordo com o artigo, não há indicação de aprendizado visto que não há aumento contínuo das pontuações ao longo das 40 rodadas. Na verdade, o que a imagem apresenta é que a primeira experiência coletiva (rodadas 11-20) não apenas produz um melhor desempenho, mas torna os indivíduos melhores formuladores quando voltam a jogar sozinhos (21-30). Além disso, o desempenho

médio individual ainda assim é inferior ao desempenho coletivo imediatamente anterior (11-20).

Analisando estatisticamente, o aumento da pontuação média das duas partes individuais (1-10 e 21-30) foi de 3,2% com significância estatística  $t = 6,1$ . Já o aumento da pontuação média da Parte Dois e Quarto teve diferença de 0,9% com  $t = 1,6$ , estatisticamente não significativo.

O segundo experimento foi criado com a motivação de trazer o mínimo possível de bagagem intelectual anterior, para analisar a tomada de decisão mais “pura” possível. Este jogo se trata do clássico “problema da urna”, onde uma urna eletrônica contém inicialmente 50% de bolas azuis e 50% de bolas vermelhas. Os participantes foram informados que a composição da urna se alteraria para 70:30 – não sabendo qual cor apresentará qual proporção – em algum ponto selecionado aleatoriamente no experimento até o 10º sorteio.

O objetivo dos participantes é acertar qual das cores terá sua proporção aumentada dentro da urna. A função objetivo aqui definida julgou velocidade e precisão como as variáveis independentes. Todo participante iniciava o jogo com 40 pontos e teria a possibilidade de ganhar 60 se acertasse o sentido da mudança (qual cor apresenta 70%). Entretanto, a diferença de “rodadas-sorteio” entre o momento da mudança de proporção e o momento em que o participante faz seu palpite é igual à quantidade de pontos perdidos no final da rodada. Isso significa que se a programação mudou a proporção no 8º sorteio e o palpite ocorreu no 4º sorteio, o participante perde 4 pontos.

A função objetivo é descrita a seguir, cujas variáveis são  $S$  (pontuação que será convertida em incentivos monetários aos participantes no final do experimento),  $C$  (variável fictícia equivalente à 1 se o palpite é correto e 0 se incorreto) e  $L$  (atraso da decisão, definido como a diferença entre a rodada em que a composição mudou e a rodada que foi realizada o palpite).

$$(4) S = 40 + 60C - |L|$$

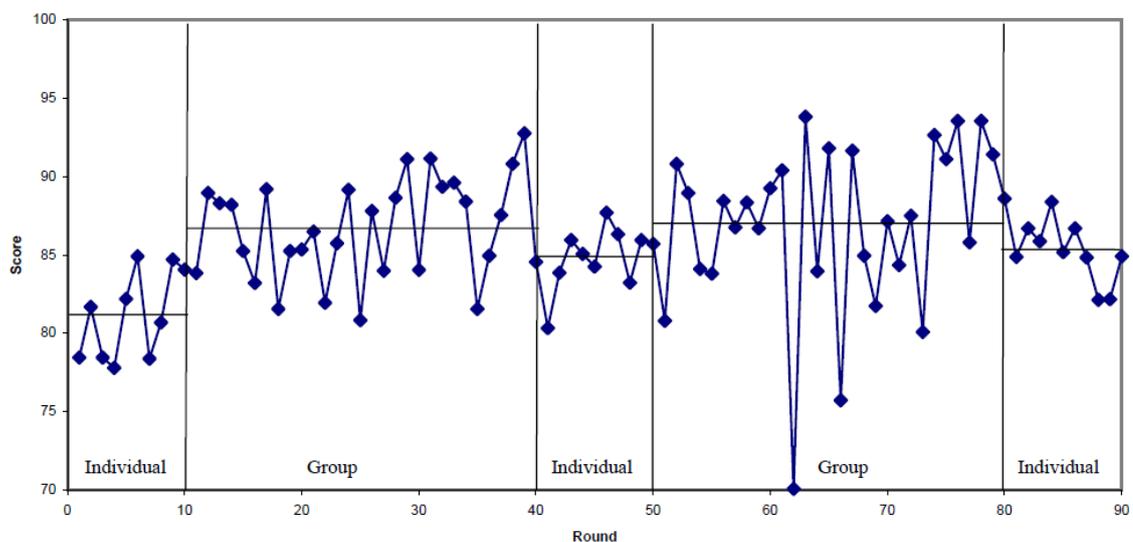
A proporção de 60:1 para o acerto teve essa alta magnitude justificada por dissuadir os participantes a se anteciparem adivinhando a cor mais “cedo” e os 40 pontos livres foram fornecidos em cada rodada a fim de tornar pontuações negativas impossíveis (o experimento não passou de 40 rodadas por seção).

Na Parte Um do experimento, cada um dos 5 participantes jogou somente 10 rodadas do jogo, com toda e qualquer comunicação entre si vedada. Na Parte Dois do experimento, os 5 participantes se uniram e jogaram coletivamente mais 30 rodadas, recebendo pontuações comuns do grupo. Na Parte Três os participantes jogaram novamente individualmente 10 rodadas e na Parte Quatro, mais 30 rodadas coletivas. Por fim, foi incluída uma Parte Cinco com mais 10 rodadas individuais. No total foram 90 rodadas jogadas por seção, alternando-se jogos individuais e coletivos e as pontuações dos participantes também foram convertidas em incentivo monetário.

Na amostra geral, a pontuação média alcançada pelos grupos foi de 86,8 contra apenas 83,7 para indivíduos – melhora de 3,7%, o que é estatisticamente significativo  $t = 4,3$ .

Em comparação com as dez rodadas anteriores na Parte Dois, as pontuações do grupo são cerca de 3,8% melhores; e essa diferença é significativa ( $t = 2,0$ ) nos níveis convencionais. Comparando as Partes Três e Quatro, os escores dos grupos ainda estão 2,3% acima dos escores individuais; mas agora a diferença não é mais significativa ( $t = 1,2$ ). Ainda assim, a conclusão geral apoia a noção de que grupos superam indivíduos.

A taxa de erro (frequência de adivinhação da cor errada) foi de 15,7% para indivíduos, mas apenas 10,7% para grupos. Essa diferença no desempenho também é estatisticamente significativa –  $t = 4,2$  com observações individuais.



BLINDER, Alan and MORGAN, John (2005)

Sobre a perspectiva de performance individual e coletiva, ambos experimentos de formulação de política monetária e do problema da urna apresentaram resultados semelhantes de que grupos superam indivíduos na resolução e previsão de problemas. Presumivelmente, o artigo insinua que existe uma sabedoria coletiva consequente de um mundo complicado e diverso, onde ninguém conhece o modelo verdadeiro ou todos os fatos. Os dados necessários para qualquer tomada de decisão podem ser difíceis de processar ou interpretar e os juízos de valor individuais podem influenciar essas decisões. Quando disposto de um grupo de diferentes indivíduos, essas influências tendem a se equilibrar, pois a agregação de decisões individuais tende a tornar a decisão coletiva inercial.

### 3.2. Netflix Prize Competition

Concluimos até então que grupos superam indivíduos na resolução de problemas. Uma questão ainda mais profunda sobre este assunto é se o nível de diversidade de um grupo faz diferença na resolução de problemas. Um estudo de caso que mostra muito bem a importância da diversidade dentro da coletividade é o *Netflix Prize Competition*, apresentado por Page em *The Hidden Factor*, (2012, página 31).

A Netflix é uma empresa global provedora de filmes e séries que atua em plataforma online. A Diretoria Executiva da empresa Netflix desenvolveu um algoritmo computacional denominado Cinematch que previa as preferências de conteúdo (filmes, séries, documentários, etc) dos seus consumidores dentro da sua plataforma. Em 2006, a empresa lançou uma competição com o objetivo de aprimorar seu algoritmo, consistindo no desafio de aprimorar em 10% a previsão de classificação de conteúdo por parte dos consumidores.

Um sistema de previsão rebuscado e preciso é importante para a empresa visto que está diretamente associado a ganhos de poder de mercado. Uma plataforma que é capaz de prever com maior precisão as preferências de conteúdo do seu consumidor faz com que ele seja menos propenso a abandonar a plataforma, o que é sinônimo de ganhos financeiros para a empresa. O Cinematch previa uma sugestão de filme a partir da classificação dada pelo consumidor sobre algum filme já visto – e talvez em conjunto com informações demográficas. Supondo que o consumidor deu nota máxima (nota 5) no filme *Thor: Ragnarok*, a sugestão mais provável seria

*Avengers: Infinity War* no lugar de 50 Tons de Cinza, por exemplo. O aumento do número de classificações pelo usuário faz com que o algoritmo seja mais eficiente, obviamente.

A Netflix ofereceu mais de 100 milhões de classificações por cerca meio milhão (500.000) de usuários que avaliaram cerca de 18.000 filmes. Todos juntos, os dados consistiram em mais de 100 milhões de classificações de filmes. Esses dados foram divididos em duas partes: *training set* e *testing set*. O primeiro *set* foi divulgado na internet para que os competidores do mundo todo pudessem formular seus modelos. O segundo *set* foi dividido em 3 (cada um com 1,5 milhões de notas de filmes): um para que os competidores pudessem praticar seus modelos já formulados; outro usado para classificações formais e, por fim, a terceira parte para a validação final e determinação do ganhador. No início da competição, os competidores formularam uma variedade de abordagens para a modelagem.

Nearest-neighbor methods definiu que para cada par de filmes, existe uma medida de similaridade. Com isso, este método previa a nota que os usuários dariam de acordo com as notas já dadas à filmes “semelhantes”, como Rocky e Rocky II. Isso significa que a plataforma indicava os filmes que seriam teoricamente similares uns aos outros.

Decomposition methods definiu que cada filme poderia ser descrito por uma lista de atributos. O filme *Titanic*, por exemplo, poderia ser discernido em romance, drama, desastre ou morte, alto orçamento, vendas de sucesso, trilha sonora melosa e históricos. O vetor de atributos de cada filme era estruturado e o algoritmo fazia suas previsões estimando qual nota entre 1 e 10 seria dada a cada atributo pelo histórico de notas do usuário – formando um vetor de preferências sobre atributos para cada usuário<sup>3</sup>. A metodologia consistia em somar as notas de cada atributo para moldar o quão satisfeito o usuário ficaria com o filme em questão, elencando assim os filmes com maior soma para a sugestão da plataforma.

Para reduzir o viés de filmes com sucesso global e de usuários com pouca propensão a dar notas aos filmes, foram usadas duas variáveis *dummy*: *movie dummy* e *person dummy*. A primeira levava em consideração a classificação média do filme por outros usuários (filmes lançamentos ou em alta), dando maior peso para filmes

---

<sup>3</sup> Se o usuário deu notas boas para filmes que apresentassem atributos similares, este atributo teria um peso maior na sugestão estimada pela plataforma.

“bons” e menos peso a filmes “ruins”. A segunda corrigia o caso de usuários que dão somente notas altas (como 4 ou 5) ou somente notas ruins (como 1 ou 2).

Matrix decomposition methods é similar ao anterior, exceto pelo fato de que o algoritmo usa todos os dados sobre os filmes como informação para construir os atributos. Conectando todos os tipos de características dos filmes, o algoritmo constrói os ditos fatores latentes. Um filme com perseguição de carros pode ter associado suspense e bombas, formando essas 3 características um só fator latente.

O competidor que liderou a competição inicialmente foi o grupo BellKor, agrupando entre 50 e 200 fatores latentes e fazendo um algoritmo que permitia que os pesos atribuídos a cada atributo mudasse com o tempo. Isso significa que o grupo BellKor permitiu abertura para mudanças de preferências dos usuários com o tempo, aumentando ainda mais a robustez de seu modelo.

O melhor modelo feito pelo grupo BellKor usou aproximadamente 50 variáveis e foi capaz de melhorar o Cinematch em 6,58%. Entretanto, o grupo formulou não somente 1 modelo, mas 107. A união dos modelos em um único grande modelo aprimorou o Cinematch em 8,43%. Em 2008, após ganhar parte do prêmio – pois 8,43% foi o modelo mais próximo de 10% - e após a finalização da competição, BellKor decidiu abraçar a diversidade e uniu-se à BigChaos, equipe essa formada por cientistas da computação australianos. Foram aprimoradas as combinações de modelos de forma não linear no lugar de atribuir pesos às atribuições e previsões.

Em 2009, a equipe Pragmatic Theory também foi adicionada ao grupo, sendo ela formada por canadenses especializados em modelar a mudança de preferências. Este grupo final, denominado BellKor’s Pragmatic Chaos (BPC) formulou um modelo final com mais de 800 variáveis. Essa diversidade de conhecimentos fez com que o modelo final construído alcançasse o nível de aprimoramento em 10%.

Com isso, o Netflix Prize Competition foi reaberto para testar este novo modelo visto que teria muitas vantagens mercadológicas com essa nova tecnologia. Com o concurso ainda em aberto, as equipes perdedoras, na esperança de que as suas diferenças (portanto diversidade) fizessem diferença, uniram-se como The Ensemble (TE). A proposta da TE era unir-se pela diversidade e, com isso, buscou estrategicamente a melhor combinação possível de modelos. Incrivelmente os TE fizeram um modelo ainda melhor que o BPC, somente melhor na 5ª casa decimal.

As conclusões que tiramos com esse estudo de caso é que a diversidade de habilidades e conhecimentos de fato trouxe uma melhora nos resultados das equipes. Como os dois modelos finais se diferem em algum nível, recombina-los provavelmente traria um novo modelo ainda mais preciso e eficiente. Os fatores que sobressaíram neste estudo de caso foram a capacidade de criação de modelos individualmente precisos e uma coletividade diversa. Em outras palavras, o erro coletivo pode ser definido como a diferença entre a média de erros individuais e a diversidade. Este é o *Diversity Prediction Theorem*, que será devidamente explicitado na seção seguinte.

## 4. Diversity Prediction Theorem

O *Diversity Prediction Theorem* (DPT) tem como autor Scott Page (2007) e encontra-se no domínio dos estudos sobre o efeito da diversidade. Como mencionado no capítulo anterior, o DPT afirma que o erro coletivo na resolução de um problema é igual à diferença entre os erros individuais médios e a diversidade do grupo em questão. Esta seção tem por objetivo expor o teorema da forma mais intuitiva possível, mas antes de tal exposição, é necessário que se definam algumas condições para a aplicação efetiva do teorema em questão. Primeiramente, a discussão se propõe inicialmente a analisar grupos de pessoas, apesar de ser possível expandir seu uso para grupos de animais, máquinas e algoritmos. Além disso, as análises sobre capacidade coletiva e individual fazem referência a tarefas especificamente de predição.

Um ponto bastante significativo é que todos os indivíduos envolvidos no estudo de predições devem ser inteligentes, no sentido de terem mínima noção sobre a tarefa preditiva a que ele está inserido. Paralelamente, as tarefas preditivas precisam ser minimamente desafiadoras. Uma tarefa demasiadamente simples ou habitual (como a previsão do tempo de uma época específica do ano) não demanda nem diversidade nem indivíduos com especialidades. Ao mesmo tempo, se os indivíduos escolhidos para prever o peso de um avião forem indivíduos ainda na primeira série do primeiro grau, as respostas individuais muito provavelmente estarão muito discrepantes da realidade, não prevendo nada efetivamente.

A descrição deste teorema, como já mencionado anteriormente, capta o quanto um conjunto de modelos preditivos difere e quão precisos são esses modelos. Aqui nos referimos a um modelo como a estrutura lógica desenhada mentalmente por um preditor que o permite chegar efetivamente à sua previsão. Qualquer estimativa, antes de feita, passa por uma lista de atributos individuais que apresentam diversidade de interpretações a partir das informações que o agente possui. É de se esperar que a diversidade tenha ligação com os diferentes modelos possíveis que as pessoas seguem para chegar em uma mesma previsão ou estimativa.

Outro fator de importância para a aplicação deste Teorema é a suposição de independência. A comunicação entre indivíduos faz com que seus modelos e previsões apresentem vieses pois tendem a fazer os indivíduos mais inseguros (ou seja, menos precisos) a remodelar sua previsão de acordo com outrem. Infectar o modelo dos indivíduos gera um comportamento convergente, amenizando de fato a diversidade do coletivo.

Para previsões binárias, Page define Propriedade de Projeção, que afirma que no caso de duas pessoas analisarem atributos diferentes da mesma perspectiva, há a tendência de equilibrar-se a quantidade de acertos e erros. Em outras palavras, quando uma pessoa está correta, a outra tem mais probabilidade de estar errada, fazendo das duas previsões negativamente correlacionadas – ou seja, os erros se cancelam.

The Projection Property: If two people base their predictive models on different variables from the same perspective (formally, if they use nonoverlapping projection interpretations) then the correctness of their predictions is negatively correlated for binary predictions. (Page, 2007, chapter 8, page 203)

Esta propriedade implica que as multidões que analisam diversos atributos são mais sábias do que as que não o fazem. Em contraponto, a dimensionalidade da perspectiva a que a tarefa de previsão está inserida define o número de interpretações possíveis (e não sobrepostas). Este fator abre espaço para interpretações de grupo no caso de uso de atributos comuns por mais de uma pessoa – o que é uma conclusão aceitável, visto que a diversidade não é infinita.

As duas medidas mencionadas para descrever inicialmente este teorema – quão precisos são os modelos e quanto se diferem – são baseados na mesma medida de precisão: erros elevados ao quadrado. A importância deste fator é crucial para que

erros positivos e negativos não se anulam estatisticamente. Isso significa que um agente que tivesse a mesma probabilidade de superestimar e subestimar uma quantia, em média, não cometeria erros sem essa correção matemática.

Primeiramente vamos construir a lógica por trás da formulação do DPT a partir de um cenário de exemplo. Suponha um grupo de duas pessoas (A e B) que desenvolveram modelos para prever a classificação de três outras pessoas (pessoa 1, 2 e 3) em um concurso. Vamos supor as previsões demonstradas a seguir, assim como a média das duas previsões feitas e as reais classificações das 3 pessoas analisadas no concurso.

	<b>Pessoa A</b>	<b>Pessoa B</b>	<b>Média</b>	<b>Classificação</b>
<b>Pessoa 1</b>	6	10	8	6
<b>Pessoa 2</b>	3	7	5	5
<b>Pessoa 3</b>	5	1	3	1

PAGE, Scott (2007)

Conforme raciocínio descrito anteriormente, o erro individual de cada uma das pessoas A e B pode ser definido como o somatório do erro quadrado de cada previsão feita – erro aqui definido como a diferença entre sua estimação e a classificação posterior dos participantes 1, 2 e 3:

$$E_A = (6 - 6)^2 + (3 - 5)^2 + (5 - 1)^2 = 0 + 4 + 16 = 20$$

$$E_B = (10 - 6)^2 + (7 - 5)^2 + (1 - 1)^2 = 16 + 4 + 0 = 20$$

Dessa forma, definimos erro médio como a média dos erros individuais, que é 20. Com objetivos comparativos, podemos também calcular o Erro Coletivo, sendo a diferença aqui traçada pelas previsões médias do grupo (coluna 3 da tabela):

$$E_{Coletivo} = (8 - 6)^2 + (5 - 5)^2 + (3 - 1)^2 = 4 + 0 + 4 = 8$$

A intuição por trás deste simples cálculo é que a previsão coletiva deste sistema é mais precisa do que qualquer uma de suas previsões individuais. Podemos afirmar que a diversidade de suas previsões é a causadora disso, pelo simples fato de que quando uma previsão individual é superestimada, outra previsão tende a se comportar com forças subestimadas, atenuando o erro grupal resultante. Com a tentativa de quantificar o fator causador desta diferença de resultados, podemos calcular quanto suas previsões diferem. A estratégia usada por Page é mensurar o

somatório do erro quadrado de cada indivíduo com a média coletiva. Ou seja, quão distante está a previsão individual da previsão média do coletivo:

$$E_A^{diversidade} = (6 - 8)^2 + (3 - 5)^2 + (5 - 3)^2 = 4 + 4 + 4 = 12$$

$$E_B^{diversidade} = (10 - 8)^2 + (7 - 5)^2 + (1 - 3)^2 = 4 + 4 + 4 = 12$$

$$E_{diversidade} = \frac{12 + 12}{2} = 12$$

Dessa forma, definimos diversidade como a média destes valores, que é 12. Conseguimos provar matematicamente – e com um exemplo bastante simples – que o Erro coletivo (8) é igual à diferença entre o erro médio (20) e a diversidade (12). Um fator importante a ser mencionado é que esta relação quantitativa pode ser expandida para qualquer número de preditores, não necessariamente limitando-se a um grupo de 2 pessoas. Dessa forma, podemos enunciar o *Diversity Prediction Theorem*, dada uma multidão de modelos preditivos, como:

Erro coletivo = erro individual médio - diversidade de previsão

$$(c - \theta)^2 = \frac{1}{n} \sum_{i=1}^n (s_i - \theta)^2 - \frac{1}{n} \sum_{i=1}^n (s_i - c)^2$$

Onde a variável  $s_i$  é a previsão do indivíduo  $i$ ,  $c$  é a média das estimativas individuais e  $\theta$  é o número a ser estimado. O real significado deste teorema é que a capacidade individual e a diversidade coletiva contribuem igualmente para a capacidade preditiva coletiva. Grosso modo, ser diferente é tão importante quanto ser bom. O aumento em 1 unidade na diversidade apresenta o mesmo peso e influência no erro coletivo que a redução de 1 unidade na capacidade média do grupo.

Uma implicação interessante formulada por Page é a denominada *The Crowd beats the average law*, que afirma que dada qualquer coleção de diversos modelos preditivos, a previsão coletiva é mais precisa do que a média das previsões individuais. Isso significa que o erro de previsão coletiva é sempre menor do que o erro individual médio, ratificando que a multidão prediz melhor do que as pessoas que a integram.

Essa conclusão relativamente simples matematicamente traz à luz do estudo da Economia a ideia central de que a interação das partes gera de fato informação adicional ao sistema a que estão inseridos – premissa base de Sistemas Complexos Adaptáveis. Um contrassenso detectado com a abordagem Tradicional é que se as

peças individuais predizem perfeitamente, elas não podem ser diversas. Isso se dá pelo fato de que se o erro individual médio é nulo, a diversidade também deve ser nula.

Outro fator de importância para a aplicação deste Teorema é a suposição de independência. A comunicação entre indivíduos faz com que seus modelos e previsões apresentem vieses pois tendem a fazer os indivíduos mais inseguros (ou seja, menos precisos) a remodelar sua previsão de acordo com o outro. Infectar o modelo dos indivíduos gera um comportamento convergente, amenizando de fato a diversidade do coletivo.

Na seção seguinte será apresentado o teste empírico feito para a construção desta monografia, onde o *Diversity Prediction Theorem* será foco.

## 5. Teste

### 5.1. Metodologia

O teste empírico formulado para analisar o *Diversity Prediction Theorem* foi realizado via *Typeform* (ferramenta Google), pela construção de um questionário que coletou dois tipos de informação: o nível de especialização dos participantes e a capacidade preditiva de cada um deles. Este formulário foi enviado via internet para pessoas de diferentes idades, áreas de atuação, Universidades, Faculdades, Estados brasileiros e até outros países. A orientação dada no início do formulário foi de responder o questionário individualmente e somente uma (1) vez.

O nível de especialização foi observado pelas informações de nível de escolaridade e formação acadêmica, pressupondo que a combinação dessas duas variáveis indique um certo nível de diversidade cognitiva e de raciocínio por parte dos participantes. Espera-se que uma amostra com diferentes níveis de escolaridade e formação acadêmica assegure diversidade de modelos preditivos – e de atributos individuais – por parte do grupo populacional a que o experimento foi realizado.

A capacidade preditiva dos participantes foi observada a partir das respostas individuais sobre três perguntas com caráter preditivo simples, sendo elas apresentadas a seguir<sup>4</sup>.

- I) Estime a partir de seu raciocínio próprio a quantidade de grãos de feijão branco presentes na imagem abaixo.
- II) Estime a partir de seu raciocínio próprio quantos quilogramas (kg) tem o cachorro da imagem abaixo.
- III) Quantos gols você acredita que a seleção brasileira fará no seu primeiro jogo da Copa do Mundo 2018 (domingo, 17/06, contra a Suíça)?

As respostas a serem preditas foram  $\theta_1 = 278$  grãos de feijão,  $\theta_2 = 19\text{kg}$  e  $\theta_3 = 1$  gol do Brasil. O teste em si foi realizado com o total de 383 pessoas, resultando em um total de 378 respostas. Três pessoas foram excluídas do experimento por terem colocado respostas muito discrepantes da realidade, como a atribuição de 6.500kg e 12.300kg para o cachorro da foto e 100.000 grãos de feijão presentes na foto. A decisão de retirada se deu pela premissa do *Diversity Prediction Theorem* que pressupõe preditores e previsões minimamente coerentes com a realidade. As outras duas observações excluídas foram assim classificadas por terem sido enviadas com respostas vazias.

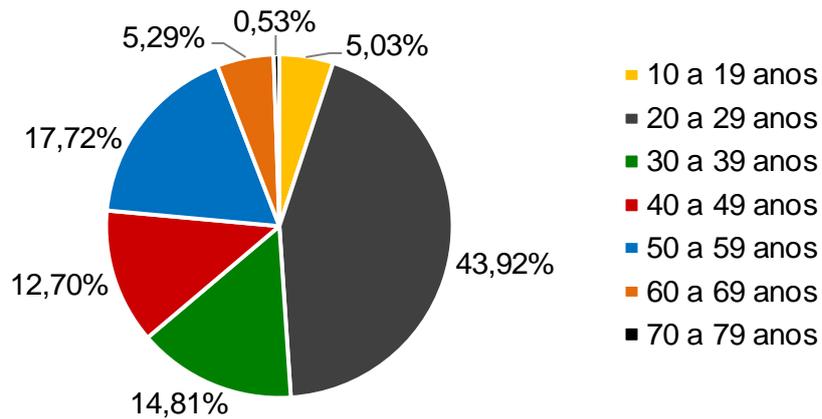
## 5.2. Teste empírico

Primeiramente fazendo uma análise demográfica sobre a amostra populacional do experimento, o total de 378 indivíduos centrou a idade no conjunto de [17,74] anos. A distribuição de idade apresentou os seguintes percentuais: 5,03% de pessoas entre 10 e 19 anos; 43,92% entre 20 e 29 anos; 14,81% entre 30 e 39 anos; 12,70% entre 40 e 49 anos; 17,72% entre 50 e 59 anos; 5,29 entre 60 e 69 anos e 0,53% entre 70 e 79 anos.

---

<sup>4</sup> As imagens referentes às perguntas se encontram no anexo ao final deste documento.

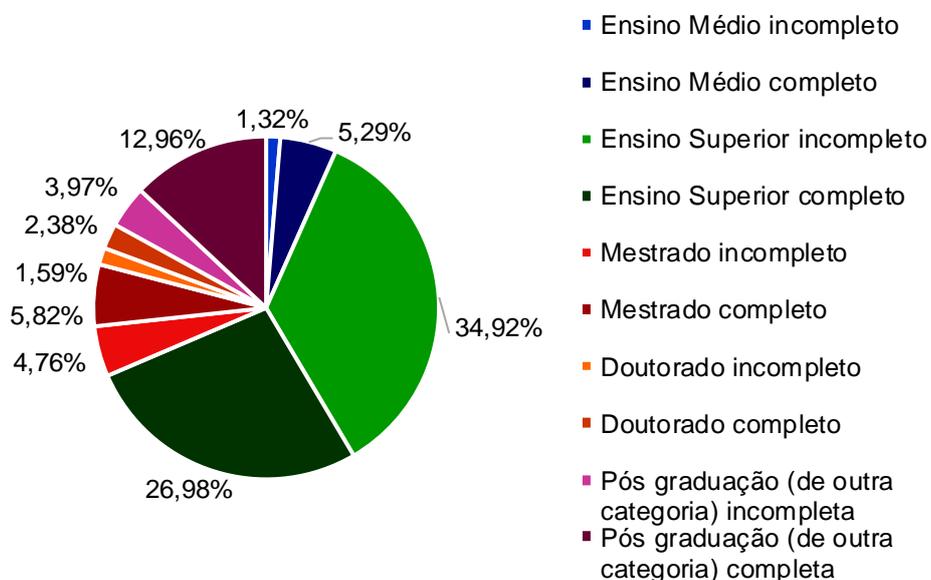
## Idade do Grupo



Elaborado pela autora

Além disso, a amostra mostrou um espectro significativo do nível de escolaridade dos indivíduos, tendo alguns o Ensino Médio ainda incompleto, outros Doutorado completo e, ainda, alguns cursos de especialização. Do total, 1,32% possui EM incompleto; 5,29% EM completo; 34,92% ES incompleto; 26,98% ES completo; 4,76% Mestrado incompleto; 5,82% Mestrado completo; 1,59% Doutorado incompleto; 2,38% Doutorado completo e, ainda, 3,97% de pessoas apresentaram pós-graduação de outra categoria incompleta e 12,96%, completa.

## Nível de Escolaridade do Grupo



Elaborado pela autora

O nível de especialização da amostra foi observado não somente ponderando o nível de escolaridade do grupo, mas unindo este dado com a formação acadêmica dos indivíduos. Do total de respostas, foram captadas 94 formações diferentes variando entre profissões com alto capital humano – como Medicina e Engenharia – e baixo capital humano – como Massoterapia e Secretariado Executivo. Apesar de quase 35% da amostra estar inclusa na categoria de Ensino Superior incompleto, foram observados 28 cursos diferentes dentro desses 34,92% da amostra, sugerindo um nível significativo de heterogeneidade.

Das previsões feitas pelo grupo, a média de cada uma das perguntas foi 251,01 grãos de feijão (com desvio padrão de 325,17), 18,92 kg (com desvio padrão de 13,54) e 2,19 gols do Brasil (com desvio padrão de 0,81). Os valores máximos observados foram de 4.000 grãos, 90 kg e 7 gols, sendo os valores mínimos observados todos nulos.

De acordo com o *Diversity Prediction Theorem*, o erro coletivo pode ser encontrado pelo cálculo do quadrado da diferença entre a previsão média do grupo ( $c$ ) e o valor real a ser estimado ( $\theta$ ), como demonstrado na tabela a seguir.

	<b>Previsões médias (c)</b>	<b>Valor (<math>\theta</math>)</b>	<b>Diferença (c – <math>\theta</math>)</b>	<b>Quadrado (c – <math>\theta</math>)<sup>2</sup></b>
<b>Grãos de feijão</b>	251,01	278	-26,99	728,29
<b>Quilogramas</b>	18,92	19	-0,08	0,01
<b>Gols do Brasil</b>	2,19	1	1,19	1,42

Elaborado pela autora

O erro individual médio e o nível de diversidade, conforme esclarecidas as suas formas de cálculo pelo teorema, formam mensuradas via Excel e os dados obtidos foram discriminados na tabela apresentada a seguir.

$$(c - \theta)^2 = \frac{1}{n} \sum_{i=1}^n (s_i - \theta)^2 - \frac{1}{n} \sum_{i=1}^n (s_i - c)^2$$

<b>FEIJÃO</b> (unidade em grãos)		<b>CACHORRO</b> (unidade em kg)		<b>BRASIL</b> (unidade em gols)	
Erro Coletivo	728,29	Erro Coletivo	0,01	Erro Coletivo	1,42
Erro Individual médio	106.187,25	Erro Individual médio	182,94	Erro Individual médio	2,08
Diversidade	105.458,97	Diversidade	182,94	Diversidade	0,66

Elaborado pela autora

## 6. Resultados e discussão

A análise demográfica realizada sugere que a diversidade cognitiva do grupo é relevante. A união de diferentes níveis de escolaridade e formações acadêmicas faz com que o grupo apresente diferentes modelos mentais para a resolução de problemas, fazendo uso de diferentes atributos e interpretações cognitivas. Essa é a premissa inicial para observar a sabedoria das multidões.

A segunda premissa é que as tarefas preditivas não sejam demasiadamente óbvias ou cotidianas, o que foi respeitado na escolha das perguntas feitas no teste empírico. Nenhuma das variáveis a serem previstas está distante da realidade brasileira, ou seja, todos os participantes com certeza têm uma noção sobre as características dimensionais de grãos de feijão, massa corpórea de um cachorro e da quantidade de gols que uma partida de futebol de campo costuma ter. Ao mesmo tempo, essas tarefas não são tão óbvias por não serem cotidianas ou básicas demais. Dessa maneira, o tipo de tarefa preditiva também respeitou as premissas necessárias para analisar a sabedoria das multidões.

O molde construído para a aplicação deste teste respeitou a simetria de informações, visto que a única informação cedida aos participantes foi de fato as fotografias mostradas ao final de cada pergunta – que foi absorvida pelos indivíduos a partir da sua própria capacidade cognitiva. No caso da previsão da quantidade de gols brasileiros, foi escolhido especificamente o resultado do primeiro jogo para evitar viés de opinião por ventura da percepção da performance do time brasileiro, uma vez já iniciada a Copa do Mundo. No caso de ter sido selecionado qualquer jogo que não o primeiro, os participantes poderiam ter usado a performance do primeiro jogo para embasar sua previsão. A escolha do primeiro jogo tem por objetivo reduzir o ruído da previsão.

De fato, como foi previsto pelo teorema, o erro coletivo foi, em todos os casos, menor do que o erro individual médio, tendo a diversidade do grupo de fato interferido no resultado grupal da previsão. Com exceção da previsão sobre a quantidade de gols, é possível observar que os desvios padrões indicam um grau elevado de dispersão sobre a média dos resultados. Isso ratifica o uso heterogêneo de atributos e modelos preditivos por parte dos participantes. Aprofundando ainda mais, essa diversidade com relativamente alta dispersão chegou a resultados muito próximos de  $\theta_i$  – indicando que os erros foram de fato amenizados quando agregados.

Dos resultados do teorema, podemos analisar que o erro individual médio da quantidade de grãos foi extremamente alto, um erro de 106.187,25 grãos sobre a quantidade real de 278 grãos presentes do pote. Nesta questão, a diversidade foi também tão alta (105.458,97), que o erro coletivo foi amenizado, deixando de estar em uma escala de milhar (728,29). Das três tarefas preditivas, essa tarefa foi a que apresentou maior ordem de grandeza no valor numérico a ser estimado e maior desvio padrão sobre as previsões individuais (dp = 325,17 grãos). Ao mesmo tempo, teve o erro coletivo muito suavizado, visto que o nível do erro individual médio e a diversidade estarem ambas em uma escala de milhar. Mesmo assim, foi a tarefa com menor precisão coletiva em seu resultado.

O erro individual médio de quilogramas atribuídos ao cachorro foi de 182,944 e a diversidade do grupo neste quesito foi incrivelmente de 182,938. A proximidade impressionante desses dois indicadores mostrou um erro coletivo de 0,007 (na tabela arredondado para 0,01) na previsão desta tarefa. É interessante analisar o desvio padrão das respostas dadas nesta tarefa, que foi de 13,54kg, e a resposta máxima e mínima dada, que foram respectivamente 90kg e 0kg. Esses dados sugerem que não houve uma convergência de respostas que geraram uma boa precisão, mas que a precisão do resultado se deu pelo nível de diversidade ter sido extremamente próxima numericamente às capacidades individuais.

Por fim, a última das tarefas preditivas mostrou um erro individual médio de 2,08 gols e uma diversidade de 0,66. Sendo a diversidade calculada no DPT com o uso da diferença entre as previsões individuais e a média das previsões individuais, vemos que o nível de heterogeneidade aqui foi o mais baixo. Talvez este fato seja atribuído à ordem de grandeza referente à esta tarefa, que variou somente entre 0 e 7. Ao mesmo tempo, aqui se aplica a ideia de dimensionalidade definida no capítulo

4, que age como limitador da diversidade. A dimensionalidade da perspectiva define a quantidade de interpretações possíveis na resolução de cada tarefa preditiva.

No caso da quantidade de gols, é de se esperar que o uso de atributos comuns pelo grupo amostral seja maior. Seja pela divulgação de estatísticas e de indicadores da área esportiva tanto sobre o time brasileiro como o time suíço – que foi feita nacionalmente e antes mesmo do início desta monografia – seja pelo momento em que foi realizada a pesquisa, na semana anterior do jogo em questão – momento em que as pessoas mais debatem sobre suas previsões do resultado do jogo.

Dessa forma, o erro coletivo captado na terceira tarefa foi de 1,42 gols, suavizando menos do que o visto nas outras tarefas, por representar cerca de 68% do erro individual médio. O erro coletivo de quilogramas representa um valor irrisório próximo de 0% do erro individual médio da segunda tarefa (0,0038%) e o erro coletivo de grãos representa um valor próximo de 0,69% do erro individual médio da primeira tarefa preditiva.

Uma observação interessante constatada na análise dos dados foi que nenhum indivíduo acertou a quantidade exata de grãos presente no pote de feijão apresentado, 2 pessoas acertaram a quantidade de quilos do cachorro e 44 pessoas previram o único gol do Brasil no seu primeiro jogo da copa.

O grupo de idade que teve menor e maior erro coletivo na previsão de grãos de feijão foi, respectivamente, 50-59 anos (799,12 grãos de erro) e 70-79 anos (com 26.569 grãos de erro). É importante ressaltar que o fato de terem participado da pesquisa somente duas pessoas neste intervalo de idade entre 70 e 79 anos, é de se esperar que seu erro seja maior. A segunda categoria de idade com maior erro coletivo foi 60-69 anos, com 20.334,76 grãos de erro e amostra de 20 pessoas. Com relação aos quilogramas do cachorro, o grupo com menor e maior erro coletivo foi 20-29 anos (com 0,13kg de erro) e 60-69 anos (com 104,04kg de erro), respectivamente. Referente aos gols do Brasil, a performance grupal com menor e maior erro coletivo foi 50-59 anos (com 0,83 gols de erro) e 70-79 anos (com 2,25 gols de erro), respectivamente.

	FEIJÃO (unidade em grãos)		CACHORRO (unidade em kg)		BRASIL (unidade em gols)	
10-19 anos	Erro Coletivo	4.474,91	Erro Coletivo	2,02	Erro Coletivo	1,73
	Erro Individual médio	57.623,95	Erro Individual médio	141,53	Erro Individual médio	2,37
	Diversidade	53.149,04	Diversidade	139,51	Diversidade	0,64
20-29 anos	Erro Coletivo	5.360,71	Erro Coletivo	<b>0,13</b>	Erro Coletivo	1,84
	Erro Individual médio	33.024,83	Erro Individual médio	169,78	Erro Individual médio	2,50
	Diversidade	27.664,12	Diversidade	169,65	Diversidade	0,66
30-39 anos	Erro Coletivo	2.704,00	Erro Coletivo	21,72	Erro Coletivo	1,23
	Erro Individual médio	33.789,50	Erro Individual médio	93,91	Erro Individual médio	1,68
	Diversidade	31.085,50	Diversidade	72,19	Diversidade	0,45
40-49 anos	Erro Coletivo	1.278,06	Erro Coletivo	2,78	Erro Coletivo	1,22
	Erro Individual médio	142.615,29	Erro Individual médio	141,54	Erro Individual médio	1,81
	Diversidade	141.337,23	Diversidade	138,76	Diversidade	0,59
50-59 anos	Erro Coletivo	<b>799,12</b>	Erro Coletivo	1,76	Erro Coletivo	<b>0,83</b>
	Erro Individual médio	149.358,63	Erro Individual médio	170,85	Erro Individual médio	1,45
	Diversidade	148.559,51	Diversidade	169,09	Diversidade	0,62
60-69 anos	Erro Coletivo	20.334,76	Erro Coletivo	<b>104,04</b>	Erro Coletivo	1,21
	Erro Individual médio	738.072,10	Erro Individual médio	724,50	Erro Individual médio	2,20
	Diversidade	717.737,34	Diversidade	620,46	Diversidade	0,99
70-79 anos	Erro Coletivo	<b>26.569,00</b>	Erro Coletivo	64,00	Erro Coletivo	<b>2,25</b>
	Erro Individual médio	27.794,00	Erro Individual médio	145,00	Erro Individual médio	2,50
	Diversidade	1.225,00	Diversidade	81,00	Diversidade	0,25

Elaborado pela autora

Com relação ao nível de escolaridade, o menor erro coletivo foi de 348,44 grãos pelo grupo de pessoas com Doutorado incompleto e o maior erro foi de 30.545,51 grãos pelo grupo de pessoas com Mestrado completo. O grupo que se aproximou mais da quantidade real de quilogramas do cachorro foi o grupo com Ensino Superior incompleto, com erro coletivo de 0,09kg. Paralelamente, o grupo que mais se distanciou foi o grupo com Ensino Médio incompleto, com 60,84kg de erro coletivo. Por fim, o menor erro coletivo referente à quantidade de gols brasileiros foi atribuído ao grupo de pessoas com Doutorado completo, com erro de 0,31 gols. Ao mesmo tempo, o maior erro coletivo nessa estimativa de gols foi referente ao grupo com Doutorado incompleto, com o erro de 2,78 gols.

	FEIJÃO		CACHORRO		BRASIL	
	(unidade em grãos)		(unidade em kg)		(unidade em gols)	
EM incompleto	Erro Coletivo	10.485,76	Erro Coletivo	<b>60,84</b>	Erro Coletivo	1,96
	Erro Individual médio	48.725,20	Erro Individual médio	112,60	Erro Individual médio	2,20
	Diversidade	38.239,44	Diversidade	51,76	Diversidade	0,24
EM completo	Erro Coletivo	12.645,00	Erro Coletivo	44,89	Erro Coletivo	1,69
	Erro Individual médio	29.813,75	Erro Individual médio	512,90	Erro Individual médio	2,20
	Diversidade	17.168,75	Diversidade	468,01	Diversidade	0,51
ES incompleto	Erro Coletivo	3.364,88	Erro Coletivo	<b>0,09</b>	Erro Coletivo	1,86
	Erro Individual médio	39.345,80	Erro Individual médio	139,55	Erro Individual médio	2,53
	Diversidade	35.980,92	Diversidade	139,45	Diversidade	0,67
ES completo	Erro Coletivo	1.127,51	Erro Coletivo	1,53	Erro Coletivo	1,14
	Erro Individual médio	60.010,17	Erro Individual médio	218,31	Erro Individual médio	1,83
	Diversidade	58.882,66	Diversidade	216,79	Diversidade	0,69
Mestrado incompleto	Erro Coletivo	1.140,94	Erro Coletivo	1,78	Erro Coletivo	1,49
	Erro Individual médio	148.937,11	Erro Individual médio	175,11	Erro Individual médio	2,11
	Diversidade	147.796,17	Diversidade	173,33	Diversidade	0,62
Mestrado completo	Erro Coletivo	<b>30.545,51</b>	Erro Coletivo	10,12	Erro Coletivo	1,09
	Erro Individual médio	408.298,14	Erro Individual médio	280,09	Erro Individual médio	1,59
	Diversidade	377.752,63	Diversidade	269,97	Diversidade	0,50
Doutorado incompleto	Erro Coletivo	<b>348,44</b>	Erro Coletivo	14,69	Erro Coletivo	<b>2,78</b>
	Erro Individual médio	38.652,33	Erro Individual médio	263,83	Erro Individual médio	3,00
	Diversidade	38.303,89	Diversidade	249,14	Diversidade	0,22
Doutorado completo	Erro Coletivo	2.988,44	Erro Coletivo	1,49	Erro Coletivo	<b>0,31</b>
	Erro Individual médio	19.632,89	Erro Individual médio	73,89	Erro Individual médio	0,56
	Diversidade	16.644,44	Diversidade	72,40	Diversidade	0,25
Pós graduação outra categoria incompleto	Erro Coletivo	6.400,00	Erro Coletivo	0,28	Erro Coletivo	1,44
	Erro Individual médio	22.465,60	Erro Individual médio	76,93	Erro Individual médio	1,87
	Diversidade	16.065,60	Diversidade	76,65	Diversidade	0,43
Pós graduação outra categoria completo	Erro Coletivo	410,68	Erro Coletivo	3,53	Erro Coletivo	1,08
	Erro Individual médio	317.860,76	Erro Individual médio	100,57	Erro Individual médio	1,78
	Diversidade	317.450,07	Diversidade	97,05	Diversidade	0,69

Elaborado pela autora

Essa análise dos erros coletivos por categoria nos mostra que não houve um padrão claro sobre quais idades e níveis de escolaridade são mais precisos. Um padrão visto foi que em cada uma das três perguntas-teste, pessoas mais novas apresentaram erros menores, assim como pessoas com maiores níveis de escolaridade.

## 7. Conclusões

A primeira e primordial conclusão da presente monografia é o reconhecimento da importância da aplicação de Sistemas Complexos Adaptáveis em Ciências Sociais no geral. A Ciência Econômica permanece elaborando modelos e intuições embasadas em raciocínios formulados no século passado, em uma era completamente diferente da atual. Muitas contribuições inovadoras estão sendo reconhecidas globalmente pela área da Psicologia e de Sistemas Complexos e elas devem ser cada vez mais fomentadas.

A importância da diversidade, muitas vezes tratada como um fator mais social do que econômico, influencia na resolução de problemas em mesma relevância que o nível de capacidades e habilidades individuais. Essa é uma conclusão com forte poder intuitivo e funcional, tendo sido reconhecida tanto para situações em que a performance coletiva foi analisada com o uso da média dos resultados individuais (como o teste realizado neste documento, variável c) como para situações em que foi unido presencialmente um grupo para resolver alguma tarefa (como no artigo *Are Two Heads Better than One? Monetary Policy by Committee*).

Apesar de precisar de uma elaboração mais rebuscada para a sua implementação, uma extensão futura da presente pesquisa é testar presencialmente a resolução de problemas a partir de composições diferentes de grupos. Assim poderíamos comparar a performance de grupos homogêneos e heterogêneos (grupos com mesma profissão, faixa de idade, nível de escolaridade e sexo contra grupos com diversidade) ou ainda grupos com pessoas conhecidas e grupos com somente uma pessoa desconhecida. A dualidade entre grupos homogêneos e heterogêneos poderia trazer à luz da pesquisa a dedução se grupos superam de fato especialistas, conforme é argumentado por Scott Page.

## 8. Anexo imagens referentes ao teste empírico



## 9. Bibliografia

BEINHOCKER, Eric. *Origins of Wealth*. First Edition. EUA, Harvard Business School Press, 2007

BLINDER, Alan and MORGAN, John. Are Two Heads Better than One? Monetary Policy by Committee. *Journal of Money, Credit, & Banking* 37, no. 5: 789–811. 2005

DARWIN, Charles. *On the Origin of Species*. Inglaterra, Down, Bromley, Kent. 1859.

GLADWELL, Malcolm. *The Tipping Point: How Little Things Can Make a Big Difference*. First Edition. EUA, Little Brown and Company, 2000

LINDGREN, Kristian. Evolutionary Phenomena in Simple Dynamics. Artificial Life II, SFI Studies in the Sciences of Complexity, Vol. X, edited by C. G. Langton, C. Taylor, J. D. Farmer, & S. Rasmussen, Addison-Wesley, 1991

MILLER, John and PAGE, Scott. The Standing Ovation Problem. Wiley Periodicals, Inc., Vol. 9, No.5. 2004

MITCHELL, Melanie. Complexity: A Guided Tour. Inglaterra, Oxford University Press Inc, 2009

PAGE, Scott. The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies. EUA, Princeton University Press, 2007

PAGE, Scott. The Hidden Factor: Why Thinking Differently Is Your Greatest Asset (Course Guidebook), Lecture 5: The Diversity Prediction Theorem Times Three, 2012

RIDLEY, Matt. The Evolution of Everything: How Ideas Emerge. First U.S Edition. Inglaterra, HarperCollins Publishers. 2015

TVERSKY, Amos e KAHNEMAN, Daniel. Judgment under uncertainty: Heuristics and Biases. *Science*, New Series, Vol. 185, No. 4157. (Sep.27, 1974), 1974