

O papel do reconhecimento do acaso no raciocínio indutivo (*)

MÁRIO BOTO FERREIRA (**)
LEONEL GARCIA-MARQUES (**)

Uma das mais importantes actividades de resolução de problemas em que as pessoas normalmente se envolvem é o raciocínio indutivo. Ou seja, chegar a proposições gerais a partir de proposições particulares.

Genericamente, considera-se que o raciocínio indutivo, para ser correcto, deve satisfazer certos princípios estatísticos. Por exemplo, devemos ter mais confiança em generalizações baseadas num grande número de instâncias (desde que estas instâncias constituam uma amostra representativa da população) do que em generalizações baseadas num número pequeno de instâncias.

Dado o carácter central do raciocínio indutivo no quotidiano de todos nós, é no mínimo perturbador verificar que muita da investigação psicológica neste domínio indica que as pessoas, face a tarefas de escolha ou tomada de decisão em situações de incerteza, usam heurísticas (processos simplificados de decisão) que não satisfazem os

princípios estatísticos requeridos. De facto, a linha de investigação iniciada por Tversky e Kahneman (1971, 1973, 1974; para uma revisão ver Einhorn & Hogarth, 1981; Kahneman, Slovic, & Tversky, 1982; Nisbett & Ross, 1980; Sherman & Corty, 1984), mostra que as heurísticas correspondem a processos de decisão altamente eficientes que frequentemente levam a respostas que se conformam com princípios estatísticos. Mas, noutros circunstâncias, levam a erros e enviesamentos. Veja-se o caso da heurística da representatividade (a mais estudada das heurísticas propostas por Tversky e Kahneman). De acordo com esta heurística, a probabilidade de uma instância A pertencer a uma categoria B é tanto maior quanto maior for a semelhança entre A e B. Em certas circunstâncias, semelhança e probabilidade estatística estão positivamente correlacionadas. Sempre que tal aconteça a heurística da representatividade leva a respostas que coincidem com a resposta estatística. Por outro lado, esta heurística negligencia de forma consistente a informação estatística relevante para a resolução adequada de problemas indutivos (por exemplo, o tamanho das amostras, os *Base-Rates*, etc.), o que leva a erros e enviesamentos sistemáticos e característicos.

No início dos anos 80, Nisbett, Krantz, Jepson e Kunda (1983), embora reconhecendo que as diversas heurísticas identificadas por Tversky e

(*) Este artigo baseia-se na dissertação de Mestrado em Psicologia Cognitiva do primeiro autor sob orientação do segundo autor. Qualquer questão relativa ao artigo deve ser dirigida mabf@fpce.ul.pt

(**) Faculdade de Psicologia e de Ciências da Educação da Universidade de Lisboa.

Kahneman têm um importante papel no raciocínio indutivo, argumentaram que por vezes as pessoas baseiam as suas inferências em intuições que respeitam princípios estatísticos. Ou seja, paralelamente às *heurísticas não estatísticas* descritas por Kahneman e Tversky, Nisbett e colaboradores consideram a existência de *heurísticas estatísticas*. As heurísticas estatísticas podem igualmente considerar-se regras simplificadas de tomar decisões de elevada eficiência que, no entanto, revelam intuição estatística. Por exemplo, o ditado «não se deve julgar um livro pela sua capa» é eventualmente revelador da intuição estatística de que julgamentos baseados em pequenas amostras (a informação veiculada pela capa) são pouco fidedignas para concluir sobre uma população (a informação veiculada no livro).

De facto, durante a década de 80, Nisbett e colaboradores desenvolveram um esforço de investigação sistemático não só no sentido de mostrar que as pessoas por vezes raciocinam recorrendo a princípios estatísticos, mas sobretudo com o objectivo de determinar em que circunstâncias as pessoas demonstram sensibilidade a variáveis de natureza estatística; até que ponto será possível promover o uso de heurísticas estatísticas através do treino formal; em que domínios é que o raciocínio estatístico é mais ou menos provável de ocorrer; que factores facilitam ou dificultam o raciocínio indutivo baseado em princípios estatísticos; etc. (ver Nisbett, 1993, para uma revisão).

No que diz respeito aos factores influentes na qualidade (estatística ou não estatística) do raciocínio indutivo humano, é importante notar que a interpretação e resolução de qualquer situação que envolva acaso ou incerteza (como é tipicamente o caso nos problemas indutivos) implica um raciocínio baseado em princípios estatísticos. Assim, pode pensar-se que os erros e enviesamentos de julgamentos caracterizados por Tversky, Kahneman e outros psicólogos (ver Gilovich, Griffin & Kahneman, 2002; Nisbett & Ross, 1980; Sherman & Corty, 1984; Tversky & Kahneman, 1974; Kahneman, Slovic & Tversky, 1982) passam pela dificuldade manifesta das pessoas em reconhecer a componente aleatória envolvida nos vários domínios sociais do quotidiano, comprometendo-se assim o uso das heurísticas estatísticas.

O principal objectivo do presente trabalho é justamente uma tentativa de verificar até que ponto tarefas de facilitação do reconhecimento da componente aleatória de eventos associados a diversos domínios sociais levam a uma melhoria do desempenho dos sujeitos, na resolução de problemas indutivos.

Nisbett, Krantz, Jepson e Kunda (1983) consideram que a uniformidade do comportamento dos sujeitos em experiências sobre raciocínio indutivo realizadas até ao fim dos anos 70, em que investigadores como Tversky e Kahneman (1971, 1973, 1974; Kahneman & Tversky, 1972) não encontram qualquer evidência de aplicação de princípios com a Lei dos grandes Números, Regressão à Média ou consideração de «Base-Rates», não implica que o ser humano não possa raciocinar de acordo com (alguns destes) princípios estatísticos. O que parece, em grande medida, estar subjacente à proliferação do uso das heurísticas não estatísticas são factores como a falta de clareza do «espaço» da amostra e do processo de obtenção desta amostra (que constitui a informação de partida do sujeito). Diz-se que o «espaço» de uma amostra é claro quando é fácil perceber quais são as possibilidades de um ensaio ou acontecimento. Por exemplo, o «espaço» da amostra de lançamentos de um dado é óbvio, pois um dado tem seis faces, ou seja seis resultados ou acontecimentos possíveis. Por outro lado, o dado pode ser lançado repetidamente, o que torna fácil para as pessoas conceptualizarem a observação de um conjunto de lançamentos do dado como uma amostra duma população (infinita) de lançamentos. Ou seja, o processo de obtenção da amostra é igualmente claro. Em domínios sociais, o espaço da amostra e processo de obtenção da mesma são frequentemente obscuros e difíceis de conceptualizar. É o caso da «ilusão da entrevista» (Tversky & Kahneman, 1974). Na área do recrutamento e selecção de pessoal, diversos técnicos (muitos deles psicólogos) desenvolvem uma crença claramente enviesada relativamente à validade de uma entrevista como prova de selecção (eventualmente baseada na heurística da representatividade). Esta ilusão existe porque as pessoas têm dificuldade em reconhecer os resultados de uma entrevista como uma pequena amostra do comportamento do entrevistado. Esta informação é, na maioria das ve-

zes, pouco credível para extrair inferências sobre a sua personalidade, motivação, etc. (dada a componente de erro associada à variabilidade do comportamento humano). Em vez disto e de acordo com o uso da heurística da representatividade, muitos técnicos consideram os resultados da entrevista como um «esboço» ou «holograma» do entrevistado que leva a uma confiança excessiva nesta técnica de selecção.

Um segundo aspecto que concorre para o uso mais ou menos generalizado das heurísticas não estatísticas prende-se com o reconhecimento da operação de factores de acaso. Este reconhecimento está especialmente facilitado quando nos referimos ao funcionamento de um qualquer aparelho que produza eventos aleatórios. É o caso da roleta, dos números do totoloto, etc. No entanto, é provável que na maioria das actividades sociais, seja mais difícil reconhecer a presença de uma componente de variabilidade não explicável através duma análise causal. A insensibilidade ao fenómeno de regressão à média em diversos domínios sociais é um bom exemplo da dificuldade em reconhecer que existe uma componente de incerteza em muitas das circunstâncias do nosso quotidiano. Por exemplo, após a obtenção de um resultado muito elevado num teste escolar (acima da média dos resultados normalmente obtidos pelo estudante em causa) é mais provável a obtenção de resultados futuros que se aproximem ou regridam para a média (do desempenho do aluno), do que a obtenção de resultados igualmente ou mais elevados. Este é um fenómeno estatístico que se prende justamente com variações de resultados devido a factores imprevisíveis (por exemplo, uma coincidência feliz entre o que o aluno estudou e o que saía para o teste, sorte nas respostas de escolha múltipla, etc.). No entanto, dada a insensibilidade que muitas vezes manifestam aos factores de acaso, as pessoas tendem considerar quaisquer resultados altamente preditivos dos futuros resultados (independentemente de serem mais ou menos extremos). Por isto, a regressão à média é normalmente interpretada como uma mudança (no desempenho dos alunos) que requer uma explicação determinística.

Em suma, a universalidade dos enviesamentos inferenciais anteriormente obtidos parece estar relacionada com o facto de os problemas apresentados aos sujeitos serem, em geral, particular-

mente difíceis (problemas em relação aos quais os próprios investigadores por vezes sentem dificuldades) no sentido em que é difícil reconhecer a componente de acaso envolvida.

Curiosamente, a abordagem das heurísticas e enviesamentos originalmente proposta por Tversky e Khaneman (1974) nunca considerou a «dificuldade» dos problemas indutivos como uma condição-limite de aplicação das *heurísticas não estatísticas*. De acordo com os autores, estas heurísticas resumem a natureza do raciocínio indutivo humano e portanto deveriam ser igualmente usadas tanto em problemas «difíceis» como em problemas mais «fáceis». Ora tal não se verifica. Usando problemas com diferentes graus de dificuldade (com «espaços» de amostragem mais claros e maior saliência da componente de acaso), Nisbett, Krantz, Jepson e Kunda (1983) verificaram que o uso de conceitos estatísticos (mesmo na ausência de treino formal em estatística): a) não é nem raro, nem universal (a sua frequência varia em função do grau de dificuldade do problema); b) varia consistentemente entre indivíduos e; c) está correlacionado com outras aptidões mentais.

Ao admitir que as pessoas possuem alguma intuição para princípios estatísticos admite-se também a possibilidade de melhorar o desempenho das pessoas (no que diz respeito ao uso eficaz destes princípios estatísticos) através de procedimentos relativamente formais de treino cujos efeitos se manifestam ao nível da representação abstracta que as pessoas possuem destes princípios estatísticos. Fong, Krantz e Nisbett (1986), mostraram que o treino em estatística (treino formal e/ou resolução de problemas referentes a certos domínios do quotidiano – ex. *desporto*) influencia fortemente a maneira como as pessoas raciocinam sobre acontecimentos do dia-a-dia que envolvam incerteza, e que este treino se transfere para outros domínios para lá do domínio do treino. No entanto, estes resultados referem-se sempre a experiências em que a resposta aos problemas ocorre logo após a sessão de treino. Em tais circunstâncias não é possível saber se os efeitos positivos das sessões de treino se reflectem a nível da representação abstracta de certos princípios estatísticos (hipótese formalista), ou se para chegar às respostas dadas, os participantes recorreram meramente a analogias directas entre os problemas apresentados na fase

de treino e os problemas-teste (hipótese antiformalista). De forma a ultrapassar esta dificuldade, Fong e Nisbett (1991) realizaram experiências em que novamente usam sessões de treino após as quais os sujeitos respondem a problemas referentes ao mesmo ou a um domínio diferente daquele usado nas sessões de treino. Porém, desta vez alguns participantes só foram testados duas semanas após o treino. Este intervalo de tempo foi estabelecido de forma a que se perdesse a memória dos detalhes dos problemas apresentados nas sessões de treino, e assim poder contrastar a hipótese formalista com a hipótese antiformalista. Os resultados obtidos indicam que os efeitos do treino estatístico mantêm-se, mesmo no caso daqueles participantes que foram testados duas semanas após o treino, e mesmo quando os problemas de treino são de um domínio diferente do domínio dos problemas-teste (Fong & Nisbett, 1991). Ou seja, parece haver uma melhoria do desempenho dos participantes que transcende a especificidade de um dado domínio. Este padrão de resultados apoia claramente a posição formalista.

Mas o padrão de resultados desta experiência apresenta um outro aspecto de grande interesse. De facto, quando os problemas de treino são do mesmo domínio do dos problemas-teste, observa-se uma maior retenção do efeito do treino do que quando os problemas de treino e os problemas-teste são de domínios diferentes. De acordo com Fong e Nisbett (1991), uma possível explicação para este resultado baseia-se na hipótese de que os problemas de treino sirvam, não só para activar e melhorar o sistema de regras inferenciais do sujeito, mas também para providenciar aos sujeitos regras de como codificar certos eventos em termos do princípio estatístico em causa. Por exemplo, uma pessoa a quem foram dados problemas de treino sobre *desporto* pode aprender a codificar os lançamentos convertidos de um jogador de Basquetebol como uma amostra representativa da capacidade de desempenho desse jogador. Na resposta a problemas-teste do mesmo domínio, esta pessoa tenderá a usar informação semelhante (e.g. golos concretizados por um jogador de Andebol) como amostra representativa de um dado evento sobre o qual deve realizar algum tipo de inferência. Esta codificação da informação vai facilitar o recurso a heurísticas estatísticas (e.g. representações intuitivas da *Lei dos Grandes Números*).

ou seja, Estas *regras de codificação* servem para ligar as heurísticas estatísticas aos domínios em que estas vão ser aplicadas. Depois de um intervalo de duas semanas, os sujeitos que respondem a problemas-teste cujo domínio é coincidente com o dos problemas-treino, dispõem não só das representações intuitivas de certos princípios estatísticos, mas também de *regras de codificação* mais específicas que ajudam os sujeitos a aceder às heurísticas estatísticas apropriadas e a aplicar estas às situações concretas de cada problema¹.

Resumindo, para Fong e Nisbett (1991), não se trata tanto de possuir, a nível intuitivo, uma representação abstracta do princípio estatístico adequado, mas sobretudo de uma questão de codificação do problema em causa. Ou seja, do reconhecimento por parte do sujeito de que se trata de uma circunstância cuja resolução adequada implica a aplicação de alguma heurística estatística. De um ponto de vista conceptual, Esta hipótese explicativa carece porém de clarificação conceptual e suporte empírico.

Conceptualmente, o conceito de *regras de codificação* aparece definido de forma um pouco genérica. Segundo Fong e Nisbett, tratam-se de ideias (decorrentes do treino com problemas num dado domínio) de como codificar acontecimentos em termos de algum princípio estatístico. Embora estas regras sejam específicas de cada domínio, todas elas parecem ter um denominador comum: a interpretação de qualquer aconte-

¹ Uma outra interpretação possível deste padrão de resultados sugere que a retenção dos efeitos do treino após o intervalo de duas semanas resultaria do facto dos sujeitos se lembrarem dos problemas-exemplo com detalhe suficiente para resolver os problemas-teste por analogia. O melhor desempenho para os problemas-teste no mesmo domínio ocorreria porque estes problemas lembram mais os sujeitos dos problemas-exemplo do que os problemas-teste do domínio diferente. De forma a avaliar a plausibilidade desta interpretação alternativa, Fong e Nisbett (1991) testaram a memória dos sujeitos para os problemas de treino após o intervalo de duas semanas. A memória para os detalhes dos problemas de treino revelou ser extremamente pobre e não correlacionada com o desempenho.

cimento ou a resolução de qualquer situação que possua uma componente aleatória envolve a consideração de princípios estatísticos. Assim, independentemente das características específicas que estas *ideias de como codificar acontecimentos* têm em função do domínio a que nos referimos, pode pensar-se que a ligação entre representações intuitivas de princípios estatísticos e domínios específicos passa pelo reconhecimento da existência duma componente aleatória (noção de acaso) nos problemas apresentados ao sujeito. Ou seja, aquilo que é abstraído pelas pessoas é justamente a noção de um factor de variabilidade aleatório que, por sua vez, facilita uma codificação adequada da informação disponível para o uso de heurísticas estatísticas. Esta é a hipótese geral do presente trabalho que pode ser enunciada da seguinte maneira, «Se é através do reconhecimento da componente de acaso que as *regras de codificação* (Fong & Nisbett, 1991) funcionam, então condições experimentais que aumentem a acessibilidade da componente de acaso levarão a uma codificação apropriada das instâncias relevantes dos problemas e terão como efeito o aumento das respostas de natureza estatística».

O objectivo central da investigação aqui apresentada é justamente verificar até que ponto tarefas de facilitação do reconhecimento da componente aleatória de eventos associados a diversos domínios sociais levam a uma melhoria do desempenho dos sujeitos na resolução subsequente de problemas cujo domínio pode coincidir ou não com aqueles para os quais as tarefas acima referidas foram criadas.

No estudo de Fong e Nisbett (1991), usam-se dois domínios de actividade: *desporto* e *testes de aptidão*, e apenas um princípio estatístico a *Lei dos Grandes Números*². Na presente investigação, a construção do material (problemas ou situações dilemáticas do quotidiano) leva em conta quatro domínios: *Desporto*, *Escola*, *Saúde*, e *Fidelidade* (fidelidade conjugal), e quatro prin-

cípios estatísticos: *Lei dos Grandes Números*, *Base-Rates*, *Regressão à Média* e *Diagnosticidade*.

Os domínios de conteúdo foram escolhidos de forma a obter problemas cujo o grau de «dificuldade» fosse relativamente abrangente. Os eventos associados ao domínio de conteúdo *Desporto* (por exemplo, o desempenho de um jogador ou de uma equipa) são, em geral, fáceis de conceptualizar em termos estatísticos. Quer por razões objectivas (maior clareza do espaço de amostragem e do processo de obtenção da amostra), quer por aquilo a que Nisbett et al. (1983) referiram como «prescrições culturais». Ou seja, de um ponto de vista cultural, domínios de conteúdo como o *Desporto* são tradicionalmente tratados recorrendo a uma linguagem mais estatística e probabilística (por exemplo, percentagem de primeiros serviços no Ténis, «goal average» no Futebol, etc.). Eventos envolvidos no domínio de conteúdo *Saúde* são também frequentemente tratados de formas que implicam o reconhecimento da relevância do papel do acaso e, consequentemente, do raciocínio estatístico (por exemplo, opção por uma cirurgia em função da sua probabilidade de sucesso, frequência relativa de pessoas afectadas por um surto de gripe, etc.), embora eventualmente menos do que no domínio *desporto*. No caso do domínio de conteúdo *Escola*, embora diversos elementos facilitadores do papel do acaso estejam também presentes (por exemplo, a percentagem de reprovação numa disciplina) existe, por outro lado, uma cultura de mérito que tende a equacionar os resultados dos estudantes com aspectos ligados ao esforço e aptidão intelectual. Assim, há muitas vezes uma tendência de «responsabilização» do estudante por quaisquer resultados que este obtenha, negligenciando frequentemente a componente de acaso envolvida no desempenho³. Por fim, a *Fidelidade* é claramente um domínio de conteúdo cujos eventos apresentam um espaço de amostragem particularmente difícil de discernir

² Alguns dos problemas apresentados por Fong e Nisbett (1991) implicam a consideração do princípio de *Regressão à Média*. No entanto os autores consideram que este princípio é uma outra «manifestação» da *Lei dos Grandes Números*.

³ O fenómeno da insensibilidade da regressão à média (anteriormente discutido) é frequente em domínios que envolvem avaliação de desempenho como o escolar. Para um outro exemplo ligado ao treino de pilotos de aviação ver Tversky e Khaneman (1974).

e, relativamente ao qual as pessoas tendem, na nossa cultura, a desenvolver explicações de natureza mais determinística. Ou seja, raramente uma pessoa abordará questões de fidelidade conjugal em termos estatísticos, como por exemplo, usar o valor médio de comportamentos reveladores de infidelidade no passado como indicador da probabilidade de infidelidade futura. Pelo contrário, há uma maior tendência para abordar o tema da infidelidade em termos de factores individuais, variáveis de personalidade, motivos idiossincráticos, etc.

Com o uso de quatro princípios normativos, pretende-se, uma avaliação mais abrangente do uso das heurísticas estatísticas e eventuais *regras de codificação* por parte dos sujeitos, uma vez que, como referido, nos estudos de Nisbett e colaboradores (ver Nisbett, 1993) usou-se como princípio normativo apenas a *Lei dos Grandes Números*.

Desenvolveram-se também duas tarefas de facilitação do reconhecimento da componente aleatória de eventos e situações associados a certos domínios:

a) A primeira tarefa correspondeu à avaliação, por parte dos sujeitos, do grau de imprevisibilidade de diversos factores que afectam o desenrolar de certos eventos associados a um dado domínio. Por exemplo para um domínio como *desporto*, factores como «o apoio do público», «as condições climáticas», «ressaltos casuais da bola», etc. são avaliados quanto ao seu grau de imprevisibilidade, numa escala de sete pontos. De acordo com a hipótese levantada acima, esta *manipulação de reconhecimento da imprevisibilidade* (R) seria suficiente para reconhecer a existência duma componente aleatória num dado domínio de conteúdo, facilitando posteriormente a aplicação das heurísticas estatísticas a problemas ou situações que ocorrem dentro deste domínio.

b) A segunda tarefa correspondeu a um problema analógico. Assim, foram apresentados, lado a lado, dois pequenos textos. Um deles apresenta uma situação onde a fonte de variabilidade dos eventos é claramente aleatória probabilística (ex. lançar uma moeda ao ar). O outro é uma história de natureza social. Após a leitura de ambas as histórias é pedido aos sujeitos que façam a correspondência entre duas listas de diversos aspectos de um e outro texto apresentados lado a lado. O objectivo desta tarefa é levar os sujeitos

a reconhecer a existência de uma componente aleatória na história de natureza social por analogia com a história de natureza probabilística. Note-se que esta manipulação por analogia (A), não só torna a componente de acaso mais saliente como fornece aos participantes «pistas» de como representar estatisticamente a informação dos problemas de teste (através da correspondência estabelecida entre espaços de amostragem, fontes de acaso, etc.). Neste sentido, dada a natureza mais específica da *manipulação por analogia* (A) quando comparada com a *manipulação de reconhecimento da imprevisibilidade* (R), pode acontecer que A tenha maior impacto na promoção e actualização da intuição estatística dos participantes do que R.

Considerou-se ainda uma tarefa em que os sujeitos devem simplesmente ler definições dos quatro princípios normativos envolvidos na construção dos problemas (*manipulação definições – Dfs*). Cada uma destas definições é seguida de um pequeno exemplo. Esta tarefa não decorre directamente da hipótese central do trabalho, tratando-se de uma tarefa muito semelhante a uma manipulação realizada por Fong, Krantz e Nisbett (1986), a que os autores chamaram «demand condition» e onde os sujeitos recebiam também uma pequena definição da *Lei dos Grandes Números* mais um breve exemplo de aplicação desta lei⁴.

1. MÉTODO

1.1. Participantes

Os 99 sujeitos que participaram no presente

⁴ Esta condição funciona, quer no estudo de Fong, Krantz e Nisbett (1986), quer no presente estudo, como uma forma de verificar se os efeitos das manipulações se devem meramente ao facto de tornar saliente certas regras estatísticas. Há, no entanto, uma diferença importante. No caso do estudo destes autores, as manipulações consistiam em tarefas de treino formal e/ou através de exemplos. No caso presente as manipulações consistem em tarefas que procuram meramente facilitar o reconhecimento duma componente aleatória, não envolvendo qualquer tipo de treino.

estudo foram 52 estudantes universitários do 5.º ano dos cursos de Línguas e Literatura Germânicas e Românicas da Faculdade de Letras da Universidade de Lisboa, 17 estudantes do 5.º ano do curso de Filosofia, e 30 estudantes do 1.º ano do curso de Ciências do Instituto Superior de Educação e Ciências. Os sujeitos, na sua esmagadora maioria, nunca frequentaram quaisquer cursos de estatística, sendo que grande parte teve o seu último contacto formal com a matemática no 9.º ano de escolaridade. A participação dos sujeitos neste estudo foi voluntária e facultativa, tendo as aplicações do material decorrido no horário da disciplina de Psicologia Educacional.

1.2. *Delineamento experimental*

Optou-se por um design factorial 4 (princípios normativos) x 4 (domínios de conteúdo) x 4 (condições: R, A, D, e Controlo). Como variáveis dependentes tem-se as respostas dos participantes em dois momentos: logo após as manipulações e duas semanas após as manipulações. Para todos os problemas excepto os que envolvem Diagnosticidade, os participantes responderam através de uma escala de avaliação de 11 pontos. O extremo superior da escala (11) traduz a opção por uma resposta marcadamente estatística, e o extremo inferior da escala (1), traduz a opção por uma resposta marcadamente determinística (não estatística). No caso dos problemas envolvendo Diagnosticidade, a variável dependente é uma escolha forçada entre três opções de resposta.

1.3. *Material de Teste*

Os participantes responderam a 16 problemas. Estes problemas diferem em função do seu conteúdo: *Saúde (S)*, *Fidelidade (F)*, *Desporto (Dp)*, e *Escola (E)*; e do princípio normativo aplicável na resposta: *Lei dos Grandes Números (LGN)*, *Regressão à Média (RM)*, *Base-Rates (BR)*, e *Diagnosticidade (D)*.

Assim, tem-se quatro domínios de conteúdo que se combinam com quatro princípios normativos, de onde resultam os 16 problemas referidos acima⁵.

No que diz respeito aos domínios de conteúdo, foram usados quatro conjuntos de quatro problemas, cujo temática se enquadra: em situações

ou assuntos directamente relacionados com serviços e/ou profissionais de saúde, medicamentos e doenças (S); situações ou assuntos directamente relacionados com a importância da fidelidade nas relações pessoais (designadamente nas relações conjugais) (F); situações ou assuntos directamente relacionados com a actividade desportiva na forma de diversas modalidades de desporto (Dp); e situações ou assuntos directamente relacionados com o contexto de avaliação do desempenho escolar (E).

No caso dos princípios normativos, tem-se também quatro conjuntos de quatro problemas, sendo que nas respostas a estes problemas são aplicáveis: raciocínios que envolvem a consideração do tamanho de amostras quando se pretende concluir sobre as características da população (LGN); raciocínios que envolvem a noção de que (em qualquer variável com uma componente aleatória) após a ocorrência de um valor extremo é mais provável a ocorrência de um valor mais próximo da média do que a ocorrência de outro valor igual ou mais extremo (RM); raciocínios que, face a um evento incerto, envolvem a consideração das probabilidades de ocorrência de partida desse evento (BR); e raciocínios que envolvem a escolha de informação que permita a avaliação ou revisão das probabilidades de um dado evento incerto ser *verdadeiro (D)*⁶.

⁵ Os problemas que envolvem a *Lei dos Grandes Números* e a *Regressão à Média* foram criados tendo por referência o tipo de problemas utilizados por Jepsen, Krantz e Nisbett (1983) e Fong, Krantz e Nisbett (1986). Os problemas que envolvem a *Diagnosticidade* foram «inspirados» nas tarefas de Doherty et al. (1979). Os problemas que envolvem Base-Rates foram criados tendo como referencial o problema do Tom W. de Tversky e Kahneman (1973).

⁶ O tempo necessário para responder a 32 problemas resultantes da combinação dos domínios de conteúdo com os princípios normativos (2 x 4 x 4) levantou dificuldades de aplicação do material que se resolveram criando três versões, cada uma com dois conjuntos de 8 problemas. Dada a impossibilidade de combinar todos os domínios de conteúdo com todos os princípios normativos apenas em três versões, optou-se por estruturar as versões da seguinte forma: na versão 1, quatro problemas resultam da combinação dos conteúdos *desporto* e *escola* com os princípios *Lei dos*

1.4. A Estrutura dos Problemas de Teste

Para além de combinar sempre um princípio normativo com um domínio de conteúdo, a estrutura dos diversos problemas desenvolvidos para o presente estudo apresenta outros aspectos comuns.

Assim, no caso dos problemas que envolvem os princípios da *Lei dos Grandes Números* ou da *Regressão à Média*, tratam-se de histórias, em relação às quais surgem duas opiniões divergentes. Uma das opiniões baseia-se em argumentos «determinísticos» e a outra em argumentos «estatísticos» (ambos os argumentos apresentados nas histórias).

No caso dos problemas que envolvem a consideração de *Base-Rates*, tratam-se de histórias onde se confrontam informação circunstancial e informação relativa às probabilidades de partida de um dado evento ou ocorrência (apresentada normalmente de forma implícita).

Os problemas que envolvem *Diagnosticidade*, são histórias onde os participantes são confrontados com duas hipóteses, H e \bar{H} , e duas fontes de informação (ou dados) para avaliar estas duas hipóteses, D e \bar{D} . Na história é dada informação sobre $P(D/H)$. De acordo com o desenrolar da história, só é possível obter um dos restantes três ti-

pos de informação: $P(D/\bar{H})$, $P(\bar{D}/H)$ ou $P(\bar{D}/\bar{H})$. Os participantes devem decidir qual a melhor informação para poder comparar as duas hipóteses⁷.

1.5. Material das manipulações experimentais

Material referente à manipulação de imprevisibilidade (R): é constituído por quatro listas de factores referentes a quatro domínios de conteúdo: *Desporto*, *Fidelidade*, *Acidentes de viação* e *Baixas militares em cenários de guerra*.

Estes factores são situações ou acontecimentos que, tendo em geral um certo grau de incerteza associado, afectam o comportamento das pessoas dentro do domínio de conteúdo em questão⁸.

É pedido aos participantes para avaliar o grau de imprevisibilidade dos factores em cada uma destas listas, numa escala de 7 pontos (de 1 - totalmente previsível; até 7 - totalmente imprevisível).

Material referente à manipulação por analogia (A): é constituído por quatro conjuntos de dois

Grandes Números e *Base-Rates* (2 x 2), e os outros quatro resultam da combinação dos conteúdos *Fidelidade* e *Saúde* com os princípios *Regressão à Média* e *Diagnosticidade* (2 x 2); na Versão 2, quatro problemas resultam da combinação dos conteúdos *Escola* e *Fidelidade* com os princípios *Base-Rates* e *Regressão à Média* (2 x 2), e os outros quatro resultam da combinação dos conteúdos *Saúde* e *Desporto* com os princípios *Diagnosticidade* e *Lei dos Grandes Números* (2 x 2); na Versão 3, quatro problemas resultam da combinação dos conteúdos *Escola* e *Saúde* com os princípios *Base-Rates* e *Diagnosticidade* (2 x 2), e os outros quatro resultam da combinação dos conteúdos *Fidelidade* e *Desporto* com os princípios *Regressão à Média* e *Lei dos Grandes Números* (2 x 2). Dentro de cada versão não há, obviamente, problemas repetidos. Criaram-se ainda duas ordenações diferentes em cada um dos conjuntos de 8 problemas, de forma a controlar eventuais efeitos de ordem.

⁷ Para consultar exemplos dos problemas, ver anexo I.

⁸ No caso do *Desporto*, a lista de factores refere-se a diversas circunstâncias ou acontecimentos que podem afectar o resultado de um jogo de futebol (por exemplo «os ressaltos casuais da bola que desencadeiam jogadas de perigo»). No caso da *Fidelidade*, a lista de factores refere-se a circunstâncias ou situações que afectam as atitudes e comportamentos de fidelidade ou infidelidade das pessoas (por exemplo «ficar isolado numa casa com uma pessoa do sexo oposto, muito interessante, divertida e atraente, durante vários dias como consequência das péssimas condições meteorológicas»). No caso dos *Acidentes de viação*, a lista de factores refere-se a circunstâncias ou situações que podem afectar o desempenho dos automobilistas (e logo a ocorrência de acidentes (por exemplo «a fraca iluminação de algumas viaturas em viagens nocturnas»). No caso das *Baixas militares em cenários de guerra*, os factores referem-se a circunstâncias que afectam o desempenho dos soldados e consequentemente a ocorrência de baixas militares (por exemplo «um momento de desatenção de algum soldado»).

problemas. Em cada um destes conjuntos, um dos problemas é identificado como «HISTÓRIA A» e o outro como «HISTÓRIA B». Os dois problemas são apresentados lado a lado numa mesma página.

A «HISTÓRIA A» é sempre um problema de natureza claramente probabilística onde é saliente que a informação ou os dados do problema são gerados através de processos que envolvem variação aleatória (as quatro «HISTÓRIAS A» usadas nesta manipulação envolvem: retirar bolas de uma urna, lançamentos de um dado, lotaria, e retirar cartas de um baralho).

A «HISTÓRIA B» é sempre um problema social, referente a situações do quotidiano, equivalente aos problemas de teste.

Na página seguinte à apresentação dos dois problemas, é identificada a questão central de uma e outra história, e são apresentadas lado a lado, duas listas de aspectos relevantes de cada uma das histórias. Aos sujeitos é pedido que procurem estabelecer a correspondência entre estes aspectos de uma e outra história (ligando os aspectos das listas com traços). O objectivo é tornar mais saliente para os participantes a existência de elementos comuns ou correspondentes em termos dos espaços de amostragem, fontes de acaso, etc., aumentando assim a relevância da aplicação do mesmo princípio normativo à «HISTÓRIA B».

Material referente à manipulação definições (Dfs): consiste numa curta definição formal de cada um dos princípios normativos seguido, em cada caso, por um exemplo ilustrativo da aplicação do princípio. Cada princípio (definição e exemplo) é apresentado numa página diferente. É pedido aos sujeitos para ler atentamente as descrições de alguns princípios estatísticos envolvidos nos problemas-teste.

1.6. Procedimento

A apresentação do material experimental ocorreu em três momentos diferentes, durante um período de três semanas.

Numa primeira vez, foram distribuídos pelos sujeitos cadernos de oito problemas (correspondendo a problemas de uma das três versões referidas na descrição do material de teste)⁹ com as seguintes instruções na primeira página:

«Seguidamente são apresentadas oito pequenas histórias ou dilemas de cariz social, algumas das quais baseadas em casos verídicos.

Após a leitura de cada uma destas histórias ser-lhe-á pedido para expressar a sua opinião sobre o assunto central da história.

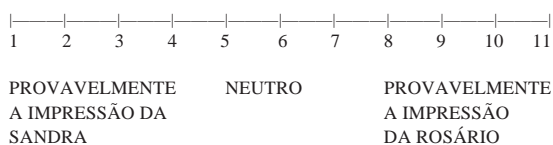
Para isto deve assinalar numa escala de 11 pontos o número que está mais de acordo com a sua opinião.

Exemplo:

A Sandra tinha estado com a Clara apenas uma vez, mas tinha ficado com uma fraca impressão dela. A Clara tinha-se revelado muito calada e desinteressada das pessoas à sua volta. Das poucas vezes que falaram a Sandra tinha achado o discurso da Clara desinteressante e enfadonho.

Ao comentar isto com a Rosário (uma amiga comum que estava frequentemente com a Clara) esta disse-lhe que, pelo contrário, achava a Clara uma boa conversadora, sociável, que exprimia as suas ideias com graça, sendo agradável e interessante discutir com ela.

Qual das impressões relativamente à Clara estará mais correcta?



Uma semana mais tarde, foram distribuídos pelos mesmos sujeitos outros cadernos constituídos por uma das três manipulações experimentais referidas na descrição do material das manipulações, mais um conjunto de quatro problemas introduzidos pelas mesmas instruções acima transcritas, excepto no caso da condição controlo em que foram distribuídos cadernos constituídos somente pelas instruções mais quatro problemas.

Dois semanas mais tarde, os mesmos sujeitos respondem aos últimos cadernos de quatro problemas.

Desta forma, cada sujeito respondeu a um

⁹ O número de sujeitos que responderam a cada versão é aproximadamente igual para as três versões.

total de 16 problemas: 8 problemas antes da manipulação, 4 problemas imediatamente após a manipulação e 4 problemas duas semanas após a manipulação.

2. RESULTADOS

A hipótese geral do presente trabalho pode agora ser retomada. De acordo com Fong e Nisbett (1991), a intuição estatística humana está normalmente dependente da aquisição (pela experiência) de *regras de codificação* adequadas que permitam o reconhecimento da aplicabilidade das heurísticas estatísticas a certas circunstâncias ou problemas sociais. Contudo, aquilo que parece estar subjacente a quaisquer *regras de codificação* é o reconhecimento de que os problemas sociais em causa envolvem uma componente de imprevisibilidade. Assim, condições experimentais que tornem saliente esta componente de imprevisibilidade (nomeadamente as tarefas referentes a R e a A) deverão facilitar, só por si, uma codificação apropriada dos problemas e consequentemente um aumento das respostas de natureza estatística.

2.1. Transformação dos dados originais

Para o teste da presente hipótese é necessário agregar os dados originais por princípio normativo ou por domínio de conteúdo¹⁰. Assim, para cada princípio normativo (excepto diagnosticidade) e domínio de conteúdo criaram-se três novas variáveis. A primeira corresponde à média aritmética das respostas aos problemas da pri-

¹⁰ Note-se que a análise estatística dos dados não agregados corresponderia a avaliar o efeito das manipulações problema a problema. Tais medidas de desempenho têm um erro de medida associado muito maior devido ao efeito que as características específicas de cada problema têm nas respostas dos participantes. Ao agregar-se as respostas por princípio normativo ou domínio de conteúdo obtêm-se medidas de desempenho mais estáveis e menos afectadas pelos enviesamentos de respostas associados às especificidades de cada problema porque os diversos erros de medida associados tendem a anular-se mutuamente (assumindo uma distribuição Normal dos erros).

meira aplicação; a segunda corresponde à média das respostas aos problemas da segunda aplicação; e a terceira variável corresponde à média das respostas aos problemas da terceira aplicação.

Após a agregação dos dados para cada uma das três aplicações (antes, logo após, e duas semanas depois da manipulação), o teste da hipótese acima explicitada foi realizado comparando cada uma das condições experimentais com a condição controlo. Para este efeito, e dado que se trata de comparações não ortogonais, recorreu-se ao teste *t* de Dunnett uma vez que este teste permite controlar os efeitos de inflação do α no caso de múltiplas comparações com o mesmo grupo controlo (Winer, 1971).

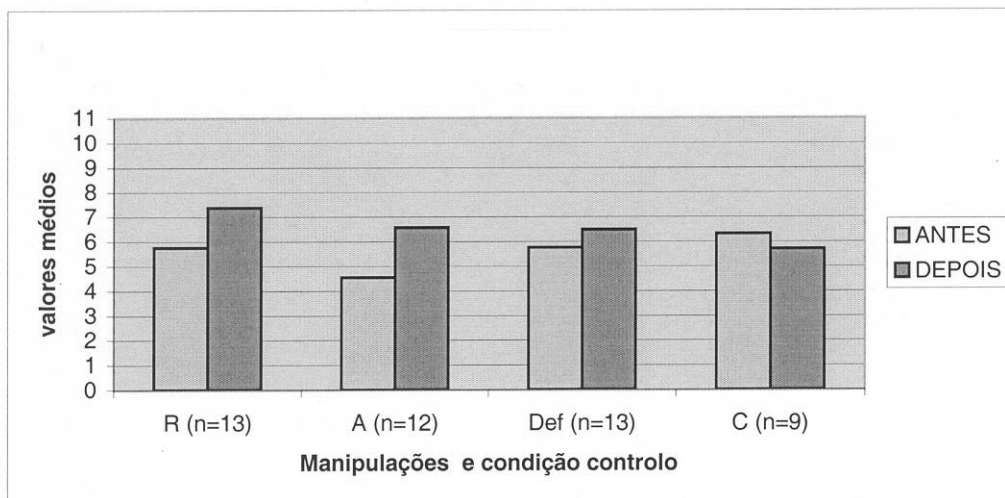
No caso dos problemas sobre *Diagnosticidade*, as respostas não foram dadas numa escala ordinal mas correspondiam antes a uma escolha forçada entre três opções de resposta. Assim, dado tratar-se de dados nominais, criou-se uma tabela de contingência referente à frequência de sujeitos que, para cada uma das quatro condições (três manipulações e uma condição controlo) passaram a optar pela escolha diagnóstica após a segunda aplicação do material (o que corresponde a uma *melhoria* do desempenho) e a frequência de sujeitos que deixaram de optar pela escolha diagnóstica após a segunda aplicação do material (o que corresponde a uma *pioria* do desempenho). Assim, tem-se uma tabela de 4 x 3 (as quatro condições versus *melhoria*, constância ou *pioria* do desempenho após a segunda aplicação). A avaliação de diferenças na frequência de resposta dos sujeitos no sentido de uma *melhoria* (ou *pioria*) no seu desempenho foi avaliada usando o teste do χ^2 e posterior análise residual da tabela de contingência (Everitt, 1977).

2.2. Efeitos das condições experimentais sobre o desempenho dos sujeitos através dos diversos princípios normativos¹¹

Em termos médios verificou-se uma ligeira melhoria do desempenho estatístico dos participantes para os quatro princípios normativos, em

¹¹ Realizaram-se duas análises de variância para cada princípio normativo. Numa escolheu-se como variável dependente a média das respostas aos problemas

FIGURA 1
Princípio normativo da *Regressão à Média*: Valores médios antes e depois das manipulações experimentais



função das condições R e A. No entanto, só no caso do princípio da *Regressão à Média* (ver Figura 1) é que esta melhoria do desempenho é estatisticamente significativa, $t(4,35) = 2,30$; $p < .05$ (teste t de Dunnet)¹² para a condição R, e $t(4,35) = 2,84$; $p < .05$, para a condição A.

Relativamente ao princípio normativo Diagnosticidade, obteve-se um resultado estatisticamente significativo para a tabela de contingência referente às quatro condições (ou quatro níveis da variável independente) versus melhoria ou pioria do desempenho após a segunda aplicação ($\chi^2 = 13.26$; $p < .05$).

(referentes ao princípio em causa) apresentados imediatamente após as manipulações (problemas da segunda aplicação); noutra escolheu-se como variável dependente a média das respostas aos problemas (referentes ao princípio em causa) apresentados imediatamente e duas semanas após as manipulações (problemas das segunda e terceira aplicações). O padrão de resultados foi igual num e noutro caso. Daí que os resultados seguidamente apresentados agreguem os problemas das segunda e terceira aplicações.

¹² Uma vez que se comparam as médias de todas as condições com a mesma condição controlo, os três testes resultantes não são independentes. Dunnet derivou a distribuição de amostragem para a estatística t que permite controlar a inflação do α , mantendo-o constante para o conjunto de todas as comparações (Winer, 1971).

A análise residual mostrou que R é a principal responsável pela obtenção deste efeito. De facto, no caso desta manipulação houve um aumento significativo ($|z| > 1.96$; $p < .05$) dos sujeitos que *melhoram* nas suas respostas aos problemas sobre *Diagnosticidade* da primeira para as segunda e terceira aplicações (ver anexo II). As restantes variações (de *melhoria* ou *pioria*) das respostas não são significativas, excepto para o caso da Dfs onde há um aumento considerável dos sujeitos que «pioram» no seu desempenho da primeira para as últimas aplicações ($|z| > 1.96$; $p < .05$).

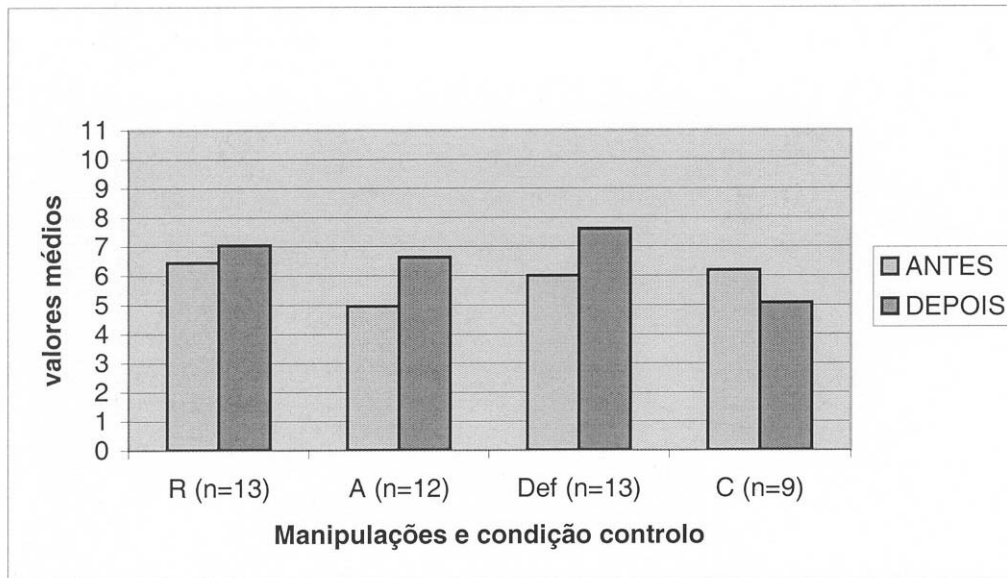
Relativamente à *Lei dos Grandes Números* e aos *Base-Rates*, não se encontraram efeitos significativos.

2.3. Efeitos das condições experimentais sobre o desempenho dos sujeitos através dos diversos domínios de conteúdo¹³

Em termos médios verifica-se uma ligeira melhoria do desempenho estatístico dos participantes para os quatro domínios de conteúdo, em função de R e A.

¹³ Realizaram-se, também neste caso, duas análises de variância para cada domínio de conteúdo. Numa es-

FIGURA 2
Domínio de conteúdo *Fidelidade*: Valores médios antes e depois das manipulações experimentais



Contudo, só para a condição A, e só nos casos dos domínios de conteúdo *Fidelidade*, $t(4,35) = 2,59$; $p < .05$ (ver Figura 2) e *Escola*, $t(4,35) = 2,38$; $p < .05$ (ver Figura 3) é que esta «melhoria» do desempenho é significativa. As restantes comparações não são significativas excepto para o domínio *Fidelidade* que também melhora significativamente em função da condição Dfs ($t(4,35) = 2,39$; $p < .05$).

Relativamente aos restantes domínios (Saúde

e Desporto), não se encontraram efeitos significativos¹⁴.

2.4. Análise do nível de dificuldade dos problemas de teste

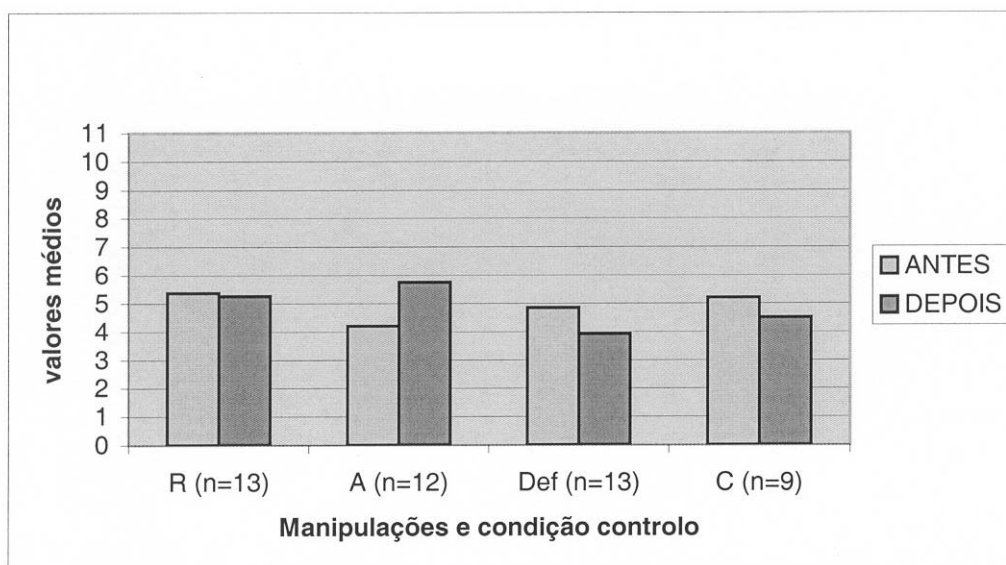
No sentido de avaliar o impacto do grau de dificuldade dos problemas nos presentes resultados computou-se a diferença entre as variáveis que agregam as respostas aos problemas por princípio normativo ou por domínio de conteúdo, antes e depois das manipulações. Tem-se assim medidas da «evolução» do desempenho dos sujeitos para cada princípio normativo e para cada domínio de conteúdo. Estas medidas estão negativamente correlacionadas com as respostas do

escolheu-se como variável dependente a média das respostas aos problemas (referentes ao domínio em causa) apresentados imediatamente após as manipulações (problemas da segunda aplicação); noutra escolheu-se como variável dependente a média das respostas aos problemas (referentes ao domínio em causa) apresentados imediatamente e duas semanas após as manipulações (problemas das segunda e terceira aplicações). O padrão de resultados foi igual num e noutro caso. Daí que os resultados seguidamente apresentados agreguem os problemas das segunda e terceira aplicações.

¹⁴ No caso dos domínios de conteúdo não foi possível avaliar o efeito das condições experimentais para o princípio *Diagnosticidade*, uma vez o número de participantes não é suficiente para computar as respectivas tabelas de contingência.

FIGURA 3

Domínio de conteúdo *Escola*: Valores médios antes e depois das manipulações experimentais



sujeito por princípio normativo e domínio de conteúdo da primeira aplicação (ver Quadros 1 e 2).

Ou seja, quanto *pior* foi o desempenho dos sujeitos nos problemas da primeira aplicação maior é a *melhoria* obtida nos problemas respondidos nas segunda e terceira aplicações, e vice-versa.

Estas observações parecem sugerir que o grau de dificuldade dos problemas, só por si, é um factor que influencia a «evolução» do desempenho dos sujeitos ao longo das aplicações.

3. DISCUSSÃO

Os resultados da investigação podem resumir-se nos seguintes pontos:

- a) Encontraram-se efeitos de facilitação das heurísticas estatísticas para os princípios normativos *Regressão à Média* (melhoria em função de A e R) e *Diagnosticidade* (melhoria em função de R), e para os domínios de conteúdo *Fidelidade* e *Escola* (melhoria em função de A). Nestes casos, como previsto, a saliência da componente de acaso promovida pelas condições experimentais levou a uma *melhoria* do desempe-

nho dos sujeitos no sentido da produção de respostas de natureza (mais) estatística;

- b) A condição Dfs não levou a efeitos de facilitação consistentes. Logo, os efeitos acima referidos não se devem meramente ao facto de tornar saliente certas regras estatísticas mas sim à promoção do reconhecimento de uma componente de acaso envolvida nos problemas de teste¹⁵;
- c) As medidas da «evolução» do desempenho dos sujeitos para cada princípio normativo e para cada domínio de conteúdo estão negativamente correlacionadas com as respostas do sujeito (por princípio normativo ou domínio de conteúdo) da primeira aplicação. Ou seja, quanto *pior* foi o desempenho dos sujeitos nos problemas da primeira aplicação maior é a *melhoria* obtida nos

¹⁵ A condição Dfs levou a dois efeitos estatisticamente significativos pouco consistentes e contraditórios (melhoria do desempenho para a *Fidelidade* e pioria do desempenho para a *Diagnosticidade*). Não temos qualquer explicação psicológica para este padrão de resultados pelo que não será mais referido.

QUADRO 1

Tabela de correlações para os princípios normativos RM, BR, e LGN. Os valores apresentados referem-se às correlações entre as respostas dadas pelos participantes antes das manipulações e a diferença das respostas antes e depois das manipulações

Respostas dadas antes das manipulações	Diferença das respostas dadas antes e depois das manipulações		
	RM	BR	LGN
RM (Regressão à Média)	-.688 p < .001	-.228 p = .123	.099 p = .509
BR (Base-Rates)	-.069 p = .646	-.724 p < .001	-.313 p < .05
LGN (Lei dos Grandes Números)	.103 p = .490	-.087 p = .561	-.566 p < .001

QUADRO 2

Tabela de correlações para os domínios de conteúdo F, E, Dp, e S. Os valores apresentados referem-se às correlações entre as respostas dadas pelos participantes antes das manipulações e a diferença das respostas antes e depois das manipulações

Respostas dadas antes das manipulações	Diferença das respostas dadas antes e depois das manipulações			
	F	E	Dp	S
F (Fidelidade)	-0,822 p = ,000	-0,4113 p = ,008	0,053 p = ,742	-0,1043 p = ,517
E (Escola)	-0,1869 p = ,242	-0,7655 p = ,000	-0,1341 p = ,403	0,0293 p = ,856
Dp (Desporto)	0,0488 p = ,762	0,1962 p = ,219	-0,7698 p = ,000	-0,0286 p = ,859
S (Saúde)	0,0912 p = ,571	-0,0731 p = ,650	0,142 p = ,376	-0,6104 p = ,000

problemas respondidos nas segunda e terceira aplicações, e vice-versa. Estas correlações revelam uma tendência para uso de heurísticas estatísticas mesmo antes das manipulações experimentais, o que obviamente dificulta a obtenção de efeitos significativos de *melhoria* do desempenho em função das manipulações. Ou seja, os resultados sugerem que o baixo grau de dificuldade dos problemas é provavelmente responsável pela obtenção de *efeitos de tecto*.

3.1. *Discussão dos efeitos da Manipulação de Imprevisibilidade (R) e da Manipulação por Analogia (A) sobre os princípios normativos e domínios de conteúdo usados*

A hipótese geral que originou a presente investigação encontra apoio experimental sobretudo através da manipulação por Analogia (A). No entanto, a manipulação de imprevisibilidade (R) foi a única que levou a uma melhoria do desempenho no caso da *Diagnosticidade*. Alguns aspectos discutidos seguidamente, podem ajudar a esclarecer este padrão de resultados.

As tarefas de manipulação A e R

A tarefa da condição A levou a melhorias de desempenho para a *Regressão à Média, Escola e Fidelidade*. A tarefa da condição R levou a uma melhoria do desempenho no caso da *Regressão à Média* (em menor grau do que a condição A) e no caso da *Diagnosticidade*. Aquilo que caracteriza as condições A e R, é o facto de ambas tornarem saliente a existência duma componente de acaso envolvida nos problemas de teste. Aquilo que mais marcadamente distingue as duas condições é o facto de a tarefa de analogia (A), para além de tornar saliente a componente aleatória, fornecer ou indicar aos participantes «formas» de representar estatisticamente os problemas de teste (analogias entre espaços de amostragem, fontes de acaso, etc.). Com efeito, na condição A os participantes são levados (pelas analogias estabelecidas entre problemas probabilísticos e problemas equivalentes aos problemas de teste), a codificar estatisticamente os problemas de teste e a perceber de que forma e em que aspectos, a informação estatística envolvida nos problemas de teste deve influenciar as suas respostas. Isto parece ser vantajoso quando a

correspondência com os problemas de teste é relativamente fácil de estabelecer (caso da *Regressão à Média, Escola, e Fidelidade*). No entanto, no caso em que esta correspondência não é imediata, como no caso dos problemas de *Diagnosticidade*, a tarefa da condição A não traz vantagens. Neste caso, a melhoria do desempenho estatístico parece ser função da tarefa de reconhecimento da componente aleatória mais abstracta, R. Este padrão de resultados está de acordo com a ideia de que as implicações da saliência duma componente de acaso no caso dos problemas envolvendo *Diagnosticidade* não são tão imediatas como no caso dos outros princípios normativos. Enquanto que reconhecimento da influência do acaso é um indicador directo da importância do tamanho da amostra, da consideração de Base-Rates e da regressão à média, a relação entre o reconhecimento do acaso e escolhas diagnósticas faz-se um nível mais abstracto. Este envolve provavelmente o reconhecimento da importância de que, nestes problemas, a distribuição na população das opções de resposta têm igual probabilidade, tornando assim mais saliente a relevância da escolha da opção diagnóstica (i.e., da opção que permite a obtenção de probabilidades condicionais comparáveis).

Os domínios de conteúdo

Jepson Krantz e Nisbett (1983) mostraram que os mesmos sujeitos, no mesmo contexto, respondendo ao mesmo tipo de problemas dão tendencialmente menos respostas «estatísticas» para problemas envolvendo atributos subjectivos do que para problemas envolvendo atributos objectivos. No presente estudo, os problemas sobre *Fidelidade e Escola* são os mais «difíceis» (i.e., parece ser mais difícil reconhecer o papel da incerteza envolvida nestes domínios).

De entre os quatro domínios de conteúdo usados, os problemas sobre *Fidelidade* (um domínio claramente subjectivo) envolvem atributos de natureza claramente mais subjectiva. Consequentemente, o papel do acaso na se torna-se neste caso bastante mais difícil de reconhecer.

O domínio de conteúdo *Escola*, embora envolvendo uma natureza mais objectiva do que o domínio *Fidelidade*, contrasta com os restantes domínios (*Saúde e Desporto*) em termos da tendência, partilhada culturalmente, de lidar com as

questões de desempenho (escolar) em termos de mérito.

De resto, a nível cultural pode estabelecer-se um contraste entre *Fidelidade* e *Escola*, por um lado; e *Desporto* e *Saúde*, por outro. Com efeito, na nossa Sociedade existe, aquilo a que Nisbett, Krantz, Jepson e Kunda (1983) chamam, «prescrições culturais» para raciocinar estatisticamente no caso de domínios de conteúdo como o Desporto («rankings», «goal average», percentagens de concretização, etc.); e a Saúde (probabilidade de apanhar uma doença, o grau de risco de uma intervenção cirúrgica, correlação entre certos sintomas e certas doenças, etc.). No entanto, no caso da *Fidelidade*, dificilmente as pessoas equacionam ou discutem assuntos deste domínio em termos da probabilidade de ser infiel nos próximos seis meses; número médio de circunstâncias que facilitam comportamentos de infidelidade por ano; etc. No caso da *Escola*, embora o domínio se preste a uma análise de natureza estatística (por exemplo, a avaliação do desempenho dos estudantes em termos da média de resultados), existe uma forte tendência de atribuição do desempenho escolar a factores internos (como o esforço e a capacidade), acompanhada da negligência de factores externos (como a sorte ou acaso). Um resultado escolar elicit normalmente uma explicação em termos do mérito (ou desmérito) do estudante, independentemente de ser positivo, negativo, mais ou menos extremo, ou mais ou menos estatisticamente representativo.

Em suma, a maior dificuldade dos problemas envolvendo a *Fidelidade*, prende-se com a sua natureza subjectiva que torna pouco transparente as condições de aplicação de princípios estatísticos, e com as «prescrições culturais» que existem para este domínio. A maior dificuldade dos problemas envolvendo a *Escola* prende-se sobretudo com «prescrições culturais», nomeadamente uma certa cultura de meritocracia que muitas vezes caracteriza a aprendizagem escolar. Assim, os efeitos de melhoria obtidos para a *Fidelidade* e a *Escola* em função de A, e a ausência de efeitos nos restantes casos deve-se provavelmente à maior dificuldade destes dois domínios por comparação com os domínios de conteúdo *Desporto* e *Saúde*. De facto, se nestes últimos, logo à partida a componente de acaso está facilmente ac-

sível, o efeito de reconhecimento do acaso resultante de R ou de A tende a ser trivial.

Os princípios normativos

No caso da *Regressão à Média*, é possível que o aumento das respostas estatísticas se deva ao facto de se tratar, à partida, de um princípio contra-intuitivo. Kahneman e Tversky (1973) referem que mesmo quando as pessoas se apercebem dos efeitos da regressão, estes são vistos como mudanças que precisam de uma explicação causal, e acrescentam que muitas explicações espúrias de mudanças que resultam meramente da regressão à média têm sido evocadas até por cientistas. Assim, ao tornar saliente (através de R e A) a componente de acaso envolvida nos problemas apresentados ter-se-á levado os sujeitos a preferirem ou a reconhecerem como mais plenamente «satisfatórios» os argumentos baseados na *Regressão à Média*.

No caso da *Lei dos Grandes Números*, a ausência de «melhoria» do desempenho dos sujeitos pode parecer intrigante dado que se trata de um princípio normativo muito semelhante à *Regressão à Média* (em termos da lógica que lhe está subjacente). No entanto, Jepson e colegas (1983) mostraram que o número de sujeitos que responde de acordo com a *Lei dos Grandes Números* varia em função da «objectividade» (facilidade de codificação) dos atributos ou instâncias relevantes de cada problema em particular.

No presente estudo, as instâncias ou atributos relevantes dos problemas envolvendo a *Lei dos Grandes Números* são «fáceis» de codificar, o que facilita o uso adequado deste princípio normativo. Com efeito, estes problemas foram os mais «fáceis» de todos¹⁶. É provável que a aplicação da *Lei dos Grandes Números* tenha sido desde o início de tal forma generalizada que os efeitos de R e A não se traduziram numa melhoria significativa do desempenho dos sujeitos. Assim, embora a mesma regra intuitiva esteja

¹⁶ Numa escala de 11 pontos em que o «11» corresponde à resposta estatística, os problemas envolvendo a *Lei dos Grandes Números* obtiveram (na primeira aplicação) uma média de 7.15.

também em causa nos problemas envolvendo a Regressão à Média, a codificação das instâncias relevantes está, neste caso, mais dificultada à partida, diminuindo a probabilidade de aplicação da regra inferencial adequada nas respostas aos problemas que antecedem R e A.

Dos problemas referentes aos quatro princípios normativos, os que envolvem a consideração de *Base-Rates* são, em média, os mais «difíceis»¹⁷, o que põe de lado uma explicação da ausência de efeitos significativos em termos de um *efeito de tecto*.

Um possível factor explicativo que pode contribuir para compreender a ausência de efeitos decorrentes do reconhecimento da componente de acaso dos problemas-teste poderá relacionar-se com a própria estrutura dos problemas que envolvem a consideração de *Base-Rates*. Em metade dos problemas sobre *Base-Rates* a informação relativa às probabilidades de partida está implícita¹⁸ (contrastando com a informação circunstancial que é sempre explicitada). Ou seja, para estes problemas o efeito da facilitação decorrente do reconhecimento do papel do acaso não é significativo eventualmente porque a informação relativa às probabilidades de partida, ao contrário da informação circunstancial tem que ser subentendida pelos participantes. Assim, a melhoria do desempenho está não só dependente do reconhecimento do acaso (como nos outros problemas) mas também da capacidade de explicitar informação latente sobre probabilidades de partida. A saliência da imprevisibilidade promovida por R ou o estabelecimento de analogias (A), não podem ter impacto sobre informação

que não sendo explicitamente apresentada, tenderá a não fazer parte da representação destes problemas.

No que diz respeito à *Diagnosticidade*, e como já referido acima, só a condição experimental R levou a um aumento significativo do número de sujeitos que escolhem a opção diagnóstica. Este resultado parece indicar que a acessibilidade da heurística estatística correspondente ao princípio normativo da *Diagnosticidade* só acontece quando se usam tarefas de reconhecimento da componente aleatória, mais abstractas (como é o caso da R). Como já discutido, o facto de a aplicação do princípio normativo *Diagnosticidade* ser aquele que tem uma relação menos directa com o reconhecimento dum componente de acaso, está de acordo com esta necessidade de tarefas de manipulação do acaso mais abstractas, e ainda com a ausência de resultados positivos no caso da A (que se baseia em analogias directas entre problemas).

3.2. *Análise das condicionantes da investigação apresentada*

Os resultados empíricos apresentados no presente trabalho foram, por vezes, condicionados pelo facto dos problemas de teste surgirem, em média, como «fáceis». Neste sentido, um primeiro aspecto a considerar para desenvolvimentos futuros do estudo agora apresentado é uma avaliação mais rigorosa do nível de dificuldade dos problemas a usar, quer a nível objectivo (através do pré-teste dos problemas de forma a construir uma hierarquia de problemas em termos do grau de dificuldade envolvido); quer a nível subjectivo acrescentando uma medida subjectiva de dificuldade (pedindo aos participantes para avaliar o grau de dificuldade dos problemas a que vão respondendo).

Desta maneira, será possível controlar a ocorrência de *efeitos de tecto* assim como de heterogeneidade do grau de dificuldade do material experimental.

3.3. *Comentários finais*

Os resultados obtidos no presente estudo oferecem algum suporte experimental à ideia de que o reconhecimento de incerteza associada aos conteúdos envolvidos nos problemas (sobretudo

¹⁷ Numa escala de 11 pontos em que o «11» corresponde à resposta estatística, os problemas envolvendo *Base-Rates* obtiveram (na primeira aplicação) uma média de 4.86.

¹⁸ Sendo que, para os quatro problemas onde a informação sobre *Base-Rates* é explicitada, em dois deles trata-se apenas dum frase que chama a atenção para a «raridade» de ocorrência de um dado fenómeno ou evento. Ainda, os *Base-Rates* nunca são apresentados sobre a forma numérica como é tradicionalmente o caso nos problemas como o dos advogados e engenheiros (Khaneman & Tversky, 1972) ou como o problema dos táxis (Khaneman & Tversky, 1973).

para domínios de natureza subjectiva onde esta componente de incerteza é menos clara e onde predominam argumentos determinísticos, como a *Fidelidade* e a *Escola*) leva a uma facilitação do uso de pelo menos certas regras de raciocínio de natureza estatística (designadamente *Regressão à Média* e *Diagnosticidade*).

De forma indirecta, os presentes resultados também suportam empiricamente o carácter intuitivo das heurísticas estatísticas, propostas por Nisbett e colegas (1983). De facto, dado que não se recorreu a quaisquer sessões de treino que pudessem eventualmente fornecer aos sujeitos «novas» formas de raciocínio, a obtenção de efeitos de «melhoria» do desempenho dos sujeitos pela mera clarificação duma componente de acaso é difícil de explicar, caso não se aceite que estas manipulações facilitam a codificação da informação dos problemas de modo a que representações intuitivas de certos princípios estatísticos possam ser (mais) usadas.

Por outro lado, o presente estudo parece sugerir que a codificação das instâncias relevantes para a aplicação da regra estatística adequada (de entre o repertório de regras intuitivas que as pessoas possuem) não está necessariamente dependente de regras de codificação específica dos conteúdos onde são «geradas», tal como proposto por Fong e Nisbett (1991). A mera saliência de factores de incerteza (conseguida através de A e R) parece ter sido condição suficiente para a *melhoria* do desempenho.

Por fim, a ênfase dada neste trabalho a questões ligadas à codificação apropriada da informação de forma a levar a representações dos problemas adequadas ou facilitadoras da aplicação de certas heurísticas em vez de outras, possibilita uma perspectiva talvez mais equidistante na avaliação dos «defeitos» e «virtudes» do raciocínio indutivo humano.

Com efeito, muitos dos erros que são associados ao uso das heurísticas não estatísticas deverão ser antes atribuídos a factores relacionados com a facilidade ou dificuldade da codificação adequada das instâncias relevantes dos problemas em causa. A importância que se tem vindo a dar às questões de codificação da informação permite perspectivar o raciocínio indutivo humano em termos de um repertório de heurísticas maior e mais diversificado.

De facto, as heurísticas que as pessoas pos-

suem (quer se trate de heurísticas estatísticas ou não estatísticas) são todas muito úteis quando devidamente aplicadas.

A presença ou ausência de erros e enviesamentos decorrentes do raciocínio indutivo humano poderá depender da aplicação heurística mais adequada para a situação em causa. Para cada circunstância de julgamento poderão competir diversas heurísticas, aquela que é usada será a que a que melhor «funciona» com a representação que a pessoa criou da circunstância em causa.

Assim, poder-se-á melhorar a qualidade do raciocínio indutivo humano agindo ao nível da facilitação duma codificação adequada das instâncias envolvidas em cada situação ou problema que envolva julgamentos em condições de incerteza.

REFERÊNCIAS

- Doherty, M. E., Mynatt R., Tweney, R. D., & Schiavo, M. D. (1979). Pseudodiagnosticity. *Acta Psychologica*, 3, 11-121.
- Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. *Annual Review of Psychology*, 32, 53-88.
- Everitt, B. S. (1977). *The analysis of contingency tables*. CH, London, New York: Halsted Press.
- Fong, G. T., & Nisbett, R. E. (1991). Immediate and delayed transfer of training effects in statistical reasoning. *Journal of Experimental Psychology*, 120, 34-45.
- Fong, G. T., Krantz, D. H., & Nisbett, R. E. (1986). The effects of statistical training on thinking about everyday problems. *Cognitive Psychology*, 18, 293-328.
- Ginossar, Z., & Trope, Y. (1987). Problem solving in judgment under uncertainty. *Journal of Personality and Social Psychology*, 52, 464-474.
- Gilovich, T., Griffin, D., & Kahneman, D. (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge: Cambridge University Press.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgment under uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3, 430-454.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80, 237-251.
- Nisbett, R. E. (1993). *Rules for reasoning*. Hillsdale, NJ: Lawrence Erlbaum, LEA.

- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E., Krantz, D. H., Jepson, S. D., & Kunda, Z. (1983). The use of statistical heuristics in everyday reasoning. *Psychological Review*, 90, 339-363.
- Sherman, S. J., & Corthy, E. (1984). Cognitive heuristics. In R. S. Wyer, & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 189-286). Hillsdale, NJ: LEA.
- Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Review*, 2, 105-110.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 4, 207-232.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Tversky, A., & Kahneman, D. (1982). Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.
- Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.

RESUMO

Segundo Tversky e Kahneman (1974), a actividade inferencial humana baseia-se em grande medida em heurísticas (regras simplificadas de tomada de decisão) que divergem dos princípios estatísticos apropriados ao julgamento na incerteza. No entanto, Nisbett, Krantz, Jepson e Kunda (1983), defenderam que paralelamente às heurísticas *não-estatísticas*, as pessoas também possuem heurísticas estatísticas (i.e. representações intuitivas e abstractas de certos princípios estatísticos). Fong e Nisbett (1991), sugerem que o uso das heurísticas estatísticas está dependente de *regras de codificação* (associadas a domínios de conteúdo específicos). Aqui considera-se que o essencial das *regras de codificação* é a facilitação do reconhecimento do componente de acaso subjacente aos problemas indutivos. Assim, condições experimentais que facilitem o reconhecimento deste componente aleatório deverão resultar numa melhoria do desempenho estatístico. Pa-

ra testar esta hipótese, foram usadas duas manipulações de facilitação do reconhecimento do acaso. Os participantes responderam a um conjunto de problemas indutivos sobre diversos domínios (*Desporto, Fidelidade conjugal, Escola, e Saúde*), envolvendo vários princípios estatísticos (*Lei dos Grandes Números, Regressão à Média, Base-Rates, e Diagnosticidade*) antes e após as manipulações (imediatamente ou duas semanas depois). Os resultados revelam uma melhoria no desempenho estatístico para os domínios *Fidelidade e Escola*, e para os princípios *regressão à média e diagnosticidade*.

Palavras-chave: Indução, julgamento na incerteza, heurísticas.

ABSTRACT

Human inference is, according to Tversky and Kahneman (1974), greatly based in heuristics (simplified rules of decision making) that diverge from the statistical principles appropriate to judgment under uncertainty. However, Nisbett, Krantz, Jepson and Kunda (1983), contend that beyond these *non-statistical heuristics*, people also possess *statistical heuristics* (i.e., intuitive counterparts of statistical principles). Fong and Nisbett (1991), suggest that the use of statistical heuristics depends from codification rules associated to specific content domains. In the present work, we propose that codification rules facilitate the recognition of the chance component inherent to inductive problems. Thus, experimental conditions that facilitate the recognition of this chance component should result in the improvement of statistical performance. To test this hypothesis, we used two manipulations meant to facilitate chance recognition. Participants answered to a number of inductive problems concerning diverse content domains (*Sports, Fidelity, School and Health*), and regarding a number of statistical principles (*Law of the great numbers, Regression to the mean, Base-Rates, and Diagnosticity*) before and after the manipulations (immediately and two weeks after). The results showed that statistical performance improved in the domains of *Fidelity and School*, and for problems concerning *Regression to the mean* and *Diagnosticity* principles.

Key words: Judgment under uncertainty, heuristics.

ANEXO I

Exemplos dos problemas que constituem o material experimental da presente investigação

PROBLEMA A (SAÚDE x LGN)

O João diz para um amigo «aqui nesta revista dizem que os estudos mostram que o aumento do cancro do pulmão entre as mulheres é causado pelo aumento do número de mulheres fumadoras. No entanto a minha mãe e a minha sogra fumaram bastante durante toda a vida e, tanto uma como outra, já passaram dos oitenta anos». O amigo do João responde «mesmo considerando o caso da tua mãe e da tua sogra, se levarmos em linha de conta os estudos citados na revista, penso que fumar é uma causa de cancro».

Qual a sua opinião?

PROBLEMA B (ESCOLA x RM)

O professor de uma escola de teatro propôs aos seus alunos como trabalho de fim de curso, a preparação de uma peça de um autor contemporâneo, a ser apresentada ao público. O professor decidiu começar por constituir um elenco para com ele formar uma verdadeira equipa de trabalho, e só depois pensar na distribuição de papéis. Ao fazer as audições para o elenco, uma aluna que o professor conhecia mal, foi brilhante. Outra aluna que já tinha trabalhado sob direcção do professor noutras duas peças do mesmo autor contemporâneo, sempre com excelência representações, não foi brilhante nas audições para a presente peça, tendo sido mesmo fraca em certas partes da audição. O professor não dispunha de mais tempo e tinha que escolher já a actriz principal, resolveu então consultar dois colegas para o ajudar a decidir. O primeiro colega disse que ele devia basear-se na qualidade das audições pois estas indicam qual das alunas está, no momento, melhor preparada para o papel. No entanto, o segundo colega disse que ele devia escolher a aluna com quem já tinha trabalhado antes pois, ao contrário das outra, o conhecimento que tinha dela não se baseava numa «curta» audição mas na experiência de duas peças.

Qual a sua opinião?

PROBLEMA C (FIDELIDADE x DIAGNOSTICIDADE)

Há já algum tempo que a Florbela procura uma relação amorosa estável. Há dias o Pedro convidou-a para sair. Ainda não se tinha decidido a aceitar e eis que lhe telefona o Rui a convidá-la para ir ao cinema. Ela sabe que tanto o Rui como o Pedro estão interessados nela. Uma amiga sua que conhece os dois dizia-lhe «olha eu diria que em 1 de cada 10 ocasiões sociais o Rui procura sempre seduzir novas raparigas, quanto a serem fiéis à pessoa com quem namoram eu acho que...» a Florbela interrompeu dizendo, «não estou a pensar aceitar nenhum dos convites!». Mais tarde, no entanto, a Florbela resolveu sair com um dos rapazes. Lembrou-se então de telefonar à amiga para saber mais sobre o Rui e o Pedro. Como não queria dar a entender à amiga que estava indecisa entre eles, resolveu que só faria uma pergunta.

Qual a pergunta que a Florbela deveria fazer?

- A Florbela deveria perguntar, de 1 a 10 que possibilidade tinha o Rui de ser fiel
- A Florbela deveria perguntar, de 1 a 10 que possibilidade tinha o Pedro de ser fiel
- A Florbela deveria perguntar, por cada 10 ocasiões sociais, em quantas delas o Pedro procura seduzir novas raparigas.

PROBLEMA D (DESPORTO x BASE-RATES)

O Futebol Clube do Porto joga com o Salgueiros para a taça de Portugal. É um facto que o Porto sai claramente em vantagem na história dos confrontos entre as duas equipas. Por outro lado, quanto mais uma equipa remata à baliza e/ou obtém pontapés de canto, maiores são as suas possibilidades de ganhar o jogo. Ora, no começo da 2.^a parte, as duas equipas estavam ainda empatadas mas o Salgueiros tinha feito 18 remates com perigo enquanto que o Porto tinha rematado apenas 6 vezes à baliza adversária; o Salgueiros tinha obtido 10 pontapés de canto, enquanto que o Porto apenas 3. Qual das duas equipas terá maiores hipóteses de ganhar o jogo?

ANEXO II

Tabelas de contingência referentes ao princípio normativo *Diagnosticidade*

TABELA 1

Tabela de contingência referente às condições experimentais versus o desempenho comparado antes e depois das manipulações

Manipulações	Melhoria (%)	Constante (%)	Pioria (%)
R	46	46	8
A	25	25	50
Dfs	8	23	69
Controlo	22	22	56

TABELA 2

Tabela dos resíduos ajustados , referente aos dados da tabela de contingência acima apresentada*

Manipulações	Melhoria	Constante	Pioria
R	2,00	1,52	-3,15
A	-0,05	-0,42	0,43
Dfs	-1,73	-0,62	2,09
Controlo	-0,25	-0,55	0,73

* O procedimento estatístico relativo à análise residual pode ser encontrado em Everitt (1977).