

UNIVERSIDADE DE LISBOA
FACULDADE DE PSICOLOGIA E DE CIÊNCIAS DA
EDUCAÇÃO



SENSIBILIDADE METACOGNITIVA: MECANISMOS DE
MONITORIZAÇÃO E REVERSÃO DE FALSAS
MEMÓRIAS

David Van der Kellen Mendes

MESTRADO INTEGRADO EM PSICOLOGIA
(Cognição Social Aplicada)

2008

UNIVERSIDADE DE LISBOA
FACULDADE DE PSICOLOGIA E DE CIÊNCIAS DA
EDUCAÇÃO



SENSIBILIDADE METACOGNITIVA: MECANISMOS DE
MONITORIZAÇÃO E REVERSÃO DE FALSAS
MEMÓRIAS

David Van der Kellen Mendes

Dissertação orientada pelo Prof. Doutor Leonel Garcia-Marques

MESTRADO INTEGRADO EM PSICOLOGIA
(Cognição Social Aplicada)

2008

RESUMO

Nas últimas décadas, têm-se alcançado grandes desenvolvimentos na modelação da memória humana, especialmente na memória de reconhecimento. Os modelos actualmente propostos podem ser divididos em dois tipos, os modelos globalistas e os modelos dualistas, ambos capazes de descrever eficazmente a grande maioria dos efeitos encontrados na literatura. Apesar destes desenvolvimentos, os modelos de memória não têm sido capazes de incorporar os mecanismos de monitorização da memória, ignorando inúmeros efeitos de supressão de memórias falsas encontrados na literatura sobre o paradigma DRM. Esta situação ocorre porque a monitorização simplesmente não é contemplada pelos modelos, e pelo facto de que medidas de memória populares como a *Teoria de Detecção de Sinal* são incapazes de distinguir entre a monitorização das memórias, e a força das memórias. Esta limitação resulta em modelos que não permitem a monitorização e consequentemente, a manifestação de reversões de resposta com uma acuidade acima do acaso. O programa experimental aqui estabelecido procura propor “novas” medidas de monitorização da memória, derivadas da *Teoria de Detecção de Sinal Tipo-2*, e através destas, identificar os diferentes componentes processuais subjacentes à monitorização da memória. A primeira experiência procura avaliar a acuidade metacognitiva dos indivíduos no paradigma DRM, e identificar os componentes processuais subjacentes desse desempenho. A segunda experiência procura testar a previsão dos modelos de memória em relação à monitorização, assumindo-se uma versão “relaxada” dos modelos, que assume a acumulação de evidências ao longo do tempo. As implicações dos estudos e medidas aqui propostas na investigação em memória de reconhecimento são discutidas.

Palavras-chave: Memória de reconhecimento, monitorização da memória, metacognição, falsas memórias, paradigma DRM, Teoria de Detecção de Sinal Tipo-2.

ABSTRACT

In the last decades, there has been a tremendous development in memory modelling literature, especially in terms of recognition memory. These models can be divided in two major classes, global models and dualist models, both accounting for most of the phenomenon encountered in the relevant literature. In spite of these developments, memory models haven't accounted for memory monitoring mechanisms, ignoring the numerous false memory suppression effects encountered in the DRM paradigm literature. This situation occurs because memory monitoring is simply not predicted by the models, and the fact that popular memory measures, like the ones provided by *signal detection theory* are unable to distinguish between memory monitoring and memory strength. This limitation results in models that don't allow for memory monitoring, and subsequently, response reversals in memory judgments whose accuracy is above-chance. The proposed experimental program sets out to consider "new" measures of memory monitoring, derived from *Type-2 Signal Detection Theory*, and through them, highlight the different processes underlying memory monitoring. The first experiment sets out to understand metacognitive accuracy in the DRM paradigm, and to assess the process components underlying the participants' performance. The second experiment tests the models' prediction of chance-level monitoring accuracy by assuming a more "relaxed" version that assumes evidence accumulation throughout time, which can theoretically account for the monitoring accuracy results. The implications of the proposed study and measures in future recognition memory research are discussed.

Keywords: Recognition memory, memory monitoring, metacognition, false memories, DRM paradigm, Type-2 Signal Detection Theory.

Quero agradecer....

Ao Prof. Dr. Leonel Garcia-Marques, por todo o apoio que me deu ao longo da elaboração desta monografia e não só, sendo uma referência no meu desenvolvimento enquanto aluno e pessoa. Nesta monografia, muitos dos seus pontos fortes (para não dizer todos) se devem ao Prof. Leonel, enquanto que os pontos fracos são todos da minha (ir)responsabilidade. É com grande prazer e entusiasmo que vejo esta monografia como um ponto de partida para projectos mais ambiciosos e para a construção de uma bela relação. Obrigado, e até já.

À Ludmila Nunes, por ter sido uma amiga e colega fantástica ao longo destes anos, tendo arranjado paciência para me aturar durante todo este tempo. Parte desta monografia deve-se à Ludmila por todo o apoio que sempre me deu, e pelo *feedback* dado, o qual me permitiu torná-la melhor e mais sólida.

Ao Prof. Dr. José Manuel Palma, não apenas por todo o apoio que sempre me deu, mas também por termos tido aquela famosa discussão há vários anos. É em parte por causa dela que me interessei por Cognição Social e acabei aqui a fazer isto.

Ao Prof. Dr. Mário Ferreira, à Prof. Dra. Sofia Santos, e ao Prof. Dr. Frederico Marques pela aprendizagem, dedicação e acompanhamento proporcionados ao longo destes dois anos. Tem sido um prazer.

Aos colegas da Secção de Cognição Social Aplicada, por termos conseguido sobreviver inteiros a isto tudo.

À Prof. Dra. Teresa Garcia-Marques, por acreditar em mim e ter-me desafiado para uma das experiências mais enriquecedoras que tive.

Ao “pessoal do ISPA”, pela amizade e por tudo o que deram na cabeça

À maltinha, por toda a força e parvoíce muita querida.

À Jette, por existir, e por isso não ter de ser inventada.

Ao “eixo do mal”, por serem gente ruim como eu.

Quero dedicar este trabalho aos meus pais, à minha avó materna, e ao meu irmão. Quero dedicar igualmente à memória do meu tio e da minha avó paterna. Por todo o amor que se sempre me deram, por tudo, obrigado.

ÍNDICE

Introdução	1
1. Modelos de memória: Propostas globalistas e dualistas.....	4
2. O problema da monitorização: Monitorização parcial e independente.....	10
3. O problema da monitorização: Falsas memórias associativas.....	16
4. O problema da monitorização: Tomada de decisão e julgamentos de confiança...	24
5. O Problema e a Abordagem.....	31
6. Metamemória: Um enquadramento para o problema.....	33
7. TDS tipo-2: Sensibilidade e critério de decisão na monitorização	35
8. Mecanismos de monitorização e modelos de memória: Um estudo empírico.....	38
Experiência 1.....	42
Método	42
Resultados e Discussão.....	45
Experiência 2.....	48
Método	48
Resultados e Discussão.....	50
Discussão Geral.....	52
Referências Bibliográficas.....	58
Anexos	73
ANEXO 1 – O motivo da composição das listas DRM.....	74
ANEXO 2 – Exemplo de lista de associados semânticos (Lista DRM).....	76

"It is only prudent never to place complete confidence in that by which we have even once been deceived."
- René Descartes

"The truth is, a person's memory has no more sense than his conscience, and no appreciation whatever of values and proportions."
- Mark Twain

"...o Dicionário de Raridades, Inverosimilhanças e Curiosidades, onde, admirável coincidência que vem a matar neste aventuroso relato, se dá como exemplo de erro a afirmação do sábio Aristóteles de que a mosca doméstica comum tem quatro patas, redução aritmética que os autores seguintes vieram repetindo por séculos e séculos, quando já as crianças sabiam, por crueldade e experimentação, que são seis as patas da mosca, pois desde Aristóteles as vinham arrancando, voluptuosamente contando, uma, duas, três, quatro, cinco, seis, mas essas mesmas crianças, quando cresciam e iam ler o sábio grego, diziam uma para as outras, A mosca tem quatro patas, tanto pode a autoridade magistral, tanto sofre a verdade com a lição dela que sempre nos vão dando."
- José Saramago em "História do Cerco de Lisboa"

Introdução

Em que podemos acreditar? Em que tipo de evidências é possível basear as nossas respostas e confiança? A dúvida sobre a capacidade dos indivíduos encontrarem verdades, ou pelo menos serem capazes de atribuir algum grau de crença, tem sido uma questão que percorre toda a história da Filosofia, tendo sido abordada não apenas por Descartes, mas por outros pensadores como Hume ou Popper. Apesar destas questões se reflectirem tradicionalmente na investigação psicológica, nomeadamente no estudo do modo como os indivíduos apreendem e utilizam a informação proveniente do meio exterior (e.g., Tolman & Brunswik, 1935), a sua integração no domínio psicológico leva a que possam ser estendidas ao meio psicológico interno do indivíduo, nomeadamente à informação representada em memória. Em que medida esta informação interna pode não ser verídica, e quais são as origens e características próprias da informação falsa?

A extensão deste problema ao meio psicológico interno é especialmente relevante para o estudo da memória humana, dado o papel fundamental desta não apenas na simples recordação de informação semântica e episódica, mas também no modo como nos descrevemos a nós próprios, interagimos uns com os outros, tomamos decisões, ou fazemos julgamentos (Tulving & Craik, 2005). A possibilidade da formação e manifestação de falsas memórias é bastante verosímil se considerarmos a literatura sobre ilusões perceptivas, na qual verificamos que esta caracteriza a percepção como um processo activo e construtivo, sendo por isso bastante vulnerável a erros e distorções (Roediger, 1996). Se um processo que tem como base a presença directa dum estímulo pode gerar diversos tipos de erros, então um processo que recorre a recordação de informação passada deve ser, pelo menos, igualmente susceptível a erros. Assim, se a informação presente em memória não for totalmente verídica, torna-se necessário compreender até que ponto os indivíduos têm percepção desta situação, e se sim, de que forma conseguem discriminar entre informação falsa e verdadeira. A possibilidade de erro levanta a questão da monitorização desse mesmo erro.

Apesar da plausibilidade das falsas memórias, estas foram negligenciadas enquanto objecto de estudo durante décadas; situação esta que teve a sua origem na própria fundação do estudo experimental da memória humana, nomeadamente nos trabalhos publicados por Hermann Ebbinghaus (1885/1964). Ebbinghaus tinha como objectivo compreender o contributo da memória na reaprendizagem de listas de pseudo-palavras, e para tal desenvolveu toda uma metodologia que lhe permitia avaliar a informação dada ao participante, o número de ensaios necessários à aprendizagem das listas, e subsequentemente, a capacidade destes em reaprenderem as listas após uma fase de latência (e esquecimento). Deste modo, foi possível a Ebbinghaus estudar as relações existentes entre a capacidade de aprendizagem, o impacto do tempo no esquecimento, e a

reaprendizagem. A ênfase dada à recordação verídica, adequada aos objectivos específicos deste estudo, não se limitou ao trabalho de Ebbinghaus e foi adoptado como objecto de estudo central nesta área de investigação, o que levou a que os erros de memória fossem negligenciados como mero ruído, intrusões sem qualquer relevância ou características específicas. Esta situação verificou-se igualmente na modelação dos mecanismos mnésicos, a qual tem normalmente por base a Teoria de Detecção de Sinal (Green & Swets, 1966; Macmillan & Creelman, 2005), segundo o qual a informação mnésica é representada segundo uma distribuição de familiaridade, sendo os erros de memória provocados por uma segunda distribuição de familiaridade associada aos distractores, designada por distribuição de “ruído”.

A noção de falsas memórias enquanto fenómeno fundamentalmente mnésico, e como tal, um objecto de estudo relevante apenas conquistou um papel central na literatura com o trabalho de Henry Roediger III e Kathleen McDermott (1995), no qual se reabilitou um procedimento experimental originalmente concebido por James Deese (1959), actualmente designado por paradigma Deese-Roediger-McDermott (DRM). Neste paradigma, os participantes são expostos a listas de palavras, todas elas associadas a uma palavra-crítica que não é apresentada na fase de estudo. Na fase de teste, é pedido aos sujeitos que recordem livremente as palavras anteriormente apresentadas ou então que as reconheçam numa lista de teste. Em ambas as condições de teste verificou-se que os sujeitos recordam a palavra-crítica, apesar desta não ter sido apresentada. Estas falsas memórias são acompanhadas não apenas de um elevado nível de confiança (Read, 1996), mas também de informação de natureza episódica (Roediger & McDermott, 1995). Um aspecto importante na literatura de falsas memórias, é o facto de se verificar a existência de mecanismos de monitorização, os quais se manifestam na capacidade dos indivíduos em avaliar as diferentes características das suas memórias, como informação conceptual, perceptiva e contextual, de modo a monitorizar a origem e veracidade das suas memórias (Johnson, Hashtroudi, & Lindsay, 1993; Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1996; Roediger, Watson, McDermott & Gallo, 2001). Apesar dos diversos desenvolvimentos alcançados nas áreas das falsas memórias (Gallo, 2006), estes não foram integrados nos modelos psicométricos de memória que procuram descrever o desempenho dos indivíduos. De facto, estes efeitos são bastante problemáticos para os modelos de memória em geral, dado que estes consideram que a informação recuperada dos traços é representada de uma forma indistinta, ao longo de uma dimensão de familiaridade. A monitorização dos conteúdos mnésicos representa assim um desafio para os modelos de memória, os quais têm sido desenvolvidos até então com base em estudos que utilizam listas de palavras não associadas, as quais acabam por produzir artificialmente a familiaridade indistinta que os modelos propõem. A questão das falsas memórias associativas e a monitorização não tem sido um aspecto

considerado na literatura que aborda os modelos de memória, pelo que não existem medidas disponíveis para a avaliação dos mecanismos de monitorização.

É proposto no presente estudo que a quantificação da capacidade dos mecanismos de monitorização pode ser alcançado com a utilização de uma derivação da TDS designada por TDS tipo-2 (Clarke, Birdsall, & Tanner, 1959; Galvin, Podd, Drga, & Whitmore, 2003). A TDS tipo-2 permite avaliar o desempenho do mesmo modo que a TDS tradicional (tipo-1), proporcionando duas medidas independentes de sensibilidade e de enviesamento de resposta. O objectivo principal neste estudo é então, através da TDS tipo-2, explorar a capacidade de monitorização dos indivíduos no contexto do paradigma DRM, no qual se verificam proporções elevadas de falsos reconhecimentos, e aos quais os participantes atribuem elevados níveis de confiança. Contudo, como será referido mais adiante, a capacidade dos indivíduos manifestarem um desempenho acima do acaso não é prevista pelos modelos de memória em geral; segundo a TDS, dado que a informação mnésica é representada de forma indistinta, os indivíduos não devem conseguir obter um desempenho superior na tarefa tipo-2, o qual implicaria necessariamente uma melhoria na capacidade discriminativa. Os modelos de memória que contemplam um processo de recolção (e.g. Yonelinas, 1997) adicional conseguem contemplar um desempenho positivo na tarefa tipo-2, contudo este deve ter apenas por base um processo de recolção o que tem consequências no tipo de função matemática que melhor descreve esse desempenho.

São então propostas duas experiências; a primeira experiência consiste na implementação do paradigma DRM, sendo incluída uma segunda fase de teste, a tarefa tipo-2. Tanto na tarefa de reconhecimento como na tarefa tipo-2 será manipulado o critério de resposta dos indivíduos através de payoffs, de modo a que se obtenham funções zROC, as quais descrevem o desempenho do indivíduo. Esta primeira experiência assume dois objectivos principais, em primeiro lugar a descrição do desempenho dos indivíduos na tarefa tipo-2, e que processos podem ser atribuídos a este desempenho com base nas características da função zROC, nomeadamente se estamos perante um processo de recolção, um processo de familiaridade, ou ambos. Em segundo lugar, esta experiência procura entender a capacidade que os modelos de memória têm em conseguir explicar o desempenho na tarefa tipo-2, tendo em conta as características da função zROC.

A segunda experiência procura avaliar mais precisamente a verdadeira capacidade dos modelos em explicarem o desempenho acima do acaso na tarefa tipo-2, assumindo-se uma versão “relaxada” dos modelos. Para tal será utilizado o paradigma sinal-resposta (e.g., Ratcliff, 2006), o qual permite a obtenção de uma função que descreve o crescimento da acuidade consoante o tempo dado para o julgamento de reconhecimento.

Estas duas experiências procuram então demonstrar a relevância da TDS tipo-2 para o estudo dos mecanismos de monitorização que operam ao nível da memória de reconhecimento, mecanismos estes que são fundamentais para uma compreensão adequada da memória humana.

Modelos de memória: Propostas globalistas e dualistas

O estudo da memória humana foi sempre acompanhado pelo desenvolvimento de modelos matemáticos que procuravam descrever o desempenho dos indivíduos. De facto, mesmo o próprio Ebbinghaus (1885/1964) procurou ajustar as suas curvas de esquecimento observadas a uma função específica. No estudo da memória de reconhecimento a situação não é diferente, sendo esta uma área em que existem debates acesos sobre que tipo de modelo melhor descreve o desempenho dos indivíduos (e.g., Diana, Reder, Arndt, & Park, 2006; Wixted, 2007).

O debate tem-se centrado em dois tipos de modelos específicos, o modelo globalista¹ e o modelo dualista. Segundo o modelo globalista (e.g., Ratcliff, Sheu, & Gronlund, 1992), a informação presente nos traços mnésicos é representada por uma variável contínua, a familiaridade, a qual assume uma distribuição normal ao longo dos itens codificados na memória. Neste modelo, os julgamentos de memória têm como base a quantificação das memórias, sendo avaliada a familiaridade de cada um dos traços. Por outro lado, o modelo dualista (Mandler, 1980; Yonelinas, 1997), postula que para além desta componente de familiaridade, a informação pode ser obtida através de uma componente de recolocção, a qual é definida por um processo de alto limiar, ou seja que a recordação é um acontecimento de “*tudo-ou-nada*”, que ocorre com a sua probabilidade de ocorrência.

Antes de considerar as especificidades de ambos os modelos, é necessário distinguir os diferentes níveis em que a modelação pode ser implementada; a um nível mais elevado, a modelação assume uma natureza psicométrica, ocorrendo ao nível da medição do desempenho geral dos indivíduos, quer seja na distinção entre processos mnésicos e processos decisórios, ou na avaliação do papel de diferentes fontes de informação. Este tipo de modelação assume um número limitado de pressupostos acerca da natureza da memória, porém não faz qualquer tipo de considerações relativamente aos processos de codificação ou à estrutura representacional envolvida. Pelo contrário a modelação pode assumir um nível inferior, procurando descrever o desempenho dos indivíduos tendo em conta determinados mecanismos e processos de codificação,

¹ Ao longo do texto, é utilizada a expressão genérica “modelo globalista”, cujos aspectos podem ser aplicadas aos diversos modelos globalistas, como o REM, TODAM, ou MINERVA2 (para uma descrição das diferenças entre os diferentes modelos, ver Raaijmakers & Shiffrin, 2002). O mesmo se aplica à designação “modelo dualista”.

representação e recuperação, e fazendo previsões relativamente ao efeito de diferentes factores em cada um dos componentes dos modelos. Os dois níveis de modelação levantam assim questões distintas, porém complementares; os modelos psicométricos consideram a possibilidade de uma determinada noção de memória, enquanto que os modelos mecanicistas preocupam-se com a implementação dessa noção de memória (Batchelder & Riefer, 1999).

O modelo psicométrico globalista tem como base a teoria de detecção de sinais (TDS; Green & Swets, 1966; Macmillan & Creelman, 2005), segundo a qual a detecção de uma determinada categoria de estímulos é feita através de uma quantificação das evidências associadas a estes, quantificação esta que é representada ao nível de uma única dimensão. A quantidade de evidências associada a cada estímulo individualmente varia, o que resulta numa distribuição dos estímulos de cada categoria ao longo da dimensão de evidências. O processo de detecção ocorre necessariamente na presença de alguma incerteza ou ruído, pelo que o desempenho de um decisor na detecção dos estímulos pertencentes a uma categoria específica (designados por *estímulos-sinal*) depende da capacidade em distingui-los do ruído envolvente. Este ruído é constituído por outros estímulos (designados por *estímulos-ruído*), que apesar de não pertencerem à categoria especificada, manifestam quantidades de evidências associadas que são por vezes semelhantes ou mesmo superiores às de alguns *estímulos-sinal*. Esta situação resulta numa sobreposição das distribuições dos *estímulos-sinal* e *estímulos-ruído* (assumidas como distribuições *gaussianas*, ver Figura 1), o que torna a sua discriminação imperfeita, e implica a presença necessária de erros no desempenho de qualquer decisor, definidos na Figura 2.

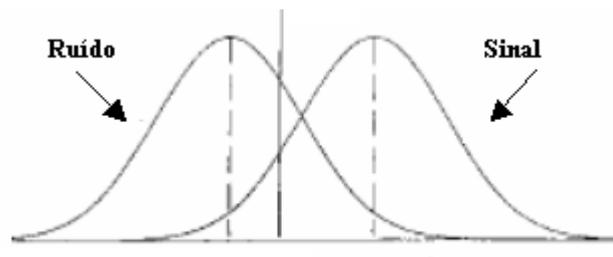


Figura 1. Distribuição dos estímulos-sinal e ruído. Adaptado de Creelman & Macmillan (2005).

A presença de incerteza no processo de detecção leva a que um decisor necessite estabelecer um critério de decisão na discriminação dos estímulos, o qual implica necessariamente custos. Se um critério for demasiado liberal, ou seja, se a detecção ocorrer com base numa quantidade de evidências baixa, vão ocorrer uma proporção elevada de acertos, porém estes serão acompanhados por uma proporção igualmente elevada de falsos alarmes, o que compromete a qualidade do desempenho global. Pelo contrário, se o critério estabelecido for bastante estrito, ou seja, apenas forem considerados estímulos que apresentem uma grande quantidade de evidências,

há uma redução drástica dos falsos alarmes, porém à custa de uma elevada proporção de rejeições incorrectas.

		RESPOSTA	
		“SINAL”	“RUÍDO”
ESTÍMULO	SINAL	ACERTOS (HITS)	OMISSÕES
	RUÍDO	FALSOS ALARMES	REJEIÇÕES CORRECTAS

Figura 2. Matriz de estímulo-resposta.

Esta condição de incerteza e a relação existente entre custos e benefícios no estabelecimento do critério de decisão, constitui a base da TDS, sendo designada por o *problema fundamental da detecção* (Swets, Tanner, & Birdsall, 1961). Assim, segundo a TDS podemos definir dois componentes subjacentes ao desempenho de um decisor, a sua sensibilidade discriminativa, que corresponde à sua capacidade em distinguir à partida os estímulos-sinal dos estímulos-ruído; e o seu critério de resposta, que define a quantidade de evidências necessária para a aceitação de um item em determinada categoria.

Consideremos um exemplo bastante simples dos princípios da TDS. Imaginemos que foram medidas as alturas de dois grupos de mulheres, sendo que num destes grupos as medições foram feitas com as mulheres descalças, e no outro, calcando sapatos de salto alto. As alturas de ambos grupos distribuem-se normalmente, apresentando contudo médias distintas dado que um dos grupos foi medido calçado, o que resulta em algo semelhante à figura 1. Em seguida, é pedido a um decisor que identifique se uma mulher foi medida calçada, tendo apenas como base a sua altura medida. Dado que existe uma sobreposição das duas distribuições (e.g., mulheres altas descalças e mulheres baixas calçadas), não é possível a um decisor obter um desempenho livre de erros – falsos alarmes e omissões. A relação entre respostas correctas e incorrectas pode ser compreendida a partir das características da distribuição normal (nomeadamente a *função densidade de probabilidade*), e da distância entre as médias das duas distribuições, distância que neste caso corresponde ao incremento causado pela utilização de sapatos na medição. Esta compreensão permite saber a probabilidade de erro associada a cada altura, tanto em termos de falsos alarmes e omissões. Se considerarmos que o indivíduo quer maximizar o seu desempenho global, este vai atribuir à distribuição de mulheres calçadas calçadas, alturas que sejam iguais ou superiores ao valor definido pela intersecção das duas distribuições, dado que só a partir desse valor, a atribuição à distribuição de mulheres calçadas permite uma probabilidade de correcta das mulheres calçadas

acima de 50%. Contudo, o decisor pode assumir outros valores mínimos para a sua resposta, podendo assumir valores mais baixos ou liberais, o que levam a uma diminuição das omissões (mulheres baixas calçadas) e a um aumento dos falsos alarmes (mulheres altas descalças), ou então valores mínimos mais elevados ou estritos, que levam a menos falsos alarmes e a mais omissões. Neste exemplo, a distância entre as duas médias corresponde então à sensibilidade discriminativa, e a altura mínima aceite pelo decisor corresponde ao critério de decisão.

Na TDS, as duas componentes subjacentes ao desempenho do decisor são avaliadas de forma independente por duas medidas, uma medida de sensibilidade discriminativa dos estímulos, d' , que corresponde à distância entre as médias das distribuições de familiaridade do ruído e do sinal; e uma segunda medida, C , que avalia o posicionamento do critério, sendo a distância em relação à intersecção das distribuições utilizada como referência. Tanto d' como C são medidas independentes, uma alteração no valor de d' não tem necessariamente no valor de C correspondente, e vice-versa.

Dada esta situação, é relevante perceber como o desempenho dos sujeitos é modificado pela variação dos valores de apenas uma destas duas medidas. Um dos modos de avaliação é através das funções de isosensibilidade, que podem estar representadas tanto sob a forma de funções ROC, ou então de funções zROC (ver Figura 3), havendo nesta últimas uma transformação das proporções de resposta em z -scores. As funções de isosensibilidade descrevem o desempenho de um decisor (em termos da proporção de acertos e falsos alarmes) com um determinado nível de sensibilidade discriminativa, *ao longo* dos diversos critérios de resposta possíveis. Dada a natureza distribucional da componente TDS, todas as funções ROC que este modelo consegue produzir são curvas, e as funções zROC são sempre lineares. A zROC é uma função bastante informativa no que diz respeito aos aspectos fundamentais da sensibilidade discriminativa; o d' corresponde à distância da sua intersecção à origem, enquanto que a razão das variâncias das distribuições de sinal e ruído corresponde ao declive da recta. Este último parâmetro é especialmente importante dado que muitas funções ROC observadas são assimétricas, pelo que o modelo globalista necessita assumir que a variância da distribuição de sinal seja normalmente superior e possa variar.

Na aplicação da TDS à memória de reconhecimento (Banks, 1970; Lockhart & Murdock, 1970), assume-se que cada traço mnésico tem associado a si uma variável contínua que corresponde ao que podemos designar por força ou familiaridade do traço. Esta familiaridade varia ao longo dos traços e é evocada pelos estímulos correspondentes durante o teste de reconhecimento. Contudo, a familiaridade não é uma propriedade exclusiva dos traços mnésicos a reconhecer, estando igualmente presente em itens novos apresentados durante o teste de reconhecimento. A familiaridade destes itens novos pode ter várias origens desde a sua semelhança com os itens codificados nos traços mnésicos até à sua ocorrência prévia num contexto diferente.

Assim, os itens apresentados na fase de estudo correspondem aos *estímulos-sinal*, enquanto que os itens que apenas foram apresentados no teste de reconhecimento equivalem aos *estímulos-ruído*. A tarefa do decisor é então de discriminar os itens anteriormente apresentados dos distractores, tendo como base a familiaridade que é evocada por estes estímulos. Para tal, o decisor assume um critério de resposta que lhe vai permitir regular o seu desempenho de acordo com os objectivos pretendidos. Como a familiaridade corresponde à única dimensão considerada na discriminação dos estímulos, o decisor vai necessariamente manifestar erros de memória, sejam falsos reconhecimentos ou rejeições incorrectas.

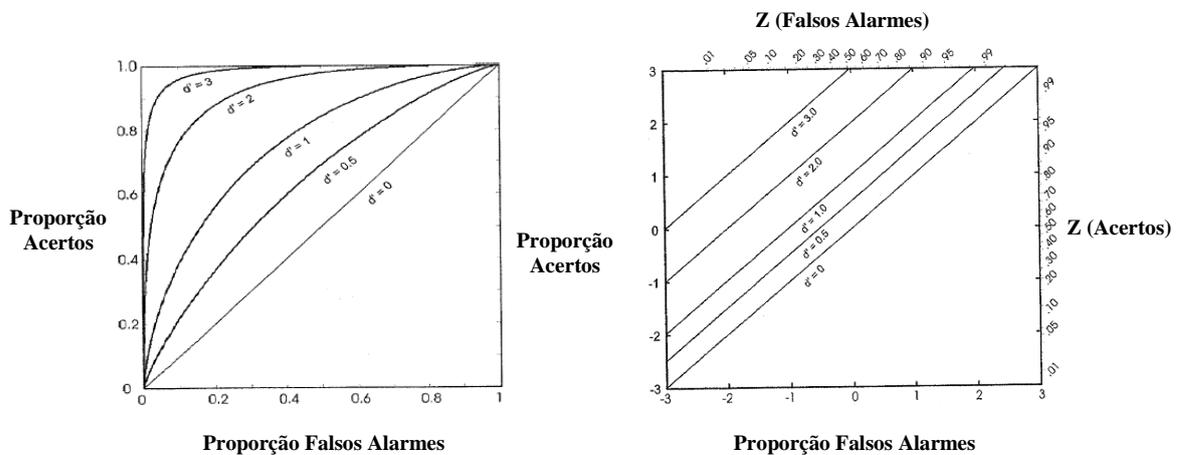


Figura 3. Exemplos de curvas ROC e de rectas zROC. Adaptado de Creelman & Macmillan (2005).

Ao contrário do modelo globalista, o modelo psicométrico dualista (e.g, Yonelinas, 1997) postula que o desempenho dos indivíduos tem por base dois componentes distintos e independentes. O primeiro componente é a familiaridade postulada pelo modelo globalista (esta familiaridade é constringida, como será explicitado adiante), enquanto que o segundo componente corresponde à recollecção, a qual é um processo de recordação “semelhante ao *recall*”, e que, como foi referido anteriormente, é definido por um processo de alto limiar, o qual é definido por um parâmetro pontual, nomeadamente a probabilidade de recordação episódica do estímulo de teste (Macmillan & Creelman, 1990). Segundo os proponentes do modelo dualista, a recollecção é um tipo de recordação controlado e lento, ao contrário da familiaridade, que é assumida como uma forma de recordação automática e rápida (Yonelinas, 2001). Estes dois componentes distinguem-se ainda na natureza da informação veiculada, a familiaridade apenas corresponde a informação indistinta, não sendo possível a diferenciação de qualquer tipo de informação dos estímulos; esta situação deve-se ao facto da familiaridade ser um construto unidimensional no qual a informação contida nos traços é agregada. Pelo contrário, a recollecção consiste numa recuperação do traço

episódico, pelo que toda a informação, contextual, perceptiva, semântica, associativa, etc., é acedida na sua totalidade, permitindo a sua avaliação individual.

O componente de recollecção, dado ser concebido como um processo de alto limiar definido por uma distribuição uniforme, produz necessariamente funções ROC lineares e zROC curvilíneas quando isolado² (Macmillan & Creelman, 2005). Porém, quando conjugado com o componente de familiaridade, o qual segundo o modelo dualista é um componente em que as variâncias de sinal e ruído são equivalentes, são produzidas as funções ROC assimétricas semelhantes às observadas empiricamente. O modelo dualista propõe então uma explicação alternativa para o problema das funções ROC assimétricas observadas ao longo da literatura – a recollecção. Esta explicação contrasta com o parâmetro de variância da distribuição de sinal postulado pelo modelo globalista. Dada a natureza distinta dos componentes do modelo dualista, é possível fazer previsões quanto ao desempenho dos indivíduos quando lhes é pedido que distingam itens consoante a sua lista de apresentação prévia (itens da lista A vs. itens da lista B). Dado que ambos os tipos de itens são igualmente familiares, a sua discriminação apenas pode ser feita com base no componente de recollecção, o que deve produzir funções ROC lineares. Segundo o modelo globalista, as funções ROC obtidas neste tipo de tarefa devem continuar a ser curvas. As previsões do modelo dualista foram confirmadas por Yonelinas (1997); o que veio consagrar este modelo como uma hipótese explicativa tão ou mais viável que o modelo globalista (no entanto ver Qin, Raye, Johnson, & Mitchell, 2001).

Os dois modelos propõem então parâmetros distintos para a explicação do declive das funções zROC, os quais se distinguem ao nível da sua importância teórica. Para o modelo de familiaridade, a variância da distribuição de sinal é um parâmetro livre que permite ajustar o modelo a funções zROC com declives inferiores a um, sendo esta a única razão da sua inclusão. A natureza desta opção é reconhecida desde os primórdios da TDS, tendo sido abordada por Green e Swets (1966):

The justification for the Gaussian model with unequal variance is, we believe, not to be made on theoretical but rather on practical grounds. It is a simple, convenient way to summarize the empirical data with the addition of a single parameter. (p. 79)

² A obtenção de funções ROC lineares e zROC curvilíneas quando são assumidas distribuições uniformes, e ROC curvilíneas e zROC lineares quando são adoptadas distribuições normais depende única e exclusivamente das características matemáticas destas distribuições, as quais não são relevantes para a presente exposição. Para uma explicação aprofundada ver Macmillan e Creelman (2005).

Pelo contrário, o modelo dualista procura explicar os declives inferiores a um através de um componente de recollecção, o qual possui toda uma concepção teórica na sua inclusão do modelo, como refere Yonelinas (2001):

Recollection is associated with auto-noetic consciousness (i.e., self-knowing or remembering), in which the episodic aspects of the study event are consciously reexperienced. (p.361)

Ambos os modelos se têm atacado e defendido numa série estudos (para uma revisão ver Malmberg, no prelo), não tendo sido ainda alcançada qualquer conclusão. Contudo, como se verificará mais adiante, ambos os tipos de modelos apresentam limitações no que diz respeito a um componente fundamental na memória de reconhecimento – a monitorização. Como será referido mais adiante, esta questão tem sido largamente ignorada pela literatura que estuda os modelos propostos, dado que estas investigações abordam a memória humana.

O problema da monitorização: Monitorização parcial, dinâmica, e independente

Uma das características comuns a ambos os modelos psicométricos de memória é o facto de não considerarem a possibilidade dos indivíduos acederem de forma estratégica, a apenas determinados tipos de informação. O modelo globalista não prevê essa hipótese dado que o componente de familiaridade consiste numa dimensão na qual toda a informação se encontra reunida de forma indiferenciada; por outro lado, o modelo dualista apenas prevê o acesso aos diferentes tipos de informação através do componente de recollecção, o que implica que os indivíduos tenham necessariamente de recordar toda a informação episódica, e não apenas características isoladas. A possibilidade de uma recollecção parcial é contemplada pelo modelo dualista, contudo esta informação é considerada como não diagnóstica, não sendo por isso abrangida pela medida de recollecção *R* (Yonelinas & Jacoby, 1996).

Apesar de esta não parecer ser uma questão central no estudo da memória de reconhecimento, a sua importância é revelada pela literatura produzida numa área designada por *monitorização da fonte*. Nesta área, assume-se que a recordação é um processo imperfeito, pelo que muitas vezes a origem ou contexto de uma determinada informação não é directamente acessível (Johnson, Hashtroudi, & Lindsay, 1993). Esta incerteza relativamente à origem da informação em memória leva a que muitas vezes os indivíduos confundam informação obtida externamente, com informação produzida internamente, como inferências lógicas. Um estudo clássico que demonstra este tipo de erros é o de Bransford & Franks (1971), no qual eram apresentadas a participantes

frases que descreviam uma determinada situação (e.g., as formigas comeram a geleia; a geleia estava na cozinha), verificando-se num teste de reconhecimento posterior que os indivíduos recordavam-se erroneamente de frases congruentes, porém não apresentadas previamente (e.g., as formigas estavam na cozinha). A literatura de monitorização da fonte centra-se então na capacidade dos indivíduos em distinguirem não apenas informação de origem externa e interna, mas também em discriminarem diferentes origens e características externas das memórias.

Segundo a literatura de monitorização da fonte (Johnson, Hashtroudi, & Lindsay, 1993; Johnson & Raye, 1981; Raye, Johnson & Taylor, 1980), a discriminação das diferentes origens da informação presente em memória tem por base dois aspectos; em primeiro lugar o tipo de evidências normalmente proporcionadas pelas diferentes fontes, e em segundo lugar o tipo de avaliação que os indivíduos fazem destas evidências. A recordação de um evento observado deve proporcionar mais evidências perceptivas e contextuais do que a recordação de um evento imaginado, pelo que se os indivíduos utilizarem a informação perceptiva ou contextual como critério de decisão, deverão obter um melhor desempenho do que se utilizassem outro critério qualquer. Quanto ao tipo de avaliação efectuado, este pode ser mais controlado, sistemático e exaustivo, sendo considerados os diferentes tipos de evidências, ou então um processo mais automático e heurístico, sendo apenas considerado uma parte da informação num julgamento frugal (e.g., Shiffrin & Schneider, 1977).

Nos estudos de monitorização que abordam a discriminação de fontes externas, os estímulos são inicialmente apresentados com múltiplas características associadas, como diferentes cores, tipos de letra, vozes distintas, ou em diferentes localizações no ecrã. Em seguida, é pedido aos participantes que identifiquem os estímulos de acordo com essas características, ou seja que estímulos foram apresentados com uma determinada cor, foram ditos por uma voz específica, ou foram apresentados numa determinada posição do ecrã. Este tipo de discriminações são mais dependentes de informação de natureza episódica, e implicam que os participantes consigam utilizar os diferentes tipos de informação que a recordação do episódio de aprendizagem proporciona. O traço episódico pode disponibilizar informação semântica, perceptiva, afectiva, ordem temporal, ou outras, e cabe ao participante utilizar as diferentes configurações informacionais de forma estratégica.

Este enquadramento vem assim enfatizar uma perspectiva multidimensional³ das evidências utilizadas na discriminação dos estímulos, a qual tem sido explorada em diversos trabalhos (Banks, 2000; Bayen & Murmane, 1996; Ferguson, Hashtroudi, & Johnson, 1992; Meiser & Bröder, 2002; Meiser, Sattler, & Weißer, 2008). Nestes estudos, algumas das dimensões (e.g. voz, localização da voz) foram “sobrepostas” (*voz feminina/localização esquerda; voz*

³ Ao longo do texto os diferentes tipos de informação presentes nos traços mnésicos serão designados tanto por dimensões ou por atributos.

masculina/localização direita) de forma a fornecerem informação redundante no sentido de uma determinada resposta. Esta sobreposição das dimensões resulta numa melhoria do desempenho global dos participantes, o que indica a presença da utilização conjunta de diferentes pistas de recuperação. Uma alternativa à sobreposição no estudo das dimensões é através do cruzamento destas, não existindo uma redundância em termos da informação proporcionada pelas dimensões. Johnson, Raye, Foley e Kim (1982) apresentaram auditivamente um conjunto de palavras, cuja apresentação era acompanhada por instruções para observarem o lado direito ou esquerdo do ecrã. Metade das palavras eram acompanhadas por uma figura correspondente, enquanto que nas outras palavras era pedido aos participantes que imaginassem os objectos correspondentes na posição do ecrã definida. Em seguida, os participantes executavam uma tarefa de reconhecimento, em que lhes era pedido que identificassem se a palavra estava associada a um objecto que “ocorreu” no lado direito do ecrã, no lado esquerdo, ou então se se tratava de uma palavra não apresentada anteriormente. O desempenho dos participantes na discriminação entre palavras antigas e novas demonstrou a existência de uma vantagem em relação às palavras associadas a objectos observados, porém não se verificaram quaisquer diferenças entre palavras associadas a objectos observados e imaginados na identificação da posição de “ocorrência” no ecrã. A vantagem das palavras associadas a objectos observados no julgamento de reconhecimento, em contraste com a identificação da localização vem mostrar que os indivíduos utilizam a informação proporcionada pelas diferentes dimensões de maneira distinta consoante a natureza do julgamento.

A predominância de diferentes dimensões nos julgamentos em julgamentos distintos pode ser encontrada em estudos que comparam sujeitos jovens e idosos; Chalfonte e Johnson (1996) pediram a um conjunto de participantes jovens e idosos que observassem um conjunto de objectos, cuja cor e localização variava. Existiam três condições principais de estudo (experiência 1a); era pedido aos participantes tentassem memorizar o objecto em si, que prestassem atenção à localização dos estímulos, ou então à sua cor (manipulação *inter-participantes*). O julgamento que era pedido a estes participantes correspondia à dimensão de estudo, ou seja, os participantes que prestaram atenção à cor eram testados relativamente à cor dos estímulos teste, e o mesmo sucedia para a para o reconhecimento ou identificação da localização. Os resultados obtidos demonstraram que tanto os jovens como idosos têm desempenhos equivalentes em julgamentos de reconhecimento⁴ e de identificação de cor, porém os idosos apresentam um desempenho significativamente inferior na identificação da localização original dos objectos. Posteriormente (experiência 1b) testou-se o efeito da codificação simultânea da presença e localização dos estímulos; ou então da presença e cores destes. Os julgamentos executados na fase de teste

⁴ Estes resultados contrastam com a noção de que os idosos são mais vulneráveis ao falso reconhecimento, contudo uma análise global dos diversos estudos efectuada por Gallo (2006) mostra que as diferenças médias não apresentam uma magnitude tão elevada como se supunha.

correspondiam às duas dimensões enfatizadas na fase de estudo. Aquilo que os dados obtidos vieram mostrar foi que os idosos eram significativamente piores em ambos os julgamentos, o que vem indiciar dificuldades por parte dos idosos em julgamentos que implicam a agregação de dimensões, mesmo quando avisados previamente na fase de estudo. Esta dificuldade foi ainda superior quando o tipo de aprendizagem foi intencional do que quando foi incidental (Experiência 3), ou seja, quando instruídos previamente acerca do tipo de estudo que devem fazer e do teste a executar subsequentemente, os idosos manifestam um desempenho inferior. Estes resultados vêm sugerir que há uma degradação diferencial nas várias dimensões ao longo do envelhecimento, e que os idosos apresentam grandes dificuldades na regulação estratégica destas dimensões de forma a otimizar o seu desempenho. A presença destas dificuldades ao nível de dimensões específicas representa um problema para a recollecção, dado que uma recordação de natureza episódica deveria permitir o acesso a todas estas dimensões.

Uma possibilidade explicativa destes resultados seria a presença de um efeito de interferência ou inibição provocado pela existência de dimensões mais acessíveis, porém um estudo posterior de Starns e Hicks (2005), veio sugerir que a recuperação de cada uma das dimensões na fase de teste é independente da acessibilidade das outras. Starns e Hicks (2005) procuram entender o efeito que a reutilização de uma dimensão específica (tamanho ou localização) na apresentação das palavras na fase de teste poderia ter na recuperação de outras dimensões (Experiências 1a e 1b). Os resultados vieram mostrar que a reutilização de uma dimensão específica não teve qualquer efeito no desempenho dos participantes. Foi igualmente testada a possibilidade de que a produção de julgamentos em relação a uma dimensão específica pode influenciar julgamentos subsequentes em relação a outras dimensões (Experiência 2). Procurou-se verificar a presença de um efeito de facilitação ou inibição quando os julgamentos eram feitos para todas as palavras numa única dimensão, e em seguida repetia-se todo o processo em relação a outra dimensão (*condição de monitorização separada*); ou então, para cada palavra, os participantes executavam os julgamentos para cada uma das dimensões de forma sequencial (*condição de monitorização simultânea*). Em ambas as condições experimentais não se verificou qualquer efeito do julgamento anterior, independentemente destes julgamentos estarem correctos ou incorrectos.

Aquilo que se pode extrair destes estudos é que os indivíduos conseguem utilizar de uma forma estratégica e independente as diferentes dimensões de informação contidas nos traços mnésicos; esta capacidade implica uma regulação de acordo com a natureza do julgamento e da informação disponível em memória, a qual não é contemplada pelos modelos de memória. O componente de familiaridade postulado pelo modelo globalista não prevê que a informação utilizada seja regulada, dado que a sua diferenciação não é contemplada; o componente de recollecção também não permite explicar estes resultados, dada a independência das dimensões,

segundo os pressupostos do componente de recolção, quando um julgamento ao nível de uma dimensão fosse correcto, a probabilidade de um julgamento correcto em relação a outra dimensão deveria ser bastante elevada. O acesso independente às dimensões leva a que se considere uma “reabilitação” da noção de recolção parcial (Mulligan & Hirshman, 1997; Yonelinas & Jacoby, 1996). De facto, existem estudos em monitorização da fonte que utilizam a recolção parcial como conceito central; Dodson, Holland e Shimamura (1998) apresentaram a participantes listas de palavras auditivamente, sendo que estas podiam ser apresentadas por quatro vezes, duas delas femininas, e as outras duas masculinas. Segundo o modelo dualista, um julgamento deste tipo deve-se basear na recolção, a qual deve permitir simultaneamente identificar o sexo da voz, bem como a voz específica. Contudo, aquilo que os resultados vieram demonstrar foi que em muitos casos, os participantes conseguiam identificar o sexo da voz associada à palavra de teste, porém não conseguiam recordar da voz específica.

A incapacidade dos modelos de memória em explicar este tipo de resultados não ocorre por acaso, e é um reflexo da forma como a memória é abordada pelos modelos em geral. A memória humana tem sido tradicionalmente abordada segundo dois critérios distintos: a quantidade de memórias recuperadas e a regulação da acuidade das memórias (Roediger, 1980). Nestas duas linhas, a literatura tem-se focado tradicionalmente na quantidade de informação armazenada e na capacidade dos indivíduos em recuperá-la, assumindo assim uma concepção da memória enquanto *armazém*. Por outro lado, a ênfase na regulação da acuidade assume uma noção da memória enquanto um *processo* cuja fidelidade necessita ser monitorizada e regulada. A diferença fundamental entre as concepções de memória destas duas linhas pode ser descrita do seguinte modo; no primeiro caso o objecto de estudo é uma memória *que podemos contar*, enquanto que no segundo caso estamos perante uma memória *com a qual podemos contar* (Koriat & Goldsmith, 1996).

Uma das principais motivações por detrás do desenvolvimento das abordagens quantitativas da memória é a objectividade e a fidelidade proporcionadas pelas medidas disponíveis. Contudo, estas vantagens são contrabalançadas pela inacessibilidade a determinados aspectos que são fundamentais à descrição dos processos cognitivos subjacentes ao desempenho dos indivíduos. Outro aspecto das abordagens quantitativas é o facto de que no desenvolvimento de modelos, determinados pressupostos que deveriam ser restritos a um determinado domínio, são promovidos a um estatuto de pressupostos conceptuais globais. Um exemplo desta situação pode ser encontrado na própria TDS, que ao consistir numa especificação da Teoria de Decisão Estatística (Pastore, Crawley, Berens, & Skelly, 2003), considera processos de decisão apenas com base nas propriedades estatísticas das evidências disponíveis, e não na origem e qualidades destas. Tal como na inferência estatística em geral, a TDS não assume qualquer pressuposto em relação à

natureza das evidências ou dos processos que as originam (sejam eles perceptivos, semânticos, etc.); a TDS é fundamentalmente um modelo de decisão e não um modelo processual. Daí que a sensibilidade discriminativa do decisor, medida por d' , não permita distinguir a origem dessa discriminação; a origem dos valores que são obtidos por d' não de todo é interpretada pela TDS, estes tanto podem ser resultado da força dos traços mnésicos dos itens a recordar, como da capacidade dos mecanismos de monitorização e regulação em distinguir os *estímulos-alvo* dos *estímulos-ruído*. Ambas são interpretações igualmente válidas de d' (Goldsmith & Koriat, 2008).

A literatura que aborda a monitorização da fonte vem mostrar que os principais modelos de memória assumem simplificações dos processos de regulação e recuperação mnésica que não se coadunam com os resultados encontrados. Estas limitações têm-se tornado cada vez mais gritantes na medida em que os mecanismos de monitorização assumem uma predominância cada vez maior em diversas áreas de estudo, especialmente naquelas em que são obtidos bastantes erros de memória.

Curiosamente, as limitações encontradas nos modelos de memória, nomeadamente ao nível da indiferenciação dos diversos tipos de informação, não se verificam em outros modelos bastante próximos, se não mesmo parcialmente sobrepostos – os modelos de categorização (Estes, 1994). Nestes modelos, os estímulos são caracterizados por uma série de atributos, os quais apresentam uma determinada diagnosticidade relativamente a uma ou outra categoria, servindo então como base para as respostas dos indivíduos. A diagnosticidade dos diversos atributos pode ser alterada de forma independente, dependendo do feedback que o indivíduo recebe em relação a cada uma das suas respostas numa tarefa de categorização (e.g., Gluck & Bower, 1988). Infelizmente, o estudo de aspectos como a diagnosticidade diferencial dos atributos nos modelos de categorização não levou a um enriquecimento dos modelos de memória, dado que a noção que a literatura sobre categorização tem da memória de reconhecimento é de que esta consiste num tipo de categorização “pobre”:

In a typical recognition study, subjects view a sequence of stimulus patterns, then are tested with a mix of previously seen and novel patterns with the task of classifying each as “old” or “new”. Here class membership carries no information about the properties of the stimulus (except time of occurrence), but otherwise the task is analogous to category learning (Estes, 1994, p.5)

Tal como na literatura sobre modelos de memória, o reconhecimento é visto como um julgamento de estímulos, em que os diferentes atributos dos traços mnésicos não são considerados de forma distinta. Como se poderá verificar em seguida, esta noção é uma simplificação errada,

fruto de uma abordagem à memória de reconhecimento que negligencia a natureza dos erros de memória. Quando ocorrem erros de memória de forma sistemática, e estes erros têm por base a avaliação da presença de determinados tipos de informação, verifica-se que a diagnosticidade dos diferentes “atributos” dos traços é considerada pelos indivíduos.

O problema da monitorização: Falsas memórias associativas

A literatura que aborda directamente o desempenho dos indivíduos em memória de reconhecimento, raramente considerou a questão dos erros de memória como um objecto de estudo que merecesse qualquer tipo de estudo sistemático. Existiu sempre uma ênfase no reconhecimento verídico, negligenciando-se as falsas memórias, e encarando-as literalmente como resultado da presença de “ruído” nos processos de recordação (Roediger, 1996). As bases para o estudo específico das falsas memórias surgiram de forma muito discreta no final da década de 50 num artigo de James Deese (1959). Neste estudo, Deese procurou entender a razão pela qual os indivíduos se recordavam sistematicamente de palavras não apresentadas anteriormente, aquando de um teste de recordação livre. A hipótese considerada foi que estes erros de memória se deviam às relações associativas existentes entre as palavras nas listas de estudo e as palavras intrusivas. Dada esta hipótese, foi apresentado a um grupo de participantes listas de palavras (e.g., *SONO*, *DORMIR*, *CAMA*), sendo que as palavras de cada uma das listas apresentavam ligações associativas a uma palavra não apresentada, designada por palavra-crítica (e.g. *SONHO*). Após a apresentação de cada uma das listas era pedido aos participantes que executassem uma tarefa de recordação livre, procurando recordar as palavras anteriormente apresentadas na lista. Os resultados obtidos por Deese vieram mostrar que as relações associativas entre as listas e as palavras-críticas (designadas por força associativa retrógrada – BAS) estavam na base deste efeito de falsas memórias, verificando-se uma elevada correlação (.87) entre a BAS de cada lista e a recordação errónea das palavras-críticas (Deese, 1959).

Este artigo de Deese manteve-se na obscuridade até à sua redescoberta em meados da década de 90 com a publicação de uma investigação de Henry L. Roediger e Kathleen McDermott (1995). Nesta investigação, os autores adaptaram o procedimento de Deese, dando origem a um paradigma experimental posteriormente designado por paradigma DRM (Deese-Roediger-McDermott). No paradigma DRM é apresentado aos participantes listas de 15 palavras associadas a uma palavra-crítica não apresentada; após a apresentação de cada uma destas listas é pedido aos participantes que executem uma tarefa de recordação livre, ou então que executem uma tarefa distractora (manipulação inter-participantes). Após a apresentação de todas as listas de palavras, os

participantes executam uma tarefa de reconhecimento, cuja lista inclui não apenas as palavras apresentadas, mas também palavras não apresentadas como as palavras-críticas. Para além das resposta sim/não, os participantes indicavam qual a natureza fenomenológica da recordação de cada uma das palavras (Gardiner, 1988; Tulving, 1985), sendo-lhes pedido que indicassem se a sua resposta se baseava numa recordação da experiência episódica (resposta “remember”), ou então esta tem por base uma familiaridade indistinta da sua ocorrência (resposta “know”). Os resultados obtidos por Roediger e McDermott vieram mostrar que os indivíduos recordavam mais frequentemente as palavras-críticas (55%) do que recordavam palavras apresentadas a meio da lista de estudo (47%). Na tarefa de reconhecimento, as palavras-críticas foram julgadas como anteriormente apresentadas na mesma proporção que as palavras da lista, sendo que estes falsos reconhecimentos tinham associados a si uma recordação episódica (resposta “remember”) em 72% das listas. Verificou-se ainda que o falso reconhecimento era superior quando os sujeitos tinham manifestado falsas memórias na tarefa de recordação livre anterior, o que revela um efeito potenciador da recordação prévia.

O paradigma DRM veio colocar as falsas memórias no centro do estudo da memória de reconhecimento, tendo produzido dezenas de artigos desde a sua introdução (para uma revisão ver Gallo, 2006). O efeito de falsas memórias tem sido encontrado mesmo quando as listas de estudo são apresentadas de forma repetida (apesar de se verificar uma diminuição do efeito; Benjamin, 2001; McDermott, 1996), ou quando a apresentação das listas é feita subliminarmente, não existindo uma activação consciente da palavra-crítica (Seamon, Luo, & Gallo, 1998; McDermott & Watson, 2001). De facto, o efeito das listas DRM na activação da palavra-crítica verifica-se mesmo em medidas de memória implícita (McDermott, 1997). Dada a sua natureza associativa, o efeito de falsas memórias é moderado pelo número de palavras associadas apresentadas em cada lista (Robinson & Roediger, 1997), sendo a força associativa das palavras da lista com a palavra-crítica – a BAS, o melhor preditor da ocorrência de falsas memórias (Roediger, Watson, McDermott, & Gallo, 2001).

Quanto a manipulações experimentais ao nível do teste de memória, estudos mostram que a repetição da fase de teste sem que hajam fases de estudo adicionais leva a um aumento das falsas memórias (Payne, Elie, Blackwell, & Neuschatz, 1996, Experiência 2); a múltipla apresentação das palavras na fase de teste aumenta a sua familiaridade, o que leva a uma maior dificuldade por parte dos participantes em identificar os estímulos, ou a lembrarem-se de que respostas deram anteriormente. Este aumento também é encontrado quando é prolongado o intervalo existente entre a fase de estudo e a fase de teste, verificando-se um aumento da probabilidade de recordação de memórias falsas, probabilidade essa superior à da recordação verdadeira (Payne et al., 1996, Experiência 1); este fenómeno deve-se ao facto de que as palavras-

críticas, para além de serem em média quase tão familiares como as palavras apresentadas, são activadas sempre que uma palavra apresentada surge na fase de teste. Esta situação leva a que a familiaridade das palavras-críticas não sofra os mesmos decréscimos verificados na recordação verídica.

Um dos aspectos principais deste efeito de falsas memórias é o facto de os erros de memória não serem resultado de um critério de resposta liberal por parte dos participantes. Os participantes manifestam falsas memórias mesmo quando são previamente avisados do efeito (Gallo, Roberts, & Seamon, 1997; McDermott & Roediger 1998), o que demonstra que estamos perante um efeito robusto e não um mero artefacto experimental, e a recordação da palavra crítica é acompanhada por níveis de confiança bastante elevados (Read, 1996). Outro aspecto surpreendente é a natureza episódica destes falsos reconhecimentos; estas falsas memórias são acompanhadas por elevados níveis de confiança, com os participantes a considerarem-se capazes de indicar uma suposta posição da palavra-crítica na lista de aprendizagem (Read, 1996), quando são utilizadas listas auditivas, em indicar qual das diferentes vozes a pronunciou na fase inicial (Payne et al., 1996), ou mesmo atribuir características físicas e espaciais específicas aos estímulos erroneamente reconhecidos (Lyle & Johnson, 2006).

A manifestação de falsas memórias episódicas levanta questões relativamente ao tipo de informação que as constitui. O facto de que as falsas memórias ocorrem de forma tão sistemática implica que estas sejam constituídas por algum tipo de informação que leva os indivíduos a aceitá-las como verdadeiras, porém, dado que estas se referem a eventos não ocorridos, devem necessariamente existir diferenças relativamente às memórias verídicas. Os estudos que se debruçaram sobre este problema (Mather, Henkel, & Johnson 1997; Norman & Schacter, 1997) sugerem que as falsas memórias são maioritariamente constituídas por informação de natureza semântica e associativa, ao contrário das memórias verídicas que são mais ricas em termos de conteúdo, incluindo igualmente informação de natureza perceptiva e contextual. Estas diferenças na constituição dos traços mnésicos revelaram-se diagnósticas na discriminação dos estímulos quando era pedido aos participantes que avaliassem o conteúdo das suas memórias, o que levou a uma maior rejeição das palavras-críticas e de distractores associados (Mather, Henkel, & Johnson, 1997).

O efeito de falsas memórias não se restringe a listas de palavras associadas, podendo ocorrer igualmente com listas de vizinhos fonológicos (e.g. PALA, TALA, BALA). Quando estas listas são apresentadas, há uma activação ao longo da vizinhança fonológica (Luce & Pisoni, 1998), o que leva a erros de memória na fase de teste, nomeadamente a falsa recordação da palavra no centro da vizinhança fonológica (e.g. MALA) (Sommers & Lewis, 1999). Este efeito de falsas memórias verificou-se mesmo quando era pedido aos participantes que indicassem a fonte (que voz

as pronunciou) das palavras (Sommers & Lewis, 1999, Experiência 2), o que demonstra semelhanças com os resultados de Payne et al. (1996). As falsas memórias fonológicas também manifestam um decréscimo quando a semelhança fonológica entre as palavras diminui, efeito que pode ser considerado equivalente ao encontrado ao nível da força associativa das palavras da lista e a palavra-crítica – a BAS.

O estudo das falsas memórias levou à elaboração de modelos teóricos que procuram explicar a totalidade do fenómeno. O modelo teórico dominante é o modelo Activação-Monitorização⁵ (Roediger, Watson, McDermott, & Gallo, 2001), o qual assume que o processo de recordação é constituído por dois componentes que operam de forma antagónica. O primeiro componente do modelo corresponde aos processos de activação presentes na codificação das palavras associadas durante a fase de estudo. Este componente assume a existência de um processo de dispersão da activação de (Collins & Loftus, 1975), no qual a activação de um item provoca a activação automática e inconsciente de outros itens na sua vizinhança semântica. Dado que as palavras de cada lista estão associadas a uma palavra-crítica, esta última deve ser repetidamente activada durante o processo de aprendizagem das listas, o que resulta na sua sobreactivação. Esta sobreactivação vai ter impacto na fase de teste, levando a que a palavra-crítica seja confundida com as palavras originalmente presentes na lista.

O segundo componente do modelo corresponde aos mecanismos de monitorização da fonte (Johnson, Hashtroudi, & Lindsay, 1993). Estes dizem respeito à capacidade dos sujeitos em discriminarem a origem de um determinado traço mnésico, e aos processos de decisão durante a recuperação que determinam essa origem, dado que informação relativamente a esta não se encontra necessariamente disponível no próprio traço. Os processos de decisão podem ser mais sistemáticos, em que é avaliada a plausibilidade da origem da memória, ou então um processo de decisão mais heurístico, em que aquilo que é avaliado para determinar se uma memória é verdadeira é um critério de familiaridade ou então a mera presença de um determinado tipo de informação. Deste modo, as falsas memórias ocorreriam devido a falhas na monitorização, em conjunto com a sobreactivação das palavras-críticas.

Segundo o modelo Activação-Monitorização, a formação das falsas memórias tem como base fundamental a dispersão de activação produzida na fase de estudo das listas de palavras associadas, o que levanta questões acerca dos efeitos que o tipo ou nível de processamento envolvido na fase de estudo poderá ter na sua ocorrência. Dado que um processamento mais “profundo” ou semântico leva geralmente a um aumento do desempenho dos indivíduos em termos

⁵ O modelo alternativo existente, a Fuzzy-Trace Theory (e.g., Reyna & Brainerd, 1995) não vai ser considerado nesta introdução visto ser um modelo cuja explicação não consegue incluir diversos fenómenos de falsas memórias, como as falsas memórias fonológicas (Sommers & Lewis, 1999) ou as falsas memórias obtidas com listas polissémicas (Hutchinson & Balota, 2005).

de recordação verídica (Craik & Lockhart, 1972; Craik & Tulving, 1975), de que forma o tipo de processamento afecta as falsas memórias? De facto, tal como ocorre com recordação verídica, a manifestação de falsas memórias é influenciada pelo tipo de processamento estabelecido pelos participantes na fase de estudo. Diversos estudos (Rhodes & Anastasi, 2000; Thapar & McDermott, 2001; Tolia, Neuschatz, & Goodwin, 1999) incorporaram no paradigma DRM manipulações do nível de processamento e verificaram que a ocorrência de falsas memórias em relação às palavras-críticas aumenta em conjunto com a recordação verídica quando as listas foram processadas de um modo “profundo” (e.g. julgamento de agradabilidade)⁶. O padrão inverso foi encontrado quando os participantes processaram superficialmente (e.g. contagem de sílabas), havendo uma diminuição conjunta das recordações verídicas e falsas. Este efeito dos níveis de processamento foi encontrado tanto na tarefa de recordação livre (Rhodes & Anastasi, 2000; Tolia, Neuschatz, & Goodwin, 1999), como na tarefa de reconhecimento (Thapar & McDermott, 2001⁷). Na literatura de falsas memórias assume-se então que processamento “profundo”, ao enfatizar a activação das redes associativas leva a uma sobreactivação das palavras críticas, se traduz num aumento da probabilidade da sua recordação errónea. Pelo contrário, o processamento superficial leva a que a rede associativa não seja tão activada, o que reduz necessariamente o nível de familiaridade das palavras críticas.

Este efeito de “níveis de processamento” apresenta ainda, uma interacção com o tipo de “material” que constitui as listas de palavras. Quando são utilizadas listas fonológicas, o nível de processamento tem um efeito inverso, as recordações falsas e verídicas aumentam com o processamento superficial, e diminuem com o processamento profundo (Chan, McDermott, Watson, & Gallo, 2005). Isto vem mostrar do mesmo modo que as recordações verídicas, a formação de falsas memórias é dependente da correspondência entre o material e o tipo de processamento, reforçando o princípio defendido por Tulving e Thomson (1973), de que a forma como a informação é recuperada depende da forma como esta foi inicialmente codificada. Contudo, a correspondência entre material e o tipo de processamento apenas explica este efeito ao nível da activação, estando igualmente presente uma componente de monitorização. De facto, a não correspondência entre material e o tipo de processamento facilita a monitorização dos estímulos, já que a informação codificada é extremamente diagnóstica da presença do estímulo na fase de estudo. O processamento semântico permite uma melhor discriminação das palavras críticas em

⁶ A utilização de julgamentos de agradabilidade será ao longo deste texto referida por *processamento profundo*, enquanto que as tarefas perceptivas como *processamento superficial*.

⁷ Os resultados encontrados por Thapar & McDermott (2001), apesar de serem aceites pela literatura em geral como evidências de um efeito de nível de processamento no falso reconhecimento, apresentam alguns problemas metodológicos que são considerados no Anexo 1. Daí que as conclusões gerais deste estudo e toda a questão dos níveis de processamento deva ser considerada com alguma cautela, nomeadamente a relação causal que é inferida entre o nível de processamento e as falsas memórias.

listas de associados fonológicos, verificando-se o mesmo para o processamento perceptivo de listas de palavras associadas.

Esta utilização estratégica de determinados tipos de informação consoante o material utilizado apresenta contudo limitações; Watson, Balota, e Roediger (2003) utilizaram listas mistas, compostas por palavras associadas e por vizinhos fonológicos, e verificaram que o número de erros cresceu drasticamente, revelando uma sobreaditividade das falsas memórias. Segundo os autores, este efeito deve-se ao facto dos participantes não conseguirem estabelecer um critério de decisão adequado dada a presença simultânea de estímulos cujo tipo de erros que produzem têm naturezas distintas.

O facto de que determinados tipos de informação podem ser altamente diagnósticos da apresentação prévia de um estímulo leva a que esta informação possa ser implementada em estratégias de decisão simples e frugais como as heurísticas. Um exemplo da utilização de heurísticas em julgamentos de reconhecimento ocorre quando são apresentadas figuras na fase de estudo. Israel e Schacter (1997) apresentavam aos sujeitos listas de figuras acompanhadas pelas palavras correspondentes, ou então listas de apenas palavras (manipulação inter-participantes). Na fase de teste, de forma a compreender o papel do estímulo pictórico na tarefa de reconhecimento, os sujeitos de ambas as condições experimentais efectuaram testes de reconhecimento em que lhes eram apenas apresentadas as palavras, ou então eram apresentadas figuras acompanhadas pelas palavras correspondentes (manipulação inter-participantes). Os resultados (experiência 1) mostraram que o reconhecimento verídico em ambas as condições de estudo era semelhante, no entanto verificou-se uma redução drástica do falso reconhecimento na condição em que os sujeitos eram expostos a figuras em relação à condição em que eram apenas apresentadas palavras. Estes resultados indiciam que os participantes estão a utilizar a presença de informação pictórica como critério de decisão; como a recordação das palavras-críticas e outros distractores associados não evoca qualquer tipo de informação de natureza pictórica, os indivíduos conseguem facilmente rejeitá-los. Esta estratégia, posteriormente designada por heurística da distintividade, verifica-se apenas quando a presença de informação pictórica é um critério diagnóstico. Quando a apresentação de figuras é manipulada intra-participantes, a utilização da heurística da distintividade enquanto estratégia de decisão deixa de fazer sentido dado que a presença de informação pictórica não é diagnóstica. De facto, quando são apresentadas listas em que apenas parte das palavras é acompanhada por uma figura correspondente, a proporção de falsos alarmes manifestada pelos participantes aumenta drasticamente, em relação à condição em que todas as palavras da lista incluem figuras (Schacter, Israel, & Racine, 1999). A utilização da heurística da distintividade é igualmente comprometida quando é pedido aos participantes que dêem as suas respostas muito

depressa (Dodson & Hege, 2005), dado que a recuperação de informação episódica, como é o caso da informação pictórica, é mais lenta (Johnson, Kounios, & Reeder, 1994).

Quanto à capacidade dos modelos em explicar os efeitos de falsas memórias, esta revela-se à partida bastante boa, contudo ambos os tipos de modelos apresentam dificuldades ao nível dos mecanismos de monitorização. Em termos gerais, as falsas memórias podem ser explicadas apenas pela TDS. O efeito de falsas memórias é descrito pela TDS como o resultado de uma distribuição de estímulos-ruído cuja sobreposição com a distribuição de estímulos-sinal é bastante grande. As palavras associadas, dada a sua grande familiaridade, têm uma grande probabilidade de serem reconhecidas, mesmo quando é assumido um critério de resposta mais estrito. Os efeitos obtidos através de manipulações ao nível do processamento da informação são igualmente previstos pela TDS no sentido em que a informação codificada pode levar a uma maior ou menor discriminabilidade dos estímulos na fase de teste; outro aspecto que pode ser interpretado à luz dos modelos baseados na TDS é a diminuição da sensibilidade discriminativa dos indivíduos.

Enquanto que a manifestação de falsas memórias é facilmente explicada pela TDS, esta apresenta algumas limitações no que diz respeito ao processo de monitorização, dado que, como já foi referido anteriormente, a medida de sensibilidade discriminativa d' não permite distinguir entre a força dos traços mnésicos e a sua monitorização (Goldsmith & Koriat, 2008). Algumas destas limitações são suplantadas pelos modelos dualistas, que procuram explicar efeitos de supressão de falsas memórias como a heurística da distintividade (Israel & Schacter, 1997; Schacter, Israel, & Racine, 1999), através do processo de recolção. Contudo, os mecanismos de monitorização encontrados na literatura de falsas memórias apresentam uma complexidade e variabilidade no modo como se podem manifestar, algo que não é compatível com a simplicidade de um processo de recolção. Um exemplo é o estudo de Mather, Henkel, e Johnson (1997), no qual era pedido aos participantes que avaliassem os diferentes tipos de informação nos traços mnésicos, através do questionário MCQ (*Memory Characteristics Questionnaire*). Neste questionário, é pedido aos participantes que avaliem as suas memórias em termos da presença de informação perceptiva, associativa, ou contextual. A utilização deste questionário levou a que os participantes indicassem em relação às palavras-críticas uma maior presença de informação associativa em detrimento de informação perceptiva e contextual, o que contrastava com a informação perceptiva e contextual presente em memória, relativamente às palavras anteriormente apresentadas na fase de estudo. Estas diferenças obtidas com o MCQ levaram a que os participantes melhorassem o desempenho, suprimindo o falso reconhecimento das palavras-críticas. Este resultado vem contrariar o pressuposto de que a informação é representada numa única dimensão indiferenciada, bem como a noção de que a informação de natureza episódica tem de obedecer a um processo de recuperação de alto-limiar, e não a uma distribuição de evidências.

Outro aspecto que causa problemas aos modelos de memória em geral é o papel que a informação adicional fornecida aos participantes, quer seja através de avisos prévios (*warnings*), quer através de *feedback*, pode ter no seu desempenho. Num estudo desenvolvido por Starns, Lane, Alonzo, e Roussel (2007), foi investigado o impacto que o aviso prévio tinha no desempenho dos indivíduos no paradigma DRM. Estudos anteriores (Gallo, Roberts, & Seamon, 1997; McCabe & Smith, 2002; Neuschatz, Payne, Lampinen, & Togliola, 2001) haviam mostrado resultados ambíguos relativamente à capacidade dos avisos prévios na melhoria do desempenho. Os autores testaram então o efeito dos avisos prévios utilizando a TDS, e verificaram que estes levam a melhorias na sensibilidade discriminativa dos participantes (d'), mesmo quando a informação era dada após a codificação das palavras. Estes resultados vêm mostrar que os indivíduos conseguem não apenas utilizar a informação contida nos avisos prévios para regular o modo como codificam as palavras na fase de estudo, mas também conseguem seleccionar de forma estratégica as evidências utilizadas no julgamento de familiaridade dos traços mnésicos.

Outras formas de informação, como o feedback, têm sido abordadas com o propósito de compreender o seu papel no desempenho dos indivíduos. Jou & Foreman (2007) procuraram perceber o impacto que o *feedback imediato*, ou seja, *feedback* dado imediatamente após cada resposta do participante, poderia melhorar o desempenho dos participantes. A hipótese dos autores era que como as palavras e os itens críticos são constituídos por tipos de informação qualitativamente distintos (Jou et al., 2004; Mather, Henkel, & Johnson, 1997; Norman, & Schacter, 1997), o *feedback* imediato deverá levar a que o sujeito consiga avaliar a diagnosticidade das pistas mnésicas nas quais está a basear as suas decisões, e deste modo calibrar a monitorização de forma a melhorar o seu desempenho na tarefa de reconhecimento. Assim, se os sujeitos executarem múltiplos blocos de *estudo-teste*, estas melhorias devem-se reflectir não apenas ao nível dos ensaios individuais na primeira lista, mas também nas listas de teste subsequentes, verificando-se uma melhoria progressiva. Neste estudo, comparou-se o desempenho dos participantes em diferentes condições experimentais, sendo que numa delas era dado *feedback* imediato após cada resposta na tarefa de reconhecimento, enquanto noutra condição não era dado qualquer feedback. Os resultados obtidos mostraram que a proporção de falsos alarmes diminuiu progressivamente ao longo dos ensaios na condição de feedback, ao contrário do que se verificou na condição de controlo, na qual a proporção de falsos alarmes se manteve elevada e relativamente constante.

As falsas memórias produzidas pelo paradigma DRM vêm demonstrar a natureza inferencial dos julgamentos de memória, e a importância que os diferentes tipos de informação presentes nos traços mnésicos podem ter enquanto evidências diagnósticas nos julgamentos de memória em geral. Outro aspecto importante é o facto de que os diferentes tipos de informação podem ser utilizados de maneiras distintas, quer seja numa avaliação mais sistemática (Mather,

Henkel, & Johnson, 1997), quer seja através de heurísticas (Schacter, Israel, & Racine, 1999). Deste modo, o paradigma DRM, ao permitir a produção de falsas memórias tanto de natureza associativa como perceptiva, apresenta-se como uma ferramenta importante no estudo da capacidade de monitorização e regulação dos conteúdos mnésicos. Quanto aos modelos de memória, estes conseguem explicar o fenómeno das falsas memórias, e alguns mecanismos de supressão, contudo são incapazes de incorporar outras formas de monitorização que se baseiam na avaliação de cada um dos tipos de evidências disponíveis nos traços mnésicos, bem como a reavaliação da diagnosticidade destes tipos de informação aquando da presença de informação adicional.

Como será demonstrado em seguida, o facto de os modelos de memória não contemplarem os processos de monitorização não se limita à integração da literatura de monitorização da fonte ou de falsas memórias associativas. De facto, esta reflecte-se mesmo na literatura que se centra na modelação, a qual não consegue propor soluções imediatas para os problemas encontrados.

O problema da monitorização: Tomada de decisão e julgamentos de confiança

Um aspecto a ter em conta no estudo dos mecanismos dos mecanismos de monitorização é a relação existente entre a acuidade dos indivíduos e o nível de confiança que estes atribuem às suas respostas. A importância deste aspecto deriva do facto de que os julgamentos de confiança devem ser um indicador da eficácia da monitorização presente na tarefa de reconhecimento. Como se pôde verificar anteriormente, os indivíduos podem manifestar níveis de confiança elevados no falso reconhecimento de palavras críticas, o que demonstra a presença de falhas ao nível da monitorização (Read, 1996; Roediger & McDermott, 1995). No entanto, esta dissociação entre acuidade e confiança pode não se limitar a apenas um conjunto limitado de palavras, podendo verificar-se de forma recorrente níveis superiores de confiança em simultâneo com decréscimos na acuidade global dos indivíduos. A ocorrência deste tipo de dissociações é bastante informativa, no sentido em que permite estudar de que modo os processos envolvidos na tomada de decisão são os mesmos por detrás dos julgamentos de confiança, e até que ponto poderão a estar a ser utilizadas diferentes fontes de informação.

Na literatura de reconhecimento de faces a dissociação entre confiança e acuidade é um fenómeno bastante comum, existindo casos em que a acuidade dos indivíduos diminui enquanto a sua confiança no seu desempenho aumenta, e vice-versa (Chandler, 1994; Clark, 1997; Lindsay, Read & Sharma, 1998). Um exemplo recente desta dissociação na literatura pode ser encontrado no

trabalho de Busey, Tunnicliff, Loftus e Loftus (2000), os quais apresentaram a um conjunto de participantes listas de faces, sendo manipulada a luminosidade destas na fase de estudo e de teste. Quando as faces eram apresentadas numa condição de *luminosidade reduzida* e eram posteriormente testadas numa condição de *luminosidade elevada*, verificou-se que a acuidade dos participantes no reconhecimento era inferior à do grupo que era testado em condições de *luminosidade reduzida*, no entanto, os níveis de confiança manifestados eram superiores.

A relação inversa entre confiança e acuidade é algo problemático para um dos pressupostos tradicionais da TDS, segundo o qual os níveis de confiança que os sujeitos atribuem às suas decisões correspondem a uma partição ao nível da dimensão de familiaridade, sendo este funcionalmente equivalente ao tipo de partição que ocorre no estabelecimento do critério de resposta (Murdock, 1965). Assim, um critério mais liberal corresponderia a um nível de confiança mais baixo e vice-versa (ver Figura 4).

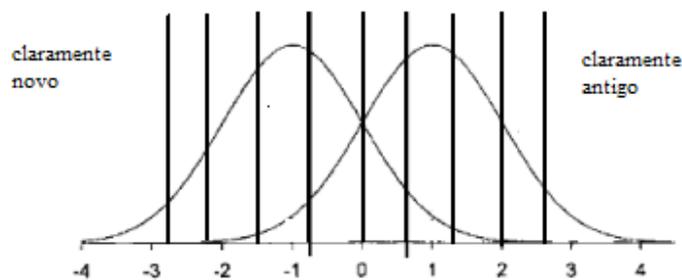


Figura 4. Segmentação das distribuições através de uma escala de avaliação. Adaptado de Creelman & Macmillan (2005).

Esta relação permite que se obtenham curvas ROC utilizando níveis de confiança ao longo de uma única sessão de teste, ao invés das múltiplas sessões de teste que a manipulação do critério de resposta implica. Este pressuposto tem como consequência a noção de que a presença de um julgamento de confiança não deve afectar o processo de decisão em si, nem o critério subjacente ao processo de decisão deve produzir qualquer tipo de efeito ao nível dos julgamentos de confiança.

O pressuposto de equivalência foi testado por Van Zandt (2000), num estudo em que era manipulado o critério de resposta dos participantes (através de *payoffs* e *base-rates*) enquanto eram pedidos julgamentos de confiança aos participantes. A existência de uma equivalência implica que os mapeamento dos níveis de confiança para cada critério na zROC se sobreponham (ver Figura 5). Porém os resultados obtidos por Van Zandt vieram mostrar que esta equivalência não se verifica, existindo diferenças ao nível do declive das rectas obtidas com os diferentes critérios de resposta. Nomeadamente, o declive das zROC obtidas com um critério mais liberal demonstrou de uma forma sistemática e significativa ser superior ao das zROC obtidas com um critério mais

conservador. Estas diferenças mostram que o pressuposto de equivalência entre os níveis de confiança e os critérios de resposta não existe, verificando-se apenas uma relação correlacional.

As diferenças ao nível do declive sugerem ainda que os participantes estão a aceder à informação mnésica de um modo distinto, podendo esta ocorrer através de uma maior proporção de informação obtida sob a forma de recollecção (Yonelinas, 1997), ou então de uma redução da variância da distribuição de sinal (Ratcliff, Shue & Gronlund, 1992). Independentemente da verosimilhança destas e de outras hipóteses (e.g., DeCarlo, 2002), o ponto fundamental resume-se ao facto de que estes resultados vêm colocar em causa a noção de que o julgamento de confiança não deve afectar as propriedades do processo de decisão, e vice-versa.

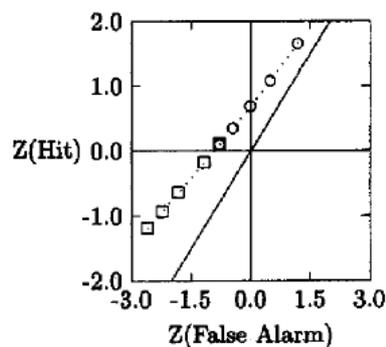


Figura 5. Previsão da TDS em relação aos z -scores da z ROC obtidos através dos níveis de confiança. Os quadrados representam os níveis de confiança na condição em que o critério de resposta é liberal, e os círculos representam os níveis de confiança na condição em que o critério é conservador. Adaptado de Van Zandt (2000).

O pressuposto de equivalência produz igualmente limitações ao nível dos tempos de resposta esperados. De acordo com este, como o processo de decisão é o mesmo que produz o julgamento de confiança; este último deve estar disponível ao mesmo tempo que a decisão, deste modo, o tempo necessário para determinar o nível de confiança deve ser constante e independente da confiança atribuída. Este pressuposto é assumido por uma classe de modelos que procuram descrever a dinâmica dos tempos de resposta (TR) em processos de tomada de decisão, designados por modelos uniprocessuais. Um exemplo desta classe de modelos é o PMM (*Probabilistic Mental Models*; Gigerenzer, Hoffrage, & Kleinbölting, 1991), segundo o qual num contexto de decisão (e.g. qual das duas cidades tem mais habitantes?), o indivíduo avalia as diferentes propriedades das alternativas (e.g. a cidade A é a capital do país). Como a validade ecológica de cada uma das propriedades deriva da sua frequência, esta representa igualmente a confiança associada a cada alternativa.

Existe contudo, uma segunda classe de modelos designados por modelos *pós-decisórios* (Juslin & Olsson, 1997; Vickers & Packer, 1982), segundo os quais o processo de decisão é

conceptualizado como um processo de acumulação de informação⁸, enquanto que o julgamento de confiança consiste numa avaliação da informação que esteve na base da decisão inicial. Estes modelos assumem que a avaliação da informação tem como base um determinado algoritmo mental que é aplicado à informação envolvida na decisão. O modelo *Balance-of-evidence* (Vickers & Packer, 1982), por exemplo, considera que a confiança atribuída a uma determinada decisão consiste na diferença entre a quantidade de informação acumulada no sentido de cada um das alternativas. Independentemente das especificidades desta classe de modelos, os processos de decisão e avaliação postulados por estes restringem igualmente os TR que devem estar associados a cada um destes. Se os processos de decisão e julgamento de confiança são independentes e ocorrem sequencialmente, então a presença de um julgamento de confiança não deve afectar o processo de decisão inicial (*SIM-NÃO*), nem este último deve afectar o TR associado ao primeiro.

Deste modo, as diferenças processuais permitem que ambos os tipos de modelos possam ser testados com base nos TR observados nas fases de resposta no julgamento de confiança subsequente; de acordo com os *modelos uniprocessuais*, não deve ocorrer qualquer processamento durante o julgamento de confiança dado que este ocorreu durante o processo de decisão, pelo que o TR do julgamento de confiança não deve variar consoante o nível de confiança atribuído. Pelo contrário, os *modelos pós-decisórios* consideram que os TR dos julgamentos de confiança deve depender dos níveis de confiança, mas não na dificuldade das decisões iniciais.

As implicações destas duas classes de modelos foram testadas por Baranski e Petrusic (1998) através de um conjunto de tarefas perceptivas, nas quais era manipulada tanto a dificuldade das tarefas como as condições de resposta, permitindo observar os efeitos do *compromisso velocidade/acuidade* (Townsend & Ashby, 1983) tanto nos TR das decisões iniciais, como nos TR dos julgamentos de confiança. Nestas tarefas perceptivas, era pedido aos participantes em cada ensaio uma decisão inicial (*SIM-NÃO*), sendo-lhes em seguida pedido que indicassem o seu nível de confiança na resposta anteriormente dada.

Aquilo que Baranski e Petrusic (1998) verificaram foi que na *condição de acuidade*, os TR das decisões iniciais variaram consoante o nível de confiança, não encontrando no entanto diferenças nos TR dos julgamentos de confiança. Na *condição de velocidade*, os TR das decisões não variaram consoante ao nível confiança, no entanto os TR dos julgamentos de confiança manifestaram diferenças consoante o nível de confiança resultante. Outro aspecto relevante diz respeito à capacidade dos sujeitos em identificar com base no julgamento de confiança as suas decisões incorrectas e levarem a cabo uma *reversão de resposta*, capacidade esta que se revelou

⁸ O princípio por detrás desta acumulação de informação consiste numa extensão dos princípios da TDS, considerando-se que há uma amostragem sequencial de informação de cada uma das distribuições segundo um processo estocástico, e que é a acumulação desta que leva à tomada de decisão (ver Ratcliff, 1978).

superior na *condição de velocidade* em relação à *condição de acuidade*. De facto, os participantes corrigiram mais respostas incorrectas quando a resposta inicial foi dada numa *condição de velocidade*. Uma possível interpretação destes resultados é que os participantes na *condição de acuidade* avaliaram a sua confiança durante o processo de decisão de forma a regular as suas respostas, enquanto que na *condição de velocidade*, os julgamentos de confiança ocorriam após os processos de decisão. Estes dois tipos de julgamento de confiança, demonstraram igualmente diferenças ao nível da sua diagnosticidade em relação aos erros de decisão cometidos, levando a que se assuma que o julgamento de confiança é um processo funcionalmente distinto e dinâmico, podendo ocorrer sob diferentes formas e em diferentes momentos consoante as características e exigências da tarefa em causa.

Os resultados de Baranski e Petrusic (1998) vêm deste modo pôr em causa a validade das duas classes de modelos; apesar dos resultados irem ao encontro dos modelos pós-decisórios, ou seja, da presença de processos distintos e independentes. No entanto, a lentificação geral das respostas na *condição de acuidade* leva a que pelo menos parte do julgamento de confiança ocorre em paralelo com o processo de decisão, ou então corresponde a este. O processos subjacentes aos julgamentos de confiança são assim apresentados como processos dinâmico cuja implementação varia consoante as exigências da tarefa em causa e cuja presença tem um impacto considerável no processo de decisão em si. Uma interpretação plausível é que a informação acumulada durante o processo de decisão é monitorizada e avaliada ao longo de todo o processo, produzindo um julgamento de confiança que evolui com o decorrer da acumulação de informação.

As questões relativas às reversões de resposta foram exploradas no domínio da memória de reconhecimento por Van Zandt & Maldonado-Molina (2004), que para além de replicarem o efeito de reversão de resposta encontrado por Baranski e Petrusic (1998), manipularam alguns aspectos da decisão inicial, nomeadamente o critério de resposta utilizado pelos participantes (através das *base-rates*). A relação entre a reversão de respostas e o critério de resposta é relevante para a validação dos *modelos pós-decisórios* (e.g. *Race model*, Townsend & Ashby, 1983, ver Figura 6), os quais são utilizados no estudo dos julgamentos de confiança em memória de reconhecimento. Segundo os modelos pós-decisórios, o critério de resposta estabelecido influencia a ocorrência de reversões de resposta, sendo mais frequente quando a resposta inicial é concordante com o enviesamento do critério. O *Race Model* assume que existem dois processos de acumulação de informação simultâneos e opostos (X_n e X_o); enquanto um dos processos acumula informação no sentido da resposta “antiga”, o outro acumula no sentido da resposta “nova”. Cada um destes processos tem associado um limiar de resposta específico (K_n e K_o), que quando alcançado, desencadeia a tomada de decisão e a manifestação da resposta do indivíduo. Este modelo assume também que os julgamentos de confiança têm como base uma comparação entre a quantidade de

informação em cada acumulação, do mesmo modo que o *Balance-of-Evidence* (Vickers & Packer, 1982).

Esta previsão tem como base o facto de que o enviesamento do critério de resposta se manifesta através de uma diminuição do limiar de resposta associado a essa resposta, o que limita a quantidade de informação que pode ser acumulada no sentido desta. Quando os limiares são semelhantes, as reversões podem ocorrer em ambos os sentidos, no entanto, quando os limiares não são semelhantes, a reversão de resposta só pode ocorrer num único sentido, no da resposta preterida pelo critério. Isto porque a quantidade de informação necessária para atingir a resposta preferida pelo critério é menor, existindo a possibilidade de haver mais informação codificada no sentido da resposta alternativa. Já no caso de uma resposta inicial que vai no sentido oposto ao do critério, esta não deve poder ocorrer dado que a informação necessária para atingir o limiar correspondente é superior. Outro aspecto a considerar é o *compromisso velocidade-acuidade*, o qual é representado pelo modelo através de um distanciamento dos limiares de resposta na condição de acuidade, pelo que a proporção de reversões de resposta deve diminuir, mantendo-se no entanto as previsões relativas ao enviesamento de resposta.

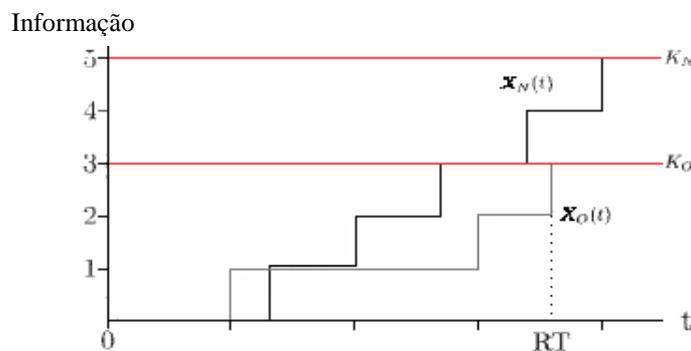


Figura 6. Representação gráfica do processo de acumulação de informação ao longo do tempo postulada pelo Race Model. K_n e K_o representam os limiares de decisão das respostas “novo” e “antigo”, respectivamente, enquanto $X_n(t)$ e $X_o(t)$ correspondem são funções que representam a quantidade de informação acumulada no sentido destas num determinado momento t . Adaptado de Van Zandt & Maldonado-Molina (2004)

Os resultados encontrados mostram no entanto que a ocorrência de reversões ocorre maioritariamente no sentido das respostas incorrectas para respostas correctas, e que a frequência das reversões de resposta não é afectada pela manipulação do critério de resposta, algo que vem comprometer os *modelos pós-decisórios* enquanto hipóteses explicativas. Quanto ao *compromisso velocidade-acuidade*, verificou-se que na condição de acuidade as reversões de resposta diminuem drasticamente, passando a ocorrer em cerca de 1% das decisões. Estes resultados vêm apoiar os modelos pós-decisórios como o *Race Model*, segundo os quais o afastamento dos limiares de resposta devem eliminar a ocorrência de reversões de resposta. Contudo, não foi encontrado em

nenhuma das condições de resposta qualquer efeito do enviesamento do critério de resposta nas reversões de resposta, o que contrasta com as previsões feitas pelo modelo. A ausência de um efeito do critério de resposta representa um problema para o modelo, que prevê reversões de resposta no sentido da resposta preterida pelo enviesamento da decisão inicial.

Podem ser apontadas diversas limitações a este estudo, tanto ao nível do material, como do procedimento. Quanto ao material, foram apenas utilizadas palavras não associadas, com alta frequência linguística, o que limita qualquer discussão relativamente ao tipo de evidências acumuladas no sentido da discriminação dos itens de teste; como se pode constatar nos estudos sobre falsas memórias, o tipo de evidências utilizadas é um aspecto fundamental para a compreensão do desempenho dos indivíduos (e.g., Johnson, Hashtroudi, & Lindsay, 1993; Mather, Henkel, & Johnson, 1997). No que diz respeito ao procedimento, os julgamentos de confiança eram feitos imediatamente a seguir à decisão inicial, o que restringe a reversão de resposta aos erros imediatamente detectados pelos sujeitos; dado que as reversões obtidas não foram afectadas pelo enviesamento de decisão, podemos afirmar com alguma confiança que estes erros corrigidos não se trataram de erros causados por antecipação ou por alguma falha na implementação da resposta, e sim reversões baseadas na avaliação das evidências acumuladas. Porém, estas reversões remetem para uma avaliação imediata das evidências, não contemplando aspectos importantes como o julgamentos que os indivíduos fazem do seu julgamento global, e no modo como este se pode reflectir nas reversões de resposta; novamente a questão do material pode ser colocada, dado que este julgamento pode levar uma avaliação com bases em tipos de evidências específicos, e que o acesso a diferentes tipos de evidências aparenta ser independente (Starns & Hicks, 2005). Outra crítica importante que pode ser feita em relação a este estudo é o facto de os autores não considerarem qualquer explicação que tenha como base uma diferenciação da informação ou dos processos envolvidos. Uma explicação plausível para os efeitos encontrados é que as respostas rápidas são baseadas na familiaridade dos estímulos, enquanto que as reversões de resposta são provocadas por um processo recollecção mais lento, que permite rejeitar ou aceitar itens anteriormente avaliados (Yonelinas, 1997; 2002); dado que nas respostas iniciais mais lentas é possível a presença de recollecção, as reversões de resposta descem drasticamente. Outras alternativas não-dualistas seriam igualmente possíveis, dada a capacidade explicativa dos modelos globalistas actuais (e.g., Wixted, 2007). Os resultados obtidos por Baranski e Petrusic (1998), bem como por Van Zandt e colaboradores (Van Zandt, 2000; Van Zandt & Maldonado-Molina, 2004) vêm mostrar que muitos dos pressupostos assumidos pelos modelos são muitas vezes simplificações incorrectas de processos bastante complexos. Contudo, muitas das abordagens adoptadas por estes estudos não são as mais adequadas, e as explicações oferecidas são bastante pobres tendo em conta todo o conhecimento acumulado em outras linhas de investigação, que

apesar de terem objectivos próprios, abordam o desempenho dos indivíduos exactamente no mesmo domínio – a memória de reconhecimento.

O problema da monitorização revela-se não apenas na literatura que aborda fenómenos mnésicos em que os mecanismos de monitorização ocupam um papel central, mas também nos estudos que abordam aspectos específicos dos modelos. Verifica-se que esta é uma questão que está presente, mesmo que de uma forma não manifesta, ao longo de toda a literatura sobre memória de reconhecimento, a qual surpreendentemente se encontra fragmentada e dispersa. Surge então a necessidade de integrar o conhecimento presente na literatura, e propor uma noção diferente de memória que proporcione formas alternativas de pensar o problema da monitorização, e novas medidas que permitam não apenas avaliar a capacidade desta, mas também a natureza dos componentes subjacentes.

O Problema e a Abordagem

Como se pôde verificar ao longo desta introdução, os modelos de memória em geral, apresentam limitações na descrição processual de efeitos recorrentes na literatura. Estas limitações derivam em parte da incapacidade que a TDS tem em distinguir conceptualmente e operacionalmente a força dos traços mnésicos da sua monitorização, bem como da inaptidão dos processos recolectivos propostos pelos modelos dualistas, em descrever o dinamismo e complexidade dos mecanismos de monitorização.

As críticas apresentadas, apesar de serem dirigidas tanto ao modelo globalista, como ao modelo dualista, centram-se fundamentalmente na componente de familiaridade, e como tal, na TDS. Contudo, é possível responder imediatamente às críticas utilizando o argumento de que a TDS assume uma dimensão de familiaridade indiferenciada porque a informação utilizada no processo de recordação é variável, não sendo por isso relevante a sua especificação. Tal como foi dito anteriormente, a TDS é fundamentalmente um modelo de decisão e não um modelo processual. Este argumento é falacioso; é verdade que a TDS não assume qualquer pressuposto relativamente à natureza da informação, à excepção da sua distribuição, e que este aspecto permitiu à TDS enquanto modelo genérico ser utilizada com grande sucesso em domínios tão distintos como o diagnóstico por Raio-X (Macmillan & Creelman, 2005). O problema não está na TDS em si, mas no facto reside no facto de que a TDS é directamente importada para a Psicologia Cognitiva como um modelo de memória viável, sem que hajam quaisquer considerações quanto à plausibilidade de *todos* os seus aspectos. Esta aceitação *a priori* da TDS como um bom modelo da memória humana pode

ser observada nos diversos modelos computacionais globalistas encontrados na literatura (e.g. MINERVA2, TODAM, CHARM, REM; para uma revisão destes modelos ver Raaijmakers & Shiffrin, 2002), os quais concretizam o modelo psicométrico globalista a um nível processual.

Consideremos então alguns pormenores de um dos modelos computacionais globalistas mais desenvolvido, o REM – *Retrieving Effectively from Memory* (Criss & Shiffrin, 2004; Shiffrin & Steyvers, 1997). Neste modelo, a informação é codificada num traço de memória, o qual é representado por um vector, sendo que cada atributo genérico codificado corresponde a um integral positivo numa entrada do vector, e cada atributo não codificado corresponde a um valor 0. O valor de cada integral positivo representa a qualidade de codificação do atributo, pelo que o valor será tanto maior quanto melhor tiver sido a codificação. Aquando da recuperação na fase de teste, um vector correspondente ao item de teste é emparelhado com cada um dos traços presentes em memória, sendo comparados os atributos. Para cada atributo que esteja presente com o mesmo valor em ambos os vectores, é quantificada a evidência no sentido do reconhecimento do item, sendo que esta evidência é calibrada segundo uma função geométrica, a qual define a probabilidade *a priori* de um determinado valor ser encontrado no mesmo atributo de ambos os vectores; quanto maior for o valor maior será a evidência no sentido do reconhecimento do item de teste. Esta função geométrica pode sofrer variações, consoante as expectativas de memorabilidade que o indivíduo tem em relação a um determinado tipo de itens. Após o vector do item de teste ter sido comparado com os traços disponíveis em memória, o somatório das evidências é avaliado segundo um decisor *bayesiano*, o qual define o reconhecimento ou não do item.

No modelo REM, os atributos são definidos de um modo genérico, não sendo feita qualquer distinção quanto à natureza da informação codificada. Até a probabilidade com que são calibrados os valores dos atributos é semelhante para todos os atributos, não sendo contemplada qualquer diagnosticidade diferencial, nem a sua avaliação individual. Os itens críticos do paradigma DRM são erroneamente reconhecidos dada a sua elevada familiaridade, a qual é constituída por informação de natureza associativa e semântica. Porém, os indivíduos são capazes de utilizar outros tipos de informação presentes em memória (e.g. atributos perceptivos) como base para a discriminação dos itens (Mather, Henkel & Johnson, 1997), ou então informação semântica para rejeitar itens perceptivamente semelhantes, porém semanticamente distintos (Gallo, Meadow, Johnson, & Foster, 2008). Esta avaliação dos itens de teste, tendo apenas como base uma parte dos atributos não é possível segundo o REM, apesar de ser facilmente encontrada na literatura de falsas memórias.

Estas limitações levam a que não sejam contemplados os mecanismos de monitorização e regulação presentes ao longo do desempenho dos indivíduos, e que se podem manifestar de maneiras distintas. A monitorização pode ocorrer ao nível da recuperação dos traços, verificando-se

uma recuperação estratégica de informação presente em memória em função dos itens a recordar e do tipo de distractores esperados (Jacoby Shimizu, Daniels, & Rhodes, 2005; Starns et al., 2007). Igualmente possível é a monitorização ocorrer após a recuperação dos traços mnésicos, sendo avaliada a distribuição da informação constituinte dos traços de forma a obterem uma melhor discriminação entre recordações verídicas e falsas (e.g., Norman & Schacter, 1997). Outro meio pelo qual a monitorização metacognitiva se podem manifestar é através de heurísticas, através das quais podem ser avaliadas de um modo rápido e frugal as características dos traços mnésicos. Esta avaliação pode ocorrer com base num critério de “*tudo-ou-nada*”, como se verifica na heurística da distintividade (Israel & Schacter, 1997; Schacter, Israel & Racine, 1999). Todos estes fenómenos apresentam características específicas ao nível do tipo de monitorização implementado, que vão desde processos de regulação e avaliação da informação recuperada, até julgamentos baseados em características dicotómicas. Deste modo, qualquer modelo (quer seja um modelo de decisão ou processual) que assuma como objectivo a descrição do desempenho dos indivíduos tem de encontrar formas adequadas para conseguir lidar com a complexidade presente nos mecanismos de monitorização metacognitiva.

O grande motivo para estas limitações é o facto de existirem linhas de investigação que, apesar de terem o mesmo objecto de estudo, raramente têm em conta os desenvolvimentos alcançados pelas outras. De facto, a literatura sobre falsas memórias raramente é citada na literatura que estuda memória de reconhecimento de palavras, e o mesmo ocorre com o estudo do reconhecimento de faces. Por outro lado, muitos dos modelos que procuram explicar o desempenho em tomada de decisão centram-se actualmente no estudo dos processos em tarefas perceptivas, apesar de terem como objectivo principal explicar o desempenho dos indivíduos em tarefas de memória⁹. A integração do conhecimento acumulado nestas diferentes linhas e a combinação de diferentes procedimentos utilizados é algo que terá necessariamente que ocorrer.

Metamemória: Um enquadramento para o problema

Tendo em consideração o estado actual da literatura, a procura por soluções surge como o próximo passo lógico; contudo, o cerne da questão permanece na forma como o problema deverá ser estruturado e abordado. A perspectiva aqui proposta é de que o primeiro passo a ser tomado não deve ser ao nível da modelação dos processos em si, mas sim numa compreensão exaustiva do

⁹ Um exemplo clássico é o *Diffusion Model* (Ratcliff, 1978; Ratcliff & Rouder, 1998; Ratcliff, Van Zandt, & McKoon, 1999), que originalmente procurava explicar o desempenho dos sujeitos em memória de reconhecimento, e actualmente é maioritariamente aplicado a tarefas de discriminação de letras, discriminação de luminosidade, ou decisão lexical.

alcance que a regulação metacognitiva tem, em que medida esta ocorre de forma consciente, e quais são os limites ao nível da regulação e controlo da informação processada que os indivíduos manifestam.

O enquadramento teórico que se apresenta como mais adequado ao problema da monitorização é a teoria hierárquica de Nelson e Narens (1990), desenvolvido na literatura de metamemória, com o objectivo de compreender a relação entre o desempenho percebido e o desempenho objectivo dos indivíduos. Segundo esta teoria, existem dois níveis de processamento cognitivo, um *nível básico* e um *meta-nível*, existindo entre estes níveis dois fluxos de informação distintos, um fluxo de informação ascendente, designado por *monitorização*, e um fluxo descendente designado por *controlo* (ver Figura 7).

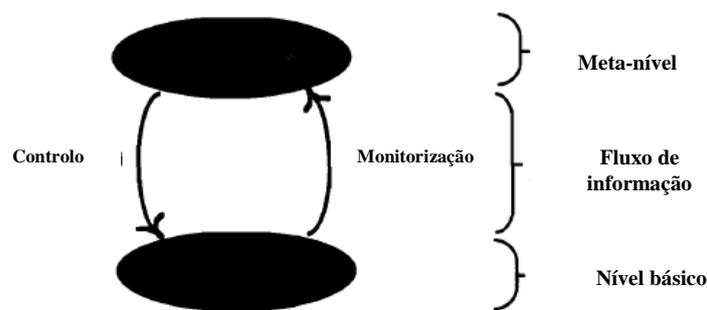


Figura 7. Ilustração da teoria hierárquica. Adaptado de Nelson & Narens (1990).

O *nível básico* refere-se ao processamento cognitivo que ocorre na execução de uma determinada tarefa, tanto ao nível dos recursos utilizados, bem como da quantidade de informação disponível. Esta tarefa tanto pode ser um julgamento de memória, como um julgamento de reconhecimento, ou então uma tarefa de codificação, como a aprendizagem de listas de estímulos. O *meta-nível* corresponde a uma representação do nível básico, a qual integra não apenas o processamento cognitivo subjacente ao desempenho do nível básico, mas também uma avaliação dos recursos disponíveis, e as expectativas em relação à tarefa em si. O *meta-nível* permite então uma avaliação do desempenho do indivíduo, tendo em conta os recursos e informação utilizados numa determinada tarefa, e os recursos e informação ainda disponíveis.

Os dois fluxos de informação definem a relação existente entre os dois níveis de processamento; o fluxo de *controlo* permite que o *meta-nível* altere o funcionamento do *nível básico*, tanto em termos de recursos cognitivos aplicados, como da utilização da informação disponível. Quanto à *monitorização*, esta consiste num processo de *feedback* por parte do *nível básico*, o que leva a uma alteração do *meta-nível*, e da representação que este faz do *nível básico*. Esta alteração leva igualmente a alterações no fluxo de *controlo*, levando a modificação do *nível*

básico. A teoria hierárquica define um fluxo circular, em que a percepção do indivíduo em relação ao seu desempenho influencia os processos subjacentes a este, o que por sua vez leva a uma alteração da percepção do indivíduo. Esta teoria permite enquadrar muitos dos efeitos encontrados na literatura, como o efeito de *feedback* no paradigma DRM (Jou & Foreman, 2007), no qual se verifica que o *feedback* após cada resposta dos participantes levava a um aumento do *d'* ao longo de toda a fase de teste, em contraste com os participantes a que não era dado *feedback*. O *feedback* deve levar a uma actualização do *meta-nível*, o que por sua vez tem consequências no funcionamento do *nível básico*, por via do fluxo de *controlo*. Este e outros efeitos, como o efeito de *warning* (Starns et al., 2007), heurística da distintividade (Schacter, Israel & Racine, 1999), a gestão de dimensões (Chalfonte & Johnson, 1996), ou a utilização estratégica da recollecção parcial (Dodson, Holland, & Shimamura, 1998) podem ser interpretados como manifestações da relação entre os dois níveis da teoria hierárquica.

Apesar de contrastar com os princípios postulados pelos modelos de memória, especialmente com a TDS, a noção da teoria hierárquica de que a informação utilizada na recordação pode ser regulada não é inédita nos modelos de memória, podendo ser encontrada no trabalho de Atkinson e Shiffrin (1971): “memory is best described in terms of the flow of information into and out of short-term storage and the subject’s control of that flow” (p. 83).

Apesar das vantagens que a teoria hierárquica traz para a interpretação dos efeitos encontrados na literatura, mantém-se o problema da medição da monitorização, tanto ao nível da sua capacidade, como da sua relação com o desempenho dos indivíduos. Para que uma medida possa ser desenvolvida, é necessário que se conceba uma tarefa que à partida permita avaliar a monitorização de forma objectiva, de preferência uma medida com as mesmas qualidades métricas encontradas na TDS.

TDS tipo-2: Sensibilidade e critério de decisão na monitorização

Um meio para se observar as particularidades e limites da monitorização consiste na observação de como esta possibilita aos indivíduos uma avaliação do seu desempenho, quando estes estão perante uma situação experimental em que diferentes fontes de informação estão disponíveis simultaneamente, fontes cuja diagnosticidade é questionável. Esta avaliação do desempenho pode ser feita através de um *juízo de correcção*, sendo pedido aos participantes, após terem completado a tarefa de reconhecimento, que indiquem quais das suas respostas estão correctas e quais estão incorrectas. Esta abordagem permite não apenas enquadrar os diversos aspectos da regulação metacognitiva numa tarefa de classificação simples, mas também encarar a

questão da monitorização como uma extensão do *problema fundamental da detecção* (Swets, Tanner, & Birdsall, 1961), tornando deste modo possível o seu estudo à luz da TDS, mais especificamente da TDS tipo-2.

A Teoria de Detecção de Sinal tipo-2 (TDS tipo-2; Clarke, Birdsall, & Tanner, 1959; Pollack, 1959) utiliza medidas equivalentes às da TDS tradicional (também designada por TDS tipo-1), permitindo do mesmo modo a distinção entre sensibilidade discriminativa (d') e critério de resposta (C_2). A diferença entre os dois tipos de tarefas reside nos significados atribuídos às medidas. Enquanto que nas decisões tipo-1 os sujeitos determinam se um determinado estímulo ocorreu previamente ou não, ou seja discriminar entre “sinal” e “ruído”, aquilo que está em causa na decisão tipo-2 é a capacidade que os sujeitos têm em discriminar que respostas dadas estão correctas ou não (Galvin et al., 2003). O aspecto central desta distinção é que na tarefa tipo-1 (e.g. tarefa de reconhecimento), aquilo que está em causa é a discriminação entre estímulos, estímulos esses cuja natureza (sinal ou ruído) é definida independentemente do observador. Pelo contrário, na tarefa tipo-2 a discriminação consiste numa avaliação das evidências presentes nos processos que levaram a uma determinada decisão (decisão correcta vs. decisão incorrecta). Enquanto que na tarefa tipo-1 a avaliação ocorre apenas ao nível da discriminação entre sinal e ruído, na tarefa tipo-2, a avaliação feita no contexto da resposta previamente dada pelo próprio sujeito. Assim, na tarefa tipo-1 as probabilidades de acerto (H) e de falso alarme (FA) são definidas com base na natureza do estímulo em si:

$$H = P(\text{“sim”} | \text{sinal})$$

$$FA = P(\text{“sim”} | \text{ruído})$$

Pelo contrário, na tarefa tipo-2 as probabilidades de acerto e falso e de falso alarme são baseadas na escolha prévia do indivíduo:

$$H = P(\text{“sim”} | \text{correcto})$$

$$FA = P(\text{“sim”} | \text{incorrecto})$$

A TDS tipo-2 não tem sido utilizada no estudo da memória de reconhecimento, tendo sido aplicada de uma forma mais sistemática no estudo da *percepção inconsciente*. Kunimoto, Miller, & Pashler (2001) testaram o desempenho de um grupo de participantes na identificação de estímulos apresentados subliminarmente, sendo manipulada a sua velocidade de apresentação. Era pedido a estes participantes que respondem se determinado estímulo foi apresentado ou não, sendo também questionados acerca da correcção das suas respostas. Aquilo que se verificou foi que

existiam dois limiares distintos, um relativo ao desempenho na tarefa tipo-1, e outro referente aos julgamentos de correcção. O primeiro limiar é mais baixo do que o segundo, existindo assim sujeitos que apesar de terem um desempenho acima do nível do acaso, demonstraram ser incapazes de distinguir entre as suas respostas correctas e incorrectas. Esta situação corresponde a um caso de percepção inconsciente, que pode ser identificado e medido de uma forma objectiva através da TDS tipo-2.

Apesar das possibilidades que a TDS tipo-2 traz para o estudo dos mecanismos de monitorização das memórias, esta não é considerada pela grande maioria das investigações que abordam mecanismos metacognitivos, sendo considerada uma mera curiosidade por muitos investigadores (Macmillan & Creelman, 2005; no entanto ver Goldsmith & Koriat, 2008). Uma das poucas utilizações da TDS tipo-2 no domínio da memória ocorre no estudo da regulação metacognitiva na tarefa de recordação livre (Higham, 2007); neste caso, através de uma conjugação das palavras reproduzidas numa tarefa de recordação livre, e numa tarefa de recordação forçada, é possível obter as proporções necessárias ao cálculo de $d'2$ e $C2$. Estas medidas permitem uma avaliação da regulação estratégica da acuidade na recordação de palavras. Possibilitando uma avaliação independente da monitorização e da acuidade. No entanto, é necessário salientar que o facto de a TDS tipo-2 proporcionar medidas análogas às da TDS tradicional, não implica que ambas as medidas possam ser directamente comparadas, os integrais das funções ROC tipo-2 são sempre inferiores aos das ROC (Galvin et al., 2003). O modo mais adequado de avaliar ambas as medidas é através de um modelo de regressão.

Uma das razões para a pouca utilização da TDS tipo-2 é a confusão que normalmente é feita entre a tarefa tipo-2 e as escalas de confiança normalmente incluídas nas tarefas tipo-1. De facto, ambas não poderiam ser mais distintas. A TDS Tipo-2 avalia a capacidade dos indivíduos em discriminarem as suas respostas correctas e incorrectas, por isso se um indivíduo considerar todas as suas respostas anteriores como correctas, manifestará uma sensibilidade discriminativa nula, ou seja, este indivíduo obterá um $d'2 = 0$. Este exemplo vem mostrar um aspecto fundamental da TDS enquanto modelo de memória, em condições de teste em que não haja qualquer interferência externa no processo de decisão (e.g. atenção dividida, respostas rápidas), os indivíduos devem obter um desempenho, em relação ao qual são incapazes de identificar respostas correctas e incorrectas. Dado que a TDS assume um julgamento unidimensional baseado na familiaridade, esta deve conter toda a informação disponível ao indivíduo, representada de uma forma indistinta. Se os indivíduos não possuírem mais informação disponível, devem obter resultados nulos na tarefa Tipo-2. De facto, a unidimensionalidade da TDS produz uma situação algo paradoxal; um indivíduo cujo desempenho é comprometido por uma fonte externa, como por exemplo ter de executar duas tarefas simultaneamente, deve obter resultados superiores na TDS Tipo-2 do que um indivíduo a

quem são dadas condições que promovem a acuidade, como tempo para processar o estímulo e para responder. Para compreender melhor esta relação entre a TDS e a TDS tipo-2, basta considerarmos o exemplo da altura das mulheres dado anteriormente; se fosse pedido a indivíduos que identificassem as suas respostas correctas e incorrectas, tendo apenas como base a altura das mulheres, a única coisa que estes poderiam fazer seria alterar o seu critério de resposta, o qual é independente da sensibilidade discriminativa e apenas levaria a uma diminuição da proporção de falsos alarmes, acompanhada por uma alteração da proporção de omissões no sentido inverso. Apenas com uma alteração ao nível da informação disponível, e subsequente aumento da sensibilidade discriminativa, é possível obter um desempenho acima do acaso na tarefa tipo-2.

A grande maioria da literatura apresentada ao longo desta introdução tem demonstrado que os indivíduos conseguem de facto regular a informação utilizada nos julgamentos, pelo que estas previsões extraídas dos pressupostos da TDS permitem fazer um teste directo à qualidade desta enquanto modelo viável da memória de reconhecimento.

A inclusão da TDS Tipo-2 apresenta então duas grandes vantagens; em primeiro lugar permite um teste directo aos pressupostos da TDS, dado a obtenção de resultados positivos na tarefa Tipo-2 em relação a um desempenho anterior, no qual a acuidade não foi comprometida, não podem em teoria ser obtidos. Em segundo lugar, caso os pressupostos da TDS sejam violados, a TDS Tipo-2 permite avaliar de forma objectiva a capacidade de monitorização dos indivíduos, e a natureza deste desempenho. Para além das medidas proporcionadas pela TDS tipo-2, a própria tarefa tipo-2 permite a obtenção de curvas ROC, o que possibilita a avaliação dos processos subjacentes, nomeadamente se estamos perante um processo baseado apenas na recollecção, ou se o desempenho dos indivíduos implica um processo de familiaridade, o que implicaria uma regulação da informação utilizada na familiaridade.

Mecanismos de monitorização e modelos de memória: Um estudo empírico

Como foi referido ao longo da introdução, os mecanismos de monitorização manifestam-se de formas que não são contempladas pelos modelos de memória, quer ao nível das propostas globalistas, quer dualistas. Um dos problemas fundamentais é o facto de a acuidade poder resultar tanto da força dos traços mnésicos, como da actuação dos mecanismos de monitorização sobre estes (Goldsmith & Koriat, 2008), o que não permite avaliar se a sensibilidade discriminativa medida representa uma capacidade baseada na “totalidade” de informação disponível, estando fora do controlo dos indivíduos (ao contrário do critério de resposta), ou se representa apenas uma

configuração possível da informação disponível, que pode ser alterada consoante as circunstâncias. Um modo objectivo de testar esta capacidade de monitorização é a utilização da TDS tipo-2, a qual permite representar a capacidade de monitorização como uma extensão do *problema fundamental da detecção* (Swets, Tanner, & Birdsall, 1961).

Dado o problema colocado pela monitorização, e as oportunidades proporcionadas pela TDS tipo-2, são então propostas duas experiências que procuram avaliar a capacidade dos mecanismos de monitorização em identificarem os erros cometidos, e a natureza desta identificação. Em ambas as experiências existem duas fases de teste; o teste de reconhecimento, em que são apresentadas palavras anteriormente apresentadas, e palavras novas. No segundo teste, a tarefa tipo-2, é novamente apresentada aos participantes a lista de teste, porém é incluído junto a cada item de teste a resposta anteriormente dada, sendo perguntado ao participante se considera a sua resposta anterior correcta ou incorrecta.

A primeira experiência tem como objectivo avaliar a capacidade dos mecanismos de monitorização, nomeadamente se os participantes conseguem discriminar as suas respostas correctas e incorrectas, ocorridas no contexto do paradigma DRM. Segundo os modelos de memória esta capacidade não deve existir, dado que segundo a TDS toda a informação é representada de forma indiscriminada ao longo de uma dimensão de familiaridade. A única excepção é apresentada pelo modelo dualista, segundo o qual a discriminação das respostas correctas e incorrectas poderá ocorrer devido à recolção de informação anteriormente não recuperada. A hipótese estabelecida pela TDS é então de que os $d'2$ dos participantes devem ser iguais a 0, a qual pode ser testada pela medição directa dos valores desta medida. A hipótese que decorre do modelo dualista, de que a capacidade de discriminação na tarefa tipo-2 deriva da componente de recolção, pode ser testada através da forma da função zROC, a qual deve ser curvilínea quando o desempenho do indivíduo apenas tem por base a recolção. A função zROC permite igualmente identificar se o desempenho dos indivíduos tem subjacente pelo menos a componente de familiaridade, dado que a presença desta levaria a funções zROC lineares. A obtenção de funções lineares pode ser associada a uma *hipótese da monitorização*, segundo a qual os indivíduos conseguem regular de forma estratégica a informação utilizada na componente de familiaridade. Esta experiência permite então testar a adequação dos modelos de memória em relação aos mecanismos de monitorização.

Quanto às funções zROC obtidas nas duas tarefas, estas não vão ser obtidas através da utilização de níveis de confiança, sendo utilizada uma manipulação do critério de resposta através de *payoffs*, a qual os níveis de confiança se propõem a simular (Van Zandt, 2000). Existem duas razões principais para esta opção; em primeiro lugar existe o risco de os indivíduos confundirem os níveis de confiança com o tipo de julgamento pedido nos julgamentos tipo-2 (Galvin et al., 2003);

em segundo lugar, os níveis de confiança são vulneráveis a dependências sequenciais, ou seja, a atribuição de um determinado nível de confiança num dado ensaio não é independente do nível de confiança atribuído no ensaio anterior (Muller & Weidemann, no prelo).

Relativamente à tarefa tipo-2, é manipulada a informação que é dada ao indivíduo nesta tarefa, podendo ser fornecida informação relativamente ao seu desempenho anterior (% respostas correctas), ou então não ser fornecida qualquer informação. Esta manipulação é especialmente importante na medida em que existem estudos que demonstram uma alteração do d' aquando da presença de informação adicional, quer seja através de avisos prévios (Starns et al., 2007), quer seja através de *feedback* imediato (Jou & Foreman, 2007). A inclusão de uma condição de *feedback* assume duas funções; em primeiro lugar permite avaliar qual o impacto deste tipo de informação adicional na tarefa tipo-2, tanto ao nível dos $d'2$ como da forma das funções zROC. Em segundo lugar, permite despistar uma hipótese alternativa, caso sejam encontrados uma sensibilidade discriminativa nula na tarefa tipo-2, ou seja $d'2 = 0$. Segundo esta hipótese, é possível que os mecanismos de monitorização não permitam a obtenção de um desempenho positivo na tarefa porque os indivíduos não têm uma percepção objectiva do seu desempenho. Esta possibilidade é reforçada pelo facto de que os erros de memória ocorridos no paradigma DRM têm normalmente associados elevados níveis de confiança (Read, 1996; Roediger & McDermott, 1995).

A segunda experiência procura avaliar uma versão “relaxada” dos modelos de memória; uma crítica que pode ser apontada à primeira experiência é o facto de não considerar a acumulação de evidências que ocorre num processo de tomada de decisão, a qual é contemplada em modelos globalistas, como o Diffusion Model (Ratcliff, 1978), que tem em conta os tempos de resposta dos participantes. Assim, é possível que um desempenho positivo na tarefa tipo-2 não tenha como base a regulação da informação disponível, mas sim o facto das decisões dos indivíduo poderem ser tomadas mesmo quando apenas parte da informação foi acumulada (Ratcliff, 1988; Usher & McClelland, 2001). De facto, mesmo modelos mecanicistas como o REM (Shiffrin & Steyvers, 1997), que excluem os TR enquanto variável de interesse, consideram que a recuperação da informação em memória é um processo que se estende ao longo do tempo, e que por isso seria errado considerar que todas as evidências estão simultaneamente disponíveis.

De forma a testar este “relaxamento” da TDS, esta experiência incorpora na tarefa de reconhecimento o *paradigma sinal-resposta* (Doshier, 1981, 1984; McElree, Dolan, & Jacoby, 1999; Ratcliff, 2006), no qual a latência de resposta dos indivíduos é controlada experimentalmente, através de um sinal que indica ao participante o momento em que deve dar a sua resposta. O momento do sinal varia entre cada ensaio de teste, podendo surgir quase imediatamente à apresentação do item (e.g, 300ms), algum tempo depois (e.g., 1800ms), ou em tempos intermédios. Este paradigma tem sido utilizado como utilizado como forma de medir três

aspectos do processamento cognitivo; o momento temporal em que a informação acumulada permite uma acuidade acima do acaso; a taxa de crescimento da acuidade em direcção a uma assíntota; e o nível de acuidade assintótico. Esta avaliação não pode ser feita com os condicionamentos de TR normalmente utilizados, como a imposição de limites de tempos de resposta fixos, dado que estes apenas permitem uma estimativa do tempo necessário para tomar uma decisão, porém não permitindo estudar os momentos temporais em que a informação se torna disponível, nem a acuidade assintótica. Outra vantagem do paradigma sinal-resposta é o facto de o momento da apresentação do sinal variar aleatoriamente entre cada ensaio, não permitindo aos indivíduos estabelecer uma estratégia de resposta, o que pode acontecer quando são utilizados limites de resposta fixos (Ratcliff, 2006).

Este paradigma permite então testar a versão “relaxada” dos modelos, dado que segundo estes deve existir uma relação negativa entre o crescimento dos d' ao longo da latência do sinal de resposta, e os valores de d'^2 correspondentes. Os estímulos que foram processados mais tempo, devem demonstrar valores de d'^2 próximos de zero, enquanto os estímulos cujo processamento foi curto devem ser mais facilmente discriminados na tarefa tipo-2, levando a valores de d'^2 positivos. Esta previsão deve-se verificar no ajustamento dos valores de d' e d'^2 a um modelo de regressão, verificando-se uma relação linear negativa. A hipótese da monitorização, apesar de não fazer previsões específicas (entre uma relação linear positiva ou uma não relação), não prevê a existência desta relação negativa.

Estas duas experiências permitem uma avaliação preliminar da adequação dos modelos de memória actualmente propostos, e a validação da tarefa tipo-2 como um modo de medir a capacidade dos mecanismos de monitorização na regulação da informação disponível nos traços mnésicos. Estas experiências procuram ser um ponto de partida para a inclusão futura dos mecanismos de monitorização, tanto nos modelos psicométricos, como nos modelos mecanicistas de memória.

Experiência 1

A primeira experiência procura testar a existência de mecanismos de monitorização que permitam os indivíduos identificar na tarefa tipo-2 as suas respostas correctas e incorrectas. De acordo com o modelo globalista, os indivíduos não devem ser capazes de fazer este tipo de discriminação na medida em que o mesmo tipo de evidências indistintas (familiaridade) é o único tipo de informação disponível em ambas as tarefas. O modelo dualista prevê a possibilidade de existir um desempenho acima do acaso na tarefa tipo-2, contudo este desempenho terá como base a recollecção de informação não recordada na tarefa de reconhecimento. Estas previsões podem ser testadas através de determinados parâmetros calculados a partir das respostas dos participantes, nomeadamente a intersecção e a linearidade das funções zROC. Outro aspecto a ser considerado é a relação entre o desempenho na tarefa de reconhecimento, o desempenho na tarefa tipo-2. Se assumirmos uma visão “relaxada” dos modelos, ambos prevêem a existência de uma relação linear negativa. Igualmente analisada será a natureza das reversões de resposta, e a existência de possíveis diferenças nas suas proporções consoante a resposta inicial.

Método

Participantes

Nesta experiência participarão voluntariamente 40 indivíduos de ambos os sexos, estudantes universitários da Universidade de Lisboa, com as idades compreendidas entre os 18 e os 30 anos.

Plano Experimental

O plano experimental a utilizar consiste num plano factorial misto, 2 condições de informação (condição de *feedback* vs. condição de omissão), manipulação inter-participantes x 5 blocos de recompensa/punição de resposta (muito conservador vs. conservador vs. neutro vs. liberal vs. muito liberal), manipulação intra-participantes. As variáveis dependentes são as proporções de acertos e falsos alarmes na tarefa de reconhecimento e na tarefa tipo-2, as intersecções das funções zROC estimadas, correspondentes a d' e d'^2 , e a proporção de reversões de resposta em cada um dos blocos, que podem ser reversões correctas (falso alarme → rejeição correcta; omissão → acerto) ou incorrectas (rejeição correcta → falso alarme; acerto → omissão).

Material

Serão criadas 20 listas DRM, cada uma constituída por 15 palavras associadas a uma *palavra-crítica*, obtidas através de um pré-teste. As palavras de cada lista DRM são ordenadas de forma crescente segundo a respectiva BAS. Para a construção da lista de estudo são seleccionadas as palavras ímpares das 20 listas DRM, o que perfaz um total de 160 palavras. Na lista de teste, para além das palavras da lista de estudo, são adicionadas as palavras pares das listas DRM, mais os respectivos itens críticos, o que perfaz um total de 320 palavras. Para um exemplo de uma lista DRM, ver Anexo2.

Enquanto que a utilização de apenas distractores associados na fase de teste não é inédita na literatura de falsas memórias (e.g., Gunter, Ivanko, & Bodner, 2005), a utilização de quase metade das palavras de cada lista DRM como distractores é. Esta composição da lista de teste tem como objectivo o controlo de possíveis enviesamentos “semânticos” provocados pelas listas de teste normalmente utilizadas na literatura (e.g., Thapar & McDermott, 2001). Como referido anteriormente, a explicação deste enviesamento encontra-se no Anexo 1.

Procedimento

A experiência será realizada em pequenos grupos de cerca de dez participantes. Os participantes lerão as instruções que lhes serão facultadas individualmente. Na fase de apresentação, é pedido aos participantes que observem com atenção e tentem memorizar as palavras que vão surgir no computador em seguida. As listas DRM (palavras ímpares) são apresentadas aleatoriamente, sendo as palavras de cada lista apresentadas por ordem decrescente segundo a sua BAS. Cada palavra é apresentada em letras maiúsculas durante 1500ms, sendo apresentado após a palavra, um ecrã branco durante 500ms.

Após fase de apresentação das palavras, os participantes executam uma tarefa distractora, em que lhes é pedido aos que contem a quantidade de algarismos pares presentes numa sequência, e que insiram a sua resposta no computador. Sempre que introduzirem uma resposta, ser-lhes-á dada uma outra sequência, até que a tarefa distractora perfaça um total de 3 minutos.

Em seguida, é pedido aos participantes que executem uma tarefa de reconhecimento; nesta tarefa são apresentadas palavras, sendo a tarefa do participante indicar se essa palavra apareceu na fase de estudo, ou se é uma palavra nova, respondendo através dos botões indicados numa caixa de resposta (SIM/NÃO). Não é imposto qualquer limite no tempo de

resposta. Entre a apresentação de cada palavra, após ter sido dada uma resposta, é dado um intervalo de 1s, durante o qual é apresentado um alvo (+), centrado no ecrã. A tarefa de reconhecimento é constituída por 5 blocos de punição/recompensa, cada um constituído por 64 palavras; as palavras de cada bloco são seleccionadas aleatoriamente, existindo no entanto constrangimentos, nomeadamente a proporção de palavras novas e antigas ser igual, e serem apresentadas 4 palavras-críticas por bloco. Em cada um dos blocos, é estabelecida uma pontuação para cada resposta dada, sendo as respostas correctas bonificadas (+X), enquanto que às respostas incorrectas é atribuída uma pontuação negativa (-Y). Antes de iniciarem a tarefa de reconhecimento, os participantes são informados de que é atribuída uma pontuação a cada resposta, sendo que a pontuação total no final de cada tarefa corresponde a um prémio monetário que pode ascender a €8 (3072 pontos), consoante o total de pontos acumulados. A pontuação atribuída varia consoante os blocos de forma que os participantes manipulem o seu critério de resposta (Macmillan & Creelman, 2005); os blocos e respectivas recompensas e punições em termos de pontos são as seguintes: *bloco muito conservador*: +3, -9; *bloco conservador*: +3, -6; *bloco neutro*: +3, -3; *bloco liberal*: +6, -3; *bloco muito liberal*: +9, -3. Estes blocos são introduzidos sequencialmente, sendo a ordem contrabalançada entre participantes. Quando finalizam a tarefa de reconhecimento, os participantes iniciam a tarefa Tipo-2.

Na tarefa tipo-2, na condição de *feedback*, os participantes são informados do seu desempenho anterior na tarefa de reconhecimento, através de uma percentagem de respostas correctas (*X% Respostas Correctas*). Esta informação aparece numa janela no ecrã do computador, surgindo no ecrã antes do início da tarefa tipo-2, ficando visível ao longo de toda a tarefa. Na condição de *omissão*, não é dada qualquer informação relativa ao desempenho anterior dos participantes. A tarefa Tipo-2 é exactamente igual à de reconhecimento, só que as palavras são acompanhadas da resposta dada anteriormente (SIM/NÃO) É então pedido aos participantes que indiquem se consideram correcta ou incorrecta a resposta dada anteriormente a cada palavra. Tal como a tarefa de reconhecimento, a tarefa tipo-2 é dividida em cinco blocos de punição/recompensa, cada um com 64 itens de teste; estes itens são seleccionados aleatoriamente, salvo os constrangimentos referidos anteriormente.

Após terminarem a tarefa tipo-2, os participantes são dispensados, sendo-lhes agradecida a sua participação.

Resultados e discussão:

Os resultados obtidos pelos participantes serão classificados como acertos e falsos alarmes tanto para a tarefa de reconhecimento, como para a tarefa Tipo-2. Na tarefa Tipo-2, serão ainda classificadas as reversões de resposta correctas (omissão → acerto; falso alarme → rejeição correcta), e incorrectas (acerto → omissão; rejeição correcta → falso alarme). A estimação dos d' e d'' através do desempenho nos diferentes blocos será feita através da utilização do algoritmo de estimação de máxima verosimilhança ROCKIT (Metz, 1998). A capacidade predictora dos $d'1$ em relação aos $d'2$ será avaliada através de um modelo de regressão, e as interacções com a condição de informação e o bloco de teste. As reversões de resposta serão ajustadas a um modelo de regressão condicional, no qual é avaliada a capacidade dos acertos, dos falsos alarmes e da interacção entre estes dois relativos à tarefa de reconhecimento enquanto variáveis predictoras das reversões correctas, incorrectas, e a sua interacção na tarefa Tipo-2.

Nesta experiência, espera-se que os valores estimados de $d'2$ sejam significativamente superiores a zero, o que indicaria a presença dos mecanismos de monitorização. Quanto às reversões de resposta, estas permitem compreender a direcionalidade da monitorização dos indivíduos, e observar diferenças na capacidade de corrigir, dado estas serem originalmente positivas (SIM) ou negativas (NÃO). Dado que não ter sido encontrada qualquer literatura que aponte para uma assimetria deste género, a análise das reversões de resposta assume-se como completamente exploratória.

Caso não seja encontrada uma diferença significativa, a hipótese que defende a presença de mecanismos de monitorização na regulação do desempenho dos indivíduos é refutada, em detrimento da hipótese que assume os pressupostos da TDS, nomeadamente a existência de uma dimensão de familiaridade indiferenciada que não permite a regulação do tipo de informação utilizada. Uma explicação possível para a obtenção de $d'2$ nulos é que os indivíduos não estão a regular a informação disponível porque consideram que o seu desempenho foi bastante bom. Esta hipótese não é de todo descabida se tivermos em conta que os indivíduos tendem a atribuir níveis de confiança bastante elevados em relação à recordação dos itens críticos (Read, 1996; Roediger & McDermott, 1995). Assim, o desempenho dos participantes não se deve a uma incapacidade, mas sim à sua não utilização desta capacidade. A distinção entre esta explicação e a explicação proposta pelos pressupostos da TDS pode ser feita através dos resultados obtidos na condição de *feedback* e na condição de omissão. Se os participantes forem de facto capazes de regular a informação utilizada nas suas decisões, então *feedback* deverá levar a que qualquer percepção distorcida do seu desempenho seja corrigida, e que haja então uma regulação da informação na tarefa tipo-2. Pelo contrário, se os indivíduos não conseguirem regular a informação utilizada, o *feedback* apenas

deverá levar a um efeito de *probability matching*, o que não alteraria em nada a acuidade dos participantes.

O segundo objectivo desta experiência, dependente da obtenção de d' significativamente superiores a zero, é explorar a natureza do desempenho dos indivíduos na tarefa tipo-2, através da relação deste com o desempenho na tarefa de reconhecimento, bem como da análise das funções zROC. Quanto à relação entre as duas tarefas, esta será analisada através do ajustamento dos dados individuais a um modelo de regressão; caso seja encontrada uma relação linear positiva entre as duas variáveis, isto significará que quanto maior for o desempenho dos indivíduos, maior é a sua capacidade em regular a informação utilizada. À partida esta relação seria interpretada através da noção de que quanto maiores e mais fortes forem as evidências disponíveis, maior é a capacidade dos indivíduos regular o tipo de informação utilizada; contudo, dado o vazio conceptual associado à medida d' (Goldsmith e Koriat, 2008), esta relação pode ser vista como resultado da monitorização dos traços mnésicos, a qual se reflecte nas duas tarefas. Se a regressão apontar para uma relação negativa, pode-se considerar que os pressupostos da TDS representam ainda uma hipótese viável, dado que esta hipótese considera que quanto maior for a informação for incluída na decisão inicial, menor será a sensibilidade discriminativa dos indivíduos.

Quanto às funções zROC, estas são informativas dos processos envolvidos no reconhecimento dos itens de teste, nomeadamente se estamos perante um desempenho baseado num processo de alto-limiar, como é o caso do processo de recolção, ou se é necessário considerar uma variável contínua, como acontece com as distribuições de familiaridade propostas pela TDS (Macmillan & Creelman, 2005). Outro aspecto que esta experiência permite avaliar, é a relação entre o desempenho dos indivíduos na tarefa de reconhecimento, representado pela função zROC, e o desempenho dos indivíduos na tarefa tipo-2, no mesmo tipo de função. A avaliação desta relação é obviamente condicional na linearidade das funções obtidas nas duas tarefas.

Caso seja obtida uma função zROC curvilínea na tarefa tipo-2, pode-se considerar que é um processo de recolção que está por detrás do desempenho dos indivíduos. Este resultado, apesar de positivo para o estudo da regulação metacognitiva, seria algo problemático para a hipótese estabelecida, dado que seria uma evidência no sentido que os indivíduos não conseguem regular a informação utilizada ao nível do eixo de familiaridade, ao contrário do que é sugerido por alguns estudos (e.g., Starns et al., 2007; Jou & Foreman, 2007). O facto de que a regulação metacognitiva tem por base um processo de alto-limiar é bastante interessante para o estudo dos processos de recolção, na medida em que implica que os indivíduos não procuram utilizar um processo de recolção em todos os contextos; isto poderá significar que as diferenças nas assimetrias das curvas ROC podem ser atribuídas não apenas às características de estudo e de teste, mas também às expectativas de acuidade dos indivíduos. A obtenção deste resultado poderia levar

a um conjunto de investigações que procurassem avaliar a relação entre a presença de recollecção e a percepção que os indivíduos têm do seu desempenho na tarefa, do mesmo modo que alguns estudos em metamemória procuram avaliar a relação entre o desempenho percebido e desempenho real (e.g., Benjamin, 2003). Este resultado também poderia rever a interpretação dada pelos modelos dualistas à recollecção parcial (Yonelinas & Jacoby, 1996), a qual é normalmente relegada para um papel completamente inconsequente, apesar de alguns estudos em monitorização da fonte sugerirem que este tipo de recollecção pode ser utilizada para determinados tipos de discriminações (Dodson, Holland, & Shimamura, 1998).

Quanto à hipótese defendida, de que há uma regulação ao nível da familiaridade, esta seria apoiada pela obtenção de funções zROC lineares. Contudo, como já foi referido anteriormente, a linearidade das zROC não permite a distinção directa entre um modelo apenas baseado na familiaridade, ou então de um modelo dualista, em que existe uma componente de familiaridade, e outra de recollecção. A única excepção seria se fossem obtidos declives iguais a um, o que denotaria a igualdade de variâncias das distribuições de sinal e ruído, ou a ausência de recollecção.

Apesar desta limitação, os parâmetros das funções zROC individuais, nomeadamente o declive e a intersecção, permitem testar através do ajustamento a um modelo de regressão linear, a predominância dos dois componentes dos modelos dualistas no desempenho na tarefa tipo-2. Caso estes parâmetros assumam uma relação linear com declive negativo, isto será indicativo de que os aumentos da acuidade na tarefa tipo-2 estão associados a uma diminuição do declive, o qual pode ser interpretado como um aumento da predominância da recollecção. Por outro lado, a não existência de uma associação significativa entre as duas variáveis significaria que o aumento da acuidade não implica qualquer alteração sistemática do declive, o que sugere que a componente de recollecção não assume um papel central na acuidade na tarefa tipo-2.

Experiência 2

A segunda experiência procura testar a versão “relaxada” dos modelos, segundo a qual é possível a existência de desempenhos positivos na tarefa tipo-2, porque quando os indivíduos tomam as suas decisões na tarefa de reconhecimento, estes não processaram necessariamente toda a informação disponível (Ratcliff, 1988; Usher & McClelland, 2001). De forma a testar esta hipótese, é utilizado um paradigma de sinal-resposta, o qual é tradicionalmente utilizado na literatura de memória de reconhecimento como forma de estudar a dinâmica dos processos recuperação de informação mnésica (e.g., McElree, Dolan, & Jacoby, 1999). De acordo com a versão “relaxada” dos modelos, são os processos de recuperação inacabados que estão na base do desempenho na tarefa tipo-2, pelo que o desempenho na tarefa tipo-2 deve ser tanto pior quanto maior for o tempo de latência do sinal de resposta na tarefa de reconhecimento. Esta experiência propõe então o teste desta previsão dos modelos “relaxados”.

Método

Participantes

Nesta experiência participaram voluntariamente 40 indivíduos de ambos os sexos, estudantes universitários da Universidade de Lisboa, com as idades compreendidas entre os 18 e os 30 anos.

Plano Experimental

O plano experimental a utilizar consiste num plano factorial misto, 2 condições de informação (condição de *feedback* vs condição de omissão), manipulação inter-participantes x 5 sinais de resposta (400ms vs. 800ms vs. 1200ms vs. 2400ms vs. 4000ms), manipulação intra-participantes. As variáveis dependentes serão as proporções de acertos e de falsos alarmes obtidos na tarefa de reconhecimento e na tarefa tipo-2, bem como as diferentes reversões de resposta, correctas e incorrectas.

Material

Será utilizado o mesmo material da experiência 1.

Procedimento

Esta experiência será realizada em sessões individuais. O participante lerá as instruções que lhes serão facultadas individualmente. Numa fase inicial, o participante será informado de que irá participar num estudo sobre memória, no qual a velocidade de resposta será um factor a ter em conta, nomeadamente a sua capacidade em fornecer uma resposta logo após receber um sinal sonoro de aviso através de auscultadores colocados nos ouvidos.

De forma a se familiarizar com o paradigma sinal-resposta, o participante executa uma fase de treino, na qual este paradigma é implementado numa tarefa perceptiva de detecção de sinal (para mais detalhes ver Ratcliff, 2006); em cada ensaio desta tarefa é apresentado um conjunto de pontos distribuídos aleatoriamente num quadrado no centro do ecrã (10x10cm), sendo que o número de pontos varia entre os 13 e 87, inclusive. É pedido ao participante que responda numa caixa de resposta se o conjunto de pontos corresponde à categoria “grande” ou “pequena”, cada uma destas respostas correspondendo a uma tecla de uma caixa de resposta. O participante é informado de que o número de pontos produzidos pelas duas categorias varia, podendo estas produzir valores semelhantes; esta sobreposição das categorias faz com que determinados números de pontos (e.g., 50) sejam bastante difíceis de categorizar, o que leva necessariamente à ocorrência de respostas erradas. O participante deve então tentar responder o mais correctamente possível, tendo em conta o sinal de resposta, dado que a sua resposta deve ser dada imediatamente após o sinal, não devendo ultrapassar os 300ms. O participante é igualmente informado que o sinal de resposta varia entre cada ensaio, podendo ocorrer logo ao início do ensaio, ou então passado algum tempo. Após cada resposta do participante, é-lhe dado *feedback* quando a sua resposta é demasiado lenta e/ou incorrecta. Para a resposta “grande”, a probabilidade do participante receber o *feedback* de resposta incorrecta é a seguinte: 13–22 pontos: .05; 23–27 pontos: .1; 28–32 pontos: .15; 33–37 pontos: .225; 38–42 pontos: .3; 43–47 pontos: .4; 48–52 pontos: .5; 53–57 pontos: .6; 58–62 pontos: .7; 63–67 pontos: .775; 68–72 pontos: .85; 73–77 pontos: .9; 78–87 pontos: .95. As probabilidades assumem a ordem inversa para a resposta “pequena”. Quanto ao *feedback* sobre o tempo de resposta, este é dado quando as respostas são dadas quando a resposta é dada antes do sinal (*Resposta muito rápida*), ou então quando o tempo de latência entre o sinal e a resposta ultrapassa os 300ms (*Resposta muito lenta*). O participante executa a tarefa de detecção de sinal em dois blocos de 100 ensaios cada, podendo descansar 5 minutos entre cada bloco. Entre cada ensaio, há um intervalo de 1s, durante o qual apenas aparece no centro do ecrã um alvo (+).

Após terminar a fase de treino o participante descansa durante 10 minutos, sendo então iniciada a fase de estudo, exactamente igual à da experiência anterior, e a subsequente tarefa distractora. O teste de reconhecimento é igualmente constituído por 5 blocos, mas estes têm como único objectivo permitir ao participante descansar durante 1m entre cada bloco. Na tarefa tipo-2, as palavras são novamente apresentadas, porém num único bloco de teste, não existindo qualquer limitação em termos de tempos de resposta. A manipulação da apresentação do feedback na tarefa tipo-2 ocorre do mesmo modo que na experiência 1. Após terminarem a tarefa tipo-2, os participantes são dispensados, sendo-lhes agradecida a sua participação.

Resultados e Discussão

Os resultados obtidos em termos de acertos e falsos alarmes na tarefa de reconhecimento, bem como na tarefa tipo-2 são utilizados para o cálculo pontual das medidas d' e d'^2 , sendo estas medidas calculadas para cada um dos cinco sinais de resposta. Em seguida estes resultados são ajustados num modelo de regressão, de forma a avaliar se a relação linear negativa prevista pela TDS se verifica ou não.

Caso esta relação se verifique, os modelos de memória globalistas e dualistas saem então reforçados, sendo capazes de explicar os dados obtidos na tarefa tipo-2. Qualquer outro resultado, como a obtenção de uma relação linear positiva, ou a ausência de qualquer relação significativa, pode ser interpretado como uma evidência no sentido da hipótese da monitorização, no sentido em que se verifica um desempenho acima do acaso, o qual não é previsto por nenhum dos modelos.

Contudo, é possível considerar que a componente de recolção do modelo dualista, dado ser controlada de forma consciente e estratégica, pode não se manifestar mesmo quando a latência do sinal de resposta é elevada. Assim, o desempenho acima do acaso na tarefa tipo-2 seria explicado pela recolção de informação anteriormente não recordada. Esta hipótese poderia ser testada através de uma experiência em que apenas fossem considerados sinais de resposta com latências elevadas, em diferentes condições de *payoff*, o que permitiria a obtenção de funções zROC. Caso esta hipótese seja verdadeira, as funções zROC devem ser lineares. O teste desta hipótese não foi considerado na presente experiência dada a quantidade elevada de observações que são necessárias para a obtenção de medidas fiáveis. Macmillan, Rotello, e Miller (2004) recomendam para a obtenção de d' fiáveis a utilização de 50 respostas, e para o estudo do declive e linearidade das zROC 200 respostas distribuídas de forma equilibrada pelos diferentes critérios de resposta (ou níveis de confiança). Se esta hipótese fosse testada no presente estudo, a lista de teste

teria que ter no mínimo 1000 itens para que fossem obtidas funções zROC para cada um das latências de sinal, o que seria uma tarefa extremamente difícil e cansativa para os participantes.

Discussão Geral

As duas experiências aqui sugeridas são estudos preliminares e como tal, representam apenas uma pequena parte do potencial das medidas Tipo-2 para o estudo dos mecanismos de monitorização a operar ao nível da memória de reconhecimento, e uma oportunidade para que haja uma reavaliação e adaptação dos modelos de memória propostos. Cada vez mais o estudo da memória humana se está a dirigir para a questão dos mecanismos de monitorização e regulação metacognitiva, e o papel que estes têm no desempenho dos indivíduos (Goldsmith & Koriat, 2008; Malmberg, 2008). As medidas proporcionada pela TDS Tipo-2 permitem introduzir neste domínio o mesmo tipo de quantificação encontrada nos modelos de memória de reconhecimento, podendo contribuir para o desenvolvimento de modelos que englobem não apenas o desempenho dos indivíduos, mas também o modo como os indivíduos monitorizam o seu desempenho e o impacto desta monitorização a regulação da acuidade.

Para além da inclusão das medidas Tipo-2, as experiências aqui sugeridas procuram demonstrar a utilidade do paradigma DRM no desenvolvimento dos modelos de memória humana. Este tipo de paradigma permite a manifestação sistemática de falsas memórias, cuja natureza episódica leva a que os mecanismos de monitorização estejam activos, e avaliem as diferentes dimensões de informação presentes nos traços mnésicos de forma a discriminar a sua veracidade. Tradicionalmente a literatura sobre memória de reconhecimento tem abordado os diferentes problemas utilizando listas de palavras não associadas, as quais não levam a que haja uma grande necessidade de monitorização para além da avaliação da familiaridade dos estímulos. Este tipo de material pode ser bastante útil quando aquilo que se procura é o estudo de processos na sua “forma mais pura” sem que hajam interferências provocadas pelo material. Contudo, para o estudo dos mecanismos de monitorização metacognitiva, o tipo de material é essencial, pelo que o avanço neste domínio implica que haja uma rotura com esta tradição.

A presença deste tipo de falhas nos modelos de memória não é surpreendente, dada a dificuldade que existe em produzir modelos que postulem variáveis latentes, tendo muitos apenas disponíveis como dados empíricos informação relativa à acuidade e tempos de resposta. Muitas vezes é necessário relembrar uma noção defendida por George Box muitas vezes esquecida pelos investigadores: “All models are wrong, but some are useful” (citado por Wagenmakers, van der Maas, & Grasman, 2007, p. 20). Os princípios que guiam os dois principais modelos de memória, têm sido bastante úteis no esclarecimento dos processos e mecanismos subjacentes, oferecendo explicações para diferentes fenómenos

como o *list-strength effect* (Ratcliff, Clark, & Shiffrin, 1990) ou o *mirror effect* (Glanzer & Adams, 1985, 1990). Contudo, as questões colocadas por áreas de estudo como as falsas memórias ou a monitorização da fonte levam a que se questione a adequação destes modelos. Diversas propostas têm sido desenvolvidas com o objectivo de suprimir as limitações dos modelos actuais, produzindo extensões destes (Ashby, 2000; Banks, 2000; DeCarlo, 2003), contudo estes modelos apenas têm-se focado na questão da unidimensionalidade da TDS, propondo versões multidimensionais. A questão da monitorização continua então a ser negligenciada pelos investigadores. Como já foi referido anteriormente, a ausência de modelos que explorem a questão da monitorização tem que ver com o tipo de material utilizado na grande maioria dos estudos que abordam a modelação da memória de reconhecimento. A utilização de palavras não associadas pode trazer vantagens na medida em que não produzem qualquer tipo de efeitos que possam comprometer os estudos dos processos mnésicos, contudo tem a desvantagem de eliminar qualquer capacidade aos processos de monitorização em discriminar os estímulos. A familiaridade indistinta que os modelos propõem é uma familiaridade totalmente artificial, e enquanto paradigmas como o DRM não forem utilizados como ferramentas base para o desenvolvimento de modelos de memória, os modelos desenvolvidos serão sempre débeis na sua capacidade explicar os diversos fenómenos de monitorização encontrados na literatura (Johnson, Hashtroudi, & Lindsay, 1993; Jou & Foreman, 2007; Mather, Henkel, & Johnson, 1997; Norman & Schacter, 1997; Schacter, Israel, & Racine, 1999; Starns et al., 2007).

O tipo de soluções que os modelos de memória devem incorporar devido aos resultados encontrados na tarefa tipo-2 não são imediatas, em especial para os modelos mecanicistas. Um possível primeiro passo nos modelos globalistas seria considerar que os diferentes atributos de um traço de memória correspondem a características qualitativamente distintas, cuja calibragem (ver descrição do REM) pode variar de forma independente. Quanto aos modelos dualistas, a possibilidade de os indivíduos utilizarem de forma estratégica a recollecção pode levar a que as diversas medidas de recollecção (confiança, R/K, PDP, para uma revisão ver Malmberg, no prelo) sejam acompanhadas por medidas de julgamento metamnésico (e.g. γ de Nelson; Nelson, 1984), de forma a perceber a relação entre as expectativas do indivíduo e a manifestação da componente de recollecção.

A necessidade de considerar os mecanismos de monitorização não é uma característica da literatura sobre falsas memórias associativas (Roediger & McDermott, 1995), ou sobre monitorização da fonte (Chalfonte & Johnson, 1996); ocorrendo

igualmente na literatura que aborda a *fluência de processamento*, a qual remete para a facilidade com que um determinado estímulo é processado pelos sistemas cognitivos, e o impacto desta facilidade no julgamento dos indivíduos. Diversos estudos têm vindo a demonstrar ao longo das décadas, que os indivíduos utilizam de forma estratégica a fluência de processamento como uma característica de referência para os seus julgamentos, manifestando aquilo a que se pode designar por *efeito atribucional da fluência* ou *heurística de fluência* (e.g., Benjamin, Bjork, & Hirshman, 1998). Neste efeito, aquilo que é tido em conta não é a informação processada em si mesma, mas a fluência com que esta é processada. Esta utilização estratégica da fluência é encontrada em diversos tipos de julgamentos, como julgamentos de agradabilidade (Zajonc, 1968), veracidade (Begg, Anas, & Farinacci, 1992), claridade (Reber & Schwarz, 1999) ou duração temporal (Reber, Zimmermann, & Wurtz, 2004). Quanto a julgamentos de memória de reconhecimento, diversos estudos têm verificado a utilização da fluência enquanto um indicador da apresentação prévia dos estímulos, tanto através de manipulações subliminares (Jacoby & Whitehouse, 1989) ou supraliminares (Whittlesea & Leboe, 2000).

Os efeitos atribucionais da fluência foram revistos e estruturados por Whittlesea e colaboradores (Whittlesea & Leboe, 2000; Whittlesea & Williams, 2001a; 2001b) tendo sido desenvolvida uma proposta explicativa denominada SCAPE, e apresenta bastantes semelhanças com o modelo Activação-Monitorização (Roediger et al., 2001) Segundo o SCAPE, o desempenho de um participante e as avaliações feitas no decorrer deste têm por base a construção de um modelo mental acerca dos itens e do próprio contexto de teste. Esta proposta pode ser então dividida em duas funções distintas; uma função de produção, da qual resulta o desempenho dos indivíduos, e uma função de avaliação, da qual são originadas experiências subjectivas associadas ao desempenho. A função de produção tem por base a informação disponibilizada pelos traços mnésicos em geral, e as pistas de recuperação proporcionadas pelos estímulos de teste e pelo contexto. A função de avaliação apresenta características diferentes, assumindo um papel de monitorização da integridade da função de produção, levando ao desenvolvimento de expectativas e atitudes face a esta que serão determinantes nas respostas dadas pelos indivíduos.

De acordo com o modelo teórico SCAPE, a função de avaliação toma em consideração a qualidade da informação disponível, e avaliando a congruência do seu conteúdo em função das expectativas estabelecidas. Quando são detectadas discrepâncias no processamento dos estímulos, como no efeito Jacoby-Whitehouse (Jacoby & Whitehouse, 1989), a função de avaliação procura identificar uma origem para a fluência produzida pela primacção subliminar; como não é possível atribuir a uma origem específica presente na fase de teste, esta fluência é

então atribuída a uma hipotética apresentação prévia do estímulo, transformando a fluência de processamento num sentimento de familiaridade. Contudo, a função de avaliação pode basear-se nas suas expectativas face à presença de fluência perceptiva, e inibir o seu efeito no falso reconhecimento caso esta considere que determinados níveis de fluência são pouco verosímeis (e.g., Westerman, Lloyd, & Miller,., 2002). O efeito atribucional da fluência apresenta-se assim como resultado de um processo de monitorização metacognitiva, através de um processo de atribuição e de comparação entre fluências percebidas e esperadas, integrado num julgamento que inclui uma avaliação do tipo de estímulos a distinguir, do contexto de teste, e do nível de desempenho previsto. O modelo teórico SCAPE possui bastantes semelhanças com o modelo teórico Activação-Monitorização, as quais não são mera coincidência. A necessidade de incluir os mecanismos de monitorização verifica-se em ambas as áreas, necessidade essa que não é suprida pelos modelos psicométricos actuais. Deste modo, a introdução da TDS tipo-2 num paradigma que envolva a manipulação da fluência de processamento, como o desenvolvido por Jacoby e Whitehouse (1989). Dado que a fluência de processamento associada aos estímulos influencia a familiaridade destes, levanta-se a questão sobre a capacidade dos indivíduos em conseguirem ignorar a fluência enquanto uma evidência diagnóstica de forma a obterem um desempenho positivo na tarefa tipo-2, ou então se a obtenção deste desempenho positivo não pode ocorrer com base na componente de familiaridade, sendo necessária a componente de recollecção. Tal como na literatura relativa ao paradigma DRM, existe uma necessidade premente em avaliar tanto as capacidades reais, como o tipo de componentes subjacentes ao desempenho na tarefa tipo-2.

A extensão da aplicação da TDS tipo-2 não se restringe à fluência de processamento, existindo outras áreas que podem beneficiar da sua inclusão. Outro exemplo disso é o estudo do processamento distintivo e relacional (Arndt & Reder, 2003; Hunt, 2003; Hunt & Einstein, 1981; Hunt & McDaniel, 1993); o processamento distintivo está normalmente associado a níveis de acuidade superiores aos obtidos com o processamento relacional, existindo contudo uma relação de complementariedade entre eles. Hunt (2003) testou a capacidade dos indivíduos em discriminarem duas listas de palavras (reconhecendo apenas os itens da lista 2), sendo que o tipo de processamento associado a estas era manipulado. Os resultados obtidos mostraram que o desempenho dos participantes era superior quando as listas foram processadas de diferentes formas, o que sugere que a presença de diferentes evidências nos diferentes itens permite uma melhor discriminação. Contudo verificou-se uma acuidade superior nos indivíduos que processaram apenas a lista 2 de forma distintiva, do que quando apenas a lista 1 foi processada de forma distintiva. Tendo em conta que a diferença de processamento se mantém em ambas as condições, seria de esperar que não existissem diferenças significativas entre as duas. Hunt

(2003) considera que esta diferença reflecte uma diferença fundamental entre o papel do processamento distintivo na identificação correcta dos itens, e a rejeição correcta de distractores. A identificação correcta emerge das diferenças existentes num contexto de semelhança, enquanto que a rejeição é beneficiada por diferenças em termos de processamento entre os estímulos-alvo e os distractores potenciais

Contudo, podemos questionar-nos sobre até que ponto esta diferença se baseia numa utilização estratégica da informação distintiva, e não numa diferença intrínseca entre a identificação correcta de itens e a rejeição de distractores, como argumenta Hunt (2003)? Através da inclusão de uma tarefa tipo-2 é possível testar as duas explicações, dado que se estivermos perante uma diferença intrínseca entre a aceitação e rejeição, então estas diferenças devem-se manter na tarefa tipo-2. Por outro lado, se aquilo que está em causa é uma mera estratégia de resposta, então estas diferenças obtidas na tarefa de reconhecimento devem esbater-se na tarefa tipo-2, não se verificando uma diferença significativa entre as duas condições.

Outra aplicação seria interessante da TDS tipo-2 seria ao nível da heurística da distintividade, a qual não se manifesta quando a apresentação de figuras e palavras na fase de estudo é manipulada intra-participantes (Schacter, Israel, & Racine, 1999); Apesar de se considerar que uma estratégia baseada exclusivamente na presença de informação de natureza pictórica não é muito diagnóstica quando apenas parte dos itens estudados foi apresentada sob a forma de figuras, a informação pictórica deve estar presente nos traços mnésicos. Assim, é perfeitamente razoável que os participantes sejam capazes de utilizar essa informação pictórica numa tarefa tipo-2 para identificar as suas respostas incorrectas, nomeadamente as omissões cometidas. Se tal situação se verificar, devem-se observar diferenças entre a condição em que são apenas apresentadas palavras e condição mista (figuras + palavras), a qual não se verifica na tarefa de reconhecimento (Schacter, Israel, & Racine, 1999). O facto de uma estratégia frugal não ser viável dado que se baseia em informação pouco diagnóstica não implica que essa informação não seja utilizada no julgamento individual de cada item, e se for o caso, da cada resposta. Se tal for o caso, a TDS tipo-2 deve permitir a sua observação.

Para além de aplicações da TDS tipo-2 a diferentes áreas, é igualmente possível estudar o efeito que determinadas manipulações podem ter no desempenho dos indivíduos na tarefa tipo-2. Um tipo de manipulação cujo efeito é bastante importante para a compreensão dos mecanismos de monitorização é a manipulação do contexto (Murmane & Phelps, 1994); a disponibilização de pistas contextuais aquando da recuperação mnésica leva a melhorias no desempenho dos indivíduos em tarefas de reconhecimento, tendo sido encontradas diferenças com diversos contextos como música de fundo, odores, tamanhos ou tipos de letra, localizações do estímulo no ecrã, ou modalidade perceptiva (Dodson &

Shimamura, 2000). A manipulação do contexto na tarefa de reconhecimento e na tarefa tipo-2 permitiria avaliar o efeito provocado pela manipulação do contexto se verifica na tarefa tipo-2, e até que ponto a utilização de um contexto diferente na tarefa tipo-2 leva a valores inferiores de $d'2$, quando comparados com uma condição experimental em que o contexto é o mesmo encontrado na fase de teste; ou então se a reutilização do contexto de aprendizagem na tarefa tipo-2 melhora a sensibilidade discriminativa dos mecanismos de monitorização.

Existem assim diversas áreas em que a introdução da TDS tipo-2 pode trazer um contributo importante para o estudo dos mecanismos de monitorização envolvidos. Este contributo é especialmente importante se considerarmos que a capacidade de regulação da informação utilizada é uma característica que pode variar entre indivíduos. Cada vez mais, a investigação em memória é caracterizada pela noção de que a recordação é uma *skill*, que pode se pode manifestar de diferentes modos, dada a sua natureza ideográfica (Malmberg, 2008). De facto, todo o 48º volume do prestigiado título “Psychology of Learning and Motivation” (Benjamin & Ross, 2008) foi dedicado a esta questão, o que é revelador da sua predominância cada vez maior no meio científico. A TDS tipo-2 promete ser uma ferramenta importante no estudo desta questão, dado que permite medir a sensibilidade discriminativa dos mecanismos de monitorização que estão na base da regulação dos processos mnésicos, ou seja, permite uma medida objectiva dessa *skill*. A monitorização vem assim assumir um papel central na investigação da memória humana, papel que há muito lhe era devido.

Referências Bibliográficas

- Arndt, J., & Reder, L. M. (2003). The effect of distinctive visual information on false recognition. *Journal of Memory & Language*, 48, 1-15.
- Ashby, F. G. (2000). A stochastic version of general recognition theory. *Journal of Mathematical Psychology*, 44, 310-329.
- Atkinson, R. C. & Shiffrin, R. M. (1971). The control of short-term memory. *Scientific American*, 224, 82-90.
- Banks, W. P. (1970). Signal Detection theory and human memory. *Psychological Bulletin*, 74, 81-99.
- Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science*, 11, 267-273.
- Baranski, J. V., & Petrusic, W. M. (1998). Probing the locus of confidence judgments: Experiments on the time to determine confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 24, 929-945.
- Batchelder, W.H. and Riefer, D.M. (1999) Theoretical and empirical review of multinomial processing tree modeling. *Psychonomic Bulletin & Review*, 6, 57-86.
- Bayen, U. J., & Murnane, K. (1996). Aging and the use of perceptual and temporal information in source memory tasks. *Psychology and Aging*, 11, 293-303.
- Begg, I. M., Anas, A., & Farinacci, S. (1992). Dissociation of processes in belief: Source recollection, statement familiarity, and the illusion of truth. *Journal of Experimental Psychology: General*, 121, 446-458.

- Benjamin, A. S. (2001). On the dual effects of repetition on false recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 27, 941-947.
- Benjamin, A. S. (2003). Predicting and postdicting the effects of word frequency on memory. *Memory & Cognition*, 31, 297-305.
- Benjamin, A. S., Bjork, R. A., & Hirshman, E. (1998). Predicting the future and reconstructing the past: A Bayesian characterization of the utility of subjective fluency. *Acta Psychologica*, 98, 267-290.
- Benjamin, A. S., & Ross, B. H. (2008). *The Psychology of Learning of Motivation, Volume 48: Skill and Strategy in Memory Use*. London: Academic Press
- Bransford, J. D., & Franks, J. J. (1971). The abstraction of linguistic ideas. *Cognitive Psychology*, 2, 331-350.
- Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence-accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, 7, 26-48.
- Chalfonte, B. L., & Johnson, M. K. (1996). Feature memory and binding in young and older adults. *Memory & Cognition*, 24, 403-416.
- Chan, J. K. C., McDermott, K. B., Watson, J. M., and Gallo, D. A. (2005). The importance of material-processing interactions in inducing false memories. *Memory & Cognition*, 33, 389-395.
- Chandler, C. C. (1994). Studying related pictures can reduce accuracy, but increase confidence, in a modified recognition test. *Memory & Cognition*, 22, 273-280.
- Clark, S. E. (1997). A familiarity-based account of confidence-accuracy inversions in recognition memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 23, 232-238.
- Clarke, F. R., Birdsall, T. G., & Tanner, W. P., Jr. (1959). Two types of ROC curves and definitions of parameters. *Journal of the Acoustical Society of America*, 31, 629-630.

- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic memory. *Psychological Review*, 82, 407-428.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11, 671-684.
- Craik, F. I. M., & Tulving, E. (1975). Depth of processing and the retention of words in episodic memory. *Journal of Experimental Psychology: General*, 104, 268-294.
- Criss, A.H. and Shiffrin, R.M. (2005). List discrimination and representation in associative recognition. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 31, 1199-1212.
- DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: Theoretical developments with applications to recognition memory. *Psychological Review*, 109, 710-721.
- DeCarlo, L. T. (2003). Source monitoring and multivariate signal detection theory, with a model for selection. *Journal of Mathematical Psychology*, 47, 292-303.
- Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, 58, 17-22.
- Diana, R., Reder, L. M., Arndt, J., & Park, H. (2006). Models of recognition: A review of arguments in favor of a dual process account. *Psychonomic Bulletin & Review*, 13, 1-21.
- Dodson, C. S: & Hege, A. C. (2005). Speeded retrieval abolishes the false-memory suppression effect: Evidence for the distinctiveness heuristic. *Psychonomic Bulletin & Review*, 12, 726-731.
- Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of exact and partial source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 1121-1136.

- Dodson, C. S. & Shimamura, A. P. (2000). Differential effects of cue dependency on item and source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 1023-1944.
- Dosher, B. A. (1981). The effects of delay and interference: A speed-accuracy study. *Cognitive Psychology*, 13, 551-582.
- Dosher, B. A. (1984). Discriminating Preexperimental (Semantic) from Learned (Episodic) Associations: A Speed-Accuracy Study. *Cognitive Psychology* 16, 519-555.
- Ebbinghaus, H. (1885/1964). *Memory: A contribution to experimental psychology*. New York: Dover Publications.
- Estes, W. K. (1994). *Classification and Cognition*. New York: Oxford.
- Ferguson, S. A., Hashtroudi, S., & Johnson, M. K. (1992). Age differences in using source-relevant cues. *Psychology and Aging*, 7, 443–452.
- Gallo, D. A. (2006). *Associative illusions of memory: False memory research in DRM and related tasks*. New York: Psychology Press.
- Gallo, D. A., Meadow, N. G., Johnson, E. L., & Foster, K. T. (2008). Deep levels of processing elicit a distinctiveness heuristic: Evidence from the criterial recollection task. *Journal of Memory and Language*, 58, 1095-1111.
- Gallo, D. A., Perlmutter, D. H., Moore, C. D., & Schacter, d. L. (2008). Distinctive encoding reduces the Jacoby—Whitehouse illusion. *Memory & Cognition*, 36, 461-466.
- Gallo, D. A., Roberts, M. J., & Seamon, J. G. (1997). Remembering words not presented in lists: Can we avoid creating false memories? *Psychonomic Bulletin & Review*, 4, 271-276.
- Galvin, S. J., Podd, J. V., Drga, V., & Whitmore, J. (2003). Type 2 tasks in the theory of signal detectability: Discrimination between correct and incorrect decisions. *Psychonomic Bulletin & Review*, 10, 843-876.

- Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition*, *16*, 309-313.
- Gigerenzer, G., Hoffrage, U., & Kleinbolting, H. (1991). Probabilistic mental models: A Brunswikian theory of confidence. *Psychological Review*, *98*, 506-528.
- Glanzer, M. and Adams, J. K. (1985). The mirror effect in recognition memory. *Memory and Cognition* *13*, 8–20.
- Glanzer, M., and Adams, J. K. (1990). The mirror effect in recognition memory: Data and theory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *16*, 5-16.
- Gluck, M. A. & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*, 227-247.
- Goldsmith, M., & Koriat, A. (2008). The strategic regulation of memory accuracy and informativeness. in A. Benjamin and B. Ross (Eds.) *The Psychology of Learning and Motivation: Skill and Strategy in Memory Use* (pp. 1-60). London: Academic Press
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Gunter, R. W., Ivanko, S. L. , & Bodner, G. E. (2005). Can list context manipulations improve the recognition accuracy in the DRM paradigm? *Memory*, *13*, 862-873.
- Henkel, L. A. & Mather, M. (2007). Memory attributions for choices: How beliefs shape our memories. *Journal of Memory and Language*, *57*, 163-176.
- Higham, P. A. (2007). No special K! A signal detection framework for the strategic regulation of memory accuracy. *Journal of Experimental Psychology: General*, *136*, 1-22.
- Hunt, R. R. (2003). Two contributions of distinctive processing to accurate memory. *Journal of Memory and Language*, *48*, 811-825.

- Hunt, R. R., & Einstein, G. (1981). Relational and item-specific information in memory. *Journal of Verbal Learning and Verbal Behavior*, 20, 497–514.
- Hunt, R. R., & McDaniel, M. A. (1993). The enigma of organization and distinctiveness. *Journal of Memory and Language*, 32, 421–445.
- Hutchison, K. A., & Balota, D. A. (2005). Decoupling semantic and associative information in false memories: Explorations with semantically ambiguous and unambiguous critical lures. *Journal of Memory and Language*, 52, 1-28.
- Israel, L., & Schacter, D. L. (1997). Pictorial encoding reduces false recognition of semantic associates. *Psychonomic Bulletin & Review*, 4, 577-581.
- Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, 30, 513-541.
- Jacoby, L. L., & Whitehouse, K., (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimental Psychology: General*, 118, 126-135.
- Jacoby, L. L., Shimizu, Y., Daniels, K. A., & Rhodes, M. G. (2005). Modes of cognitive control in recognition and source memory: Depth of retrieval. *Psychonomic Bulletin & Review*, 12, 852–857.
- Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin*, 114, 3-28.
- Johnson, M. K., Kounios, J., & Reeder, J. A. (1994). Time-course studies of reality monitoring and recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 1409-1419.
- Johnson, M. K., & Raye, L. C. (1981). Reality monitoring. *Psychological Review*, 88, 67-85.
- Johnson, M. K., Raye, C. L., Foley, M. A., & Kim, J. K. (1982). Pictures and images: Spatial and temporal information compared. *Bulletin of the Psychonomic Society*, 19, 23-26.

- Jou, J., & Foreman, J. (2007). Transfer of learning in avoiding false memory: The roles of warning, immediate feedback, and incentive. *Quarterly Journal of Experimental Psychology*, *60*, 877-896.
- Jou, J., Matus, Y. E., Aldridge, J. W., Rogers, D. M., & Zimmerman, R. L. (2004). How similar is false recognition to veridical recognition objectively and subjectively? *Memory & Cognition*, *32*, 824-840.
- Justin, P., & Olsson, H. (1997). Thurstonian and Brunswikian origins of uncertainty in judgment: A sampling model of confidence in sensory discrimination. *Psychological Review*, *104*, 344-366.
- Koriat, A., & Goldsmith, M. (1996). Memory as something that can be counted versus memory as something that can be counted on. In D. Herrmann, C. McEvoy, C. Hertzog, P. Hertel, and M. Johnson (Eds.) *Basic and applied memory research: Practical applications* (Vol. 2, pp. 3-18). Hillsdale, NJ: Erlbaum.
- Kunimoto, C., Miller, J., & Pashler, H. (2001). Confidence and accuracy of near-threshold discrimination responses. *Consciousness & Cognition*, *10*, 294-340.
- Lindsay, D. S., Read, J. D., & Sharma, K. (1998). Accuracy and confidence in person identification: The relationship is strong when witnessing conditions vary widely. *Psychological Science*, *9*, 215-218.
- Lockhart, R. S. and Murdock, B. B., (1970). Memory and theory of signal detection, *Psychological Review*, *74*, 100-109.
- Luce, P. A., & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing*, *19*, 1-36.
- Lyle, K. B., & Johnson, M. K. (2006). Importing perceived features into false memories. *Memory*, *14*, 197-213.

- Macmillan, N. A. & Creelman, C. D. (1990). Response bias: Characteristics of detection theory, threshold theory, and “nonparametric” indexes. *Psychological Bulletin*, *107*, 401–413.
- Macmillan, N. A. & Creelman, C. D. (2005). *Detection theory: A user’s guide*. Cambridge: Cambridge University Press.
- Macmillan, N. A., Rotello, C. M., Miller, J. O. (2004). The sampling distributions of Gaussian ROC statistics. *Perception & Psychophysics*, *66*, 406–421.
- Malmberg, K. J. (2008). Towards an understanding of individual differences in episodic memory: Modeling the dynamics of recognition memory. In A. Benjamin and B. Ross (Eds.) *The Psychology of Learning and Motivation: Skill and Strategy in Memory Use*, (pp. 313-349). London: Academic Press.
- Malmberg, K. J. (no prelo). Recognition memory: A review of critical findings and an integrated proposal for relating them. *Cognitive Psychology*.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, *87*, 252-271.
- Mather, M., Henkel, L. A., & Johnson, M. K. (1997). Evaluating characteristics of false memories: Remember/know judgments and memory characteristics questionnaire compared. *Memory & Cognition*, *25*, 826-837.
- Mather, M., Shafir, E., & Johnson, M. K. (2000). Misrememberance of options past: source monitoring and choice. *Psychological Science*, *11*, 132–138.
- Mather, M., Shafir, E., & Johnson, M. K. (2003). Remembering chosen and assigned options. *Memory & Cognition*, *31*, 422–433.
- McCabe, D. P., & Smith, A. D. (2002). The effect of warnings on false memories in young and older adults. *Memory & Cognition*, *30*, 1065–1077.

- McDermott, K. B. (1996). The persistence of false memories in list recall. *Journal of Memory & Language*, 35, 212-230.
- McDermott, K. B. (1997). Priming on perceptual implicit memory tests can be achieved through presentation of associates. *Psychonomic Bulletin & Review*, 4, 582-586.
- McDermott, K. B., & Roediger, H. L., III (1998). Attempting to avoid illusory memories: Robust false recognition of associates persists under conditions of explicit warnings and immediate testing. *Journal of Memory & Language*, 39, 508-520.
- McDermott, K. B., & Watson, J. M. (2001). The rise and fall of false recall: The impact of presentation duration. *Journal of Memory & Language*, 45, 160-176.
- McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information in item recognition: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 25, 563-582.
- Meiser, T., & Bröder, A. (2002). Memory for multidimensional source information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 116-137.
- Meiser, T., Satler, C., & Weißer, K. (2008). Binding of multidimensional context information as a distinctive characteristic of *remember* judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 32-49.
- Metz, C. E. (1998). Rokit computer program (Computer software). Department of Radiology, University of Chicago. Retrieved from <http://www.radiology.uchicago.edu/krl/toppage11.htm>
- Mueller, S. T., & Weidemann, C. T. (no prelo). Decision noise: An explanation for observed violations of Signal Detection Theory. *Psychonomic Bulletin and Review*.
- Mulligan, N. W., & Hirshman, E. (1997). Measuring the bases of recognition memory: An investigation of the process-dissociation framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 280-304.

- Murnane, K., & Phelps, M. P. (1994). When does a different environmental context make a difference in recognition? A global activation model. *Memory & Cognition*, 22, 584-590
- Murdock, B. B. (1965). Signal detection theory and short-term memory. *Journal of Experimental Psychology*, 70, 443-447.
- Nelson, T. O. (1984). A comparison of current measures of feeling-of-knowing accuracy. *Psychological Bulletin*, 95, 109-133.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of Learning and Motivation*, 26, 125-173.
- Neuschatz, J. S., Payne, D. G., Lampinen, J. M., & Toggia, M. P. (2001). Assessing the effectiveness of warnings and the phenomenological characteristics of false memories. *Memory*, 9, 53-71.
- Norman, K. A., & Schacter, D. L. (1997). False recognition in younger and older adults: Exploring the characteristics of illusory memories. *Memory & Cognition*, 25, 838-848.
- Pastore, R. E., Crawley, E. J., Berens, M. S., & Skelly, M. (2003). "Nonparametric" A' and other modern misconceptions about signal detection theory. *Psychonomic Bulletin & Review*, 10, 556-569.
- Payne, D. G., Elie, C. J., Blackwell, J. M., & Neuschatz, J. S. (1996). Memory illusions: Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory & Language*, 35, 261-285.
- Pollack, I. (1959). On indices of signal and response discriminability. *Journal of the Acoustical Society of America*, 31, 1031.
- Qin, J., Raye, C. L., Johnson, M. K., & Mitchell, K. J. (2001). Source ROCs Are (Typically) Curvilinear: Comment on Yonelinas (1999). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1110-1115.

- Raaijmakers, J.G.W. & Shiffrin, R.M. (2002). Models of memory. In H. Pashler & D. Medin (Eds.), *Stevens' Handbook of Experimental Psychology, Third Edition, Volume 2: Memory and Cognitive Processes* (pp. 43-76). New York: John Wiley & Sons, Inc.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.
- Ratcliff, R. (1988). Continuous versus discrete information processing: Modeling the accumulation of partial information. *Psychological Review*, 95, 238-255.
- Ratcliff, R. (2006). Modeling response signal and response time data. *Cognitive Psychology*, 53, 195-237.
- Ratcliff, R., & Rouder, J. F. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9, 347-356.
- Ratcliff, R., Clark, S., & Shiffrin, R.M. (1990). The list-strength effect: I. Data and discussion. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 163-178.
- Ratcliff, R., Sheu, C. F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review*, 99, 518-535.
- Ratcliff, R., Van Zandt, T., & McKoon, G. (1999). Connectionist and diffusion models of reaction time. *Psychological Review*, 106, 261-300.
- Raye, C. L., Johnson, M. K., & Taylor, T. H. (1980). Is there something special about memory for internally generated information? *Memory & Cognition*, 8, 141-148.
- Read, J. D. (1996). From a passing thought to a false memory in 2 minutes: Confusing real and illusory events. *Psychonomic Bulletin & Review*, 3, 105-111.
- Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8, 338-342.

- Reber, R., Zimmermann, T. D., & Wurtz, P. (2004). Judgments of duration, figure-ground contrast and size for words and nonwords. *Perception and Psychophysics*, *66*, 1105–1114.
- Reyna, V. F., & Brainerd, C. J. (1995). Fuzzy-trace theory: An interim synthesis. *Learning and Individual Differences*, *7*, 1-75.
- Robinson, K. J., & Roediger, H. L. (1997). Associative processes in false recall and false recognition. *Psychological Science*, *8*, 231–237.
- Rhodes, M. G., & Anastasi, J. S. (2000). The effects of a levels-of-processing manipulation on false recall. *Psychonomic Bulletin & Review*, *7*, 158-162.
- Roediger, H. L. (1980). Memory metaphors in cognitive psychology. *Memory & Cognition*, *8*, 231-246.
- Roediger, H.L. (1996). Memory illusions. *Journal of Memory and Language*, *35*, 76-100.
- Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*, 803-814.
- Roediger, H. L., Watson, J. M., McDermott, K. B., & Gallo, D. A. (2001). Factors that determine false recall: A multiple regression analysis. *Psychonomic Bulletin & Review*, *8*, 385–407.
- Schacter, D. L., Israel, L., & Racine, C. (1999). Suppressing false recognition in younger and older adults: The distinctiveness heuristic. *Journal of Memory & Language*, *40*, 1-24.
- Seamon, J. G., Luo, C. R., & Gallo, D. A. (1998). Creating false memories of words with or without recognition of list items. *Psychological Science*, *9*, 20–26.
- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*, 127–190.

- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM—retrieving effectively from memory. *Psychonomic Bulletin & Review*, *4*, 145–166.
- Sommers, M. S., & Lewis, B. P. (1999). Who really lives next door: Creating false memories with phonological neighbors. *Journal of Memory & Language*, *40*, 83-108.
- Starns, J. J., & Hicks, J. L. (2005). Source dimensions are retrieved independently in multidimensional monitoring tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 1213–1220.
- Starns, J. J., Lane, S. M., Alonzo, J. D., & Roussel, C. C. (2007). Metamnemonic control over the discriminability of memory evidence: A signal-detection analysis of warning effects in the associative list paradigm. *Journal of Memory & Language*, *56*, 592-607.
- Swets, J. A., Tanner, W. P., Jr., & Birdsall, T. G. (1961). Decision processes in perception. *Psychological Review*, *68*, 301-340.
- Thapar, A., & McDermott, K. B. (2001). False recall and false recognition induced by presentation of associated words: Effects of retention interval and level of processing. *Memory & Cognition*, *29*, 424-432.
- Toglia, M. P., Neuschatz, J. S., & Goodwin, K. A. (1999). Recall accuracy and illusory memories: When more is less. *Memory*, *7*, 233-256.
- Tolman, E.C. & Brunswik, E. (1935). The organism and the causal texture of the environment. *Psychological Review*, *42*, 43-77.
- Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. New York: Cambridge University Press.
- Tulving, E. (1985). Memory and consciousness. *Canadian Psychologist*, *26*, 1-12.
- Tulving, E., & Craik, F. I. M. (2005). *Oxford Handbook of Memory*. Oxford: Oxford University Press.

- Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*, 352–373.
- Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: The leaky, competing accumulator model. *Psychological Review*, *108*, 550–592.
- Van Zandt, T. (2000). ROC curves and confidence judgments in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 582-600.
- VanZandt, T., & Maldonado-Molina, M. M. (2004). Response reversals in recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 1147-1166.
- Vickers, D., & Packer, J. S. (1982). Effects of alternating set for speed of accuracy on response time, accuracy, and confidence in a unidimensional discrimination task. *Acta Psychologica*, *50*, 179-197.
- Wagenmakers, E.-J., van der Maas, H. L. J., & Grasman, R. P. (2007). An EZ-diffusion model for response time and accuracy. *Psychonomic Bulletin & Review*, *14*, 3-22.
- Watson, J. M., Balota, D. A., & Roediger, H. L., III. (2003). Creating false memories with hybrid lists of semantic and phonological associates: Over-additive false memories produced by converging associative networks. *Journal of Memory & Language*, *49*, 95–118.
- Westerman, D. L., Lloyd, M. L., & Miller, J. M. (2002). On the attribution of perceptual fluency in recognition memory: The role of expectation. *Journal of Memory & Language*, *47*, 607-617.
- Whittlesea, B. W. A., & Leboe, J. P. (2000). The heuristic basis of remembering and classification: Fluency, generation and resemblance. *Journal of Experimental Psychology: General*, *129*, 84–106.

- Whittlesea, B. W. A., & Williams, L. D. (2001a). The discrepancy–attribution hypothesis: I. The heuristic basis of feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1–13.
- Whittlesea, B. W. A., & Williams, L. D. (2001b). The discrepancy–attribution hypothesis: II. Expectation, uncertainty, surprise and feelings of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 14–33.
- Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, *114*, 152–176.
- Yonelinas, A. P. (1997). Recognition memory ROCs for item and associative information: The contribution of recollection and familiarity. *Memory & Cognition*, *25*, 747–763.
- Yonelinas, A. P. (2001). Consciousness, control and confidence: The three Cs of recognition memory. *Journal of Experimental Psychology: General*, *130*, 361–379.
- Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language*, *46*, 441–517.
- Yonelinas, A. P., & Jacoby, L. L. (1996). Noncriterial recollection: Familiarity as automatic, irrelevant recollection. *Consciousness and Cognition*, *5*, 131–141.
- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology Monograph Supplement*, *9*, 1–27.

ANEXOS

ANEXO 1

O MOTIVO DA COMPOSIÇÃO DAS LISTAS DRM

A questão da composição das listas DRM na fase de estudo e de teste surge de problemas que foram encontrados em alguns estudos de falsas memórias, sendo estes problemas evidentes na investigação de Thapar e McDermott (2001). Segundo estes autores, o processamento “profundo” leva a um aumento conjunto das memórias verídicas e das falsas memórias tanto em recordação livre com em reconhecimento (ver resultados no Quadro 1). Estes resultados são normalmente interpretados como uma consequência de uma activação das relações semânticas e associativas, enfatizada pelo processamento profundo, em contraste com o processamento superficial, o qual promove a codificação de informação perceptiva. Esta interpretação pode no entanto ser questionada, se reanalisarmos os resultados obtidos de Thapar e McDermott segundo a TDS (ver Quadro 1). Os índices da TDS mostram que para o processamento profundo existe uma maior sensibilidade discriminativa e um critério de resposta liberal, enquanto que para as listas processadas superficialmente se verifica uma sensibilidade discriminativa praticamente nula, acompanhada por um critério de resposta estrito.

Processamento	Acertos	Falsos Alarmes	D'	C
Superficial	.34	.32	.06	.44
Profundo	.90	.67	0,84	-.86

Quadro 1. Proporções de respostas obtidas por Thapar & McDermott (2001), e as respectivas medidas TDS .

As medidas TDS mostram um padrão curioso; apesar de existir um aumento das falsas memórias, a sensibilidade discriminativa dos participantes é superior quando estes executam um processamento profundo do quando executam um processamento superficial, ou seja, à partida os participantes conseguiriam demonstrar uma menor proporção de falsos alarmes nas listas processadas profundamente do que nas listas processadas superficialmente, o que contradiz por completo as conclusões de Thapar e McDermott. O critério de resposta liberal demonstrado nas listas processadas profundamente, em comparação com o critério estrito encontrado nas listas processadas superficialmente leva a que se levante a hipótese de que os participantes então a utilizar um critério de decisão temático ou “*gist*” na aceitação das palavras. Tendo em conta que a informação relevante para esse tipo de julgamento se encontra mais disponível para listas processadas profundamente, este

critério de decisão deste tipo deverá levar a uma aceitação generalizada das palavras pertencentes a listas processadas profundamente e distractores associados; bem como a uma rejeição das palavras pertencentes às listas processadas de forma superficial e aos seus distractores associados.

Um possível motivo para a presença deste critério de decisão pode ser encontrado na constituição das listas de teste. Estas são constituídas por apenas 24 palavras apresentadas anteriormente (uma por lista), e por 120 distractores. Estes distractores são constituídos pelas 24 palavras críticas, por 24 palavras pouco associadas às listas, e por 72 palavras não associadas. De facto, a grande maioria dos distractores não apresenta qualquer relação com a lista de teste, pelo que uma estratégia que tenha como base um critério, consegue rejeitar à partida a grande maioria dos distractores, em especial as palavras não associadas. De facto, a proporção de palavras não associadas que foi erroneamente reconhecida pelos participantes foi de apenas .04. Assim, o aumento das falsas memórias pode não ser resultado de uma estratégia de decisão e não de um aumento real das falsas memórias. É de notar que este critério de decisão tem por base a constituição da lista de teste e não diferenças no tipo de processamento, pelo que alterações na constituição da lista devem levar a mudanças na implementação deste tipo de estratégias.

Tendo estes aspectos em conta, a inclusão de palavras associadas como distractores no teste de reconhecimento surge como um modo mais adequado de estudar o desempenho dos indivíduos, dado que evita a manifestação de estratégias como as encontradas em Thapar & McDermott (2001). Apesar de tal opção implicar que sejam apresentadas menos palavras das listas na fase de estudo, o que diminui a proporção de falsos alarmes, é possível evitar a manifestação de estratégias de resposta. Quanto à divisão das listas, esta deve ser feita de forma intercalada (palavras pares vs. palavras ímpares) dado que permite uma distribuição mais homogénea.

ANEXO 2

EXEMPLO DE LISTA DE ASSOCIADOS SEMÂNTICOS (LISTA DRM)

SENTAR
MESA
DESCANSO
MADEIRA
BANCO
CONFORTO
ASSENTO
SOFÁ
OBJECTO
PERNAS
CANSAÇO
MOBÍLIA
ALMOFADA
POLTRONA
SALA

PALAVRA-CRÍTICA: CADEIRA