

Sensibilidade e bom senso: Princípios fundamentais da teoria de detecção de sinal na investigação em Psicologia

David Van der Kellen
Ludmila Duarte Nunes
Leonel Garcia-Marques

Faculdade de Psicologia e de Ciências da Educação da Universidade de Lisboa

Resumo

Com várias aplicações, a Teoria de Detecção de Sinal (TDS) proporciona um método eficaz de análise do desempenho em identificação de sinal na presença de ruído. Uma das aplicações mais conhecidas da TDS é na área da memória, especificamente nos testes de reconhecimento. Ao permitir calcular a sensibilidade à existência de determinados atributos associados ao sinal e o grau de especificidade da detecção do mesmo, a TDS facilita a compreensão dos padrões de resposta obtidos em testes de memória. Este artigo faz uma revisão detalhada da TDS e dos seus pressupostos e medidas associadas, permitindo uma utilização adequada da teoria em paradigmas de reconhecimento do tipo sim/não e em paradigmas de escolha forçada.

Palavras-chave: Análise de desempenho, Curvas ROC, Detecção de sinal, Memória, Reconhecimento.

Abstract

Providing numerous applications, the Signal Detection Theory (SDT) offers an effective performance analysis in tasks of signal identification in the presence of noise. One of the most known applications of the SDT is in the field of memory, especially in recognitions tests, making possible to understand the response patterns identified. In this sense, this paper presents a detailed revision of the SDT, its assumptions and associated measures, allowing a correct use of the theory in yes-no and forced choice paradigms.

Key words: Memory, Performance analysis Recognition, ROC curves, Signal detection.

A correspondência relativa a este artigo deverá ser enviada para: David Van der Kellen, Faculdade de Psicologia e de Ciências da Educação, Alameda da Universidade, 1649-013 Lisboa; E-mail: davekellen@gmail.com

Introdução

Imaginemos uma situação experimental em que participantes lêem uma lista de palavras, sendo-lhes em seguida pedido que reconheçam essas mesmas palavras numa lista de teste, na qual estão igualmente incluídas palavras não apresentadas, ou distractores. Nesta situação, os participantes têm de discriminar as palavras apresentadas dos distractores, o que implica que utilizem a informação armazenada em memória. Num teste de reconhecimento, habitualmente, os participantes baseiam as suas respostas no sentimento de familiaridade que a informação armazenada em memória produz e que aumenta após um episódio recente em que essa informação esteja presente. Daí que, em média, os itens produzam um sentimento de familiaridade superior ao dos distractores. Mas a tarefa dos participantes é menos simples do que parece porque quer os itens quer os distractores variam no sentimento de familiaridade que produzem antes dos participantes lerem a lista dos itens. Daí que, apesar de, em média, os itens produzirem nos participantes um sentimento de familiaridade maior do que os distractores, muitas vezes, os distractores mais familiares *a priori* podem transmitir, durante o teste de reconhecimento, um sentimento de familiaridade maior do que os itens menos familiares *a priori*. Por isso, os participantes não podem cingir as suas respostas ao componente mnésico (o sentimento de familiaridade), têm também que escolher um critério de resposta acima do qual decidem reconhecer como itens e abaixo do qual decidem rejeitar como distractores os estímulos apresentados. O estabelecimento desse critério é chamado componente decisório subjacente ao desempenho dos participantes num teste de reconhecimento. Note-se que qualquer critério que seja o critério escolhido, o participante cometerá sempre erros desde que os distractores mais familiares sejam mais familiares do que os itens menos familiares. Esta situação leva a que os indivíduos regulem o seu desempenho, estabelecendo um compromisso entre memórias verídicas e falsas sob a forma de um critério de resposta que pode ser mais liberal ou conservador consoante o contexto e os objectivos.

Assim, numa investigação que contemple reconhecimento de informação, torna-se necessário ter em conta tanto os processos mnésicos como os processos decisórios. A negligência desta distinção pode levar a interpretações distorcidas dos dados e, conseqüentemente, a conclusões erradas. Uma ferramenta que permite não apenas distinguir estes processos mas também medi-los de forma independente é a Teoria de Detecção de Sinal (TDS), cuja importância na investigação psicológica tem crescido ao longo dos anos.

Esta teoria parte do princípio de que a maioria dos processos de decisão ocorre na presença de alguma incerteza. Assim, procura proporcionar uma linguagem precisa em termos estatísticos, matemáticos e gráficos, que descreva e analise a tomada de decisão na presença de incerteza (Green & Swets, 1966). A TDS tem sido utilizada em diferentes domínios, da electrónica às telecomunicações, passando mesmo pelo diagnóstico médico e pela decisão estatística. A sua introdução na Psicologia fez-se através de investigações em Psicofísica, alargando-se a outras áreas como a memória (Macmillan & Creelman, 2005), na qual tem tido uma enorme preponderância nos últimos 40 anos.

A TDS tem sido principalmente utilizada na análise do desempenho de participantes na identificação do sinal, quando na presença de ruído. No caso específico dos testes de reconhecimento, nos quais se pede aos participantes para decidirem se o estímulo apresentado já havia sido anteriormente estudado, o sinal corresponde ao sentimento de familiaridade gerado por um estímulo memorizado, e o ruído ao sentimento de familiaridade gerado por um novo estímulo. Neste caso, os principais objectivos da TDS são estimar a força do sinal relativamente à do ruído e identificar a estratégia de resposta utilizada pelos participantes.

Origem e pressupostos da teoria de detecção de sinal

Tendo sido desenvolvida no campo da engenharia electrotécnica e tendo por base a teoria estatística da decisão, a TDS surge inicialmente durante 2ª Guerra Mundial, com o intuito de analisar a capacidade dos radares para detectar os aviões inimigos. É nesta altura que os investigadores, preocupados com os erros cometidos pelos humanos e pelas máquinas, como os radares, em tarefas de detecção de sinal, concluíram que, em qualquer sistema onde exista ruído, não é possível maximizar as respostas positivas correctas e minimizar as respostas positivas falsas, ao mesmo tempo. É então no campo das redes de comunicações que a TDS conhece os seus desenvolvimentos, quando os engenheiros electrotécnicos tentam encontrar uma forma de determinar a sensibilidade dos sinais das redes de comunicação, na presença de ruído, mas sem que o enviesamento de resposta contamine os resultados (Green & Swets, 1966).

Em 1966, a TDS é adaptada de uma forma sistematizada à Psicofísica, por John Swets e David Green. A Psicofísica, fundada por Fechner, baseava-se no relato subjectivo das sensações. Ora, este relato subjectivo trazia, entre outros, o problema de os participantes poderem dar as suas respostas baseadas numa certa “adivinhação” (*guessing*) fundamentada pela taxa de detecção de sinal e não na sensação *per se*. Deste modo, a TDS torna-se útil, ao esclarecer o comportamento do indivíduo na identificação de limiares, ou seja, numa tarefa de decisão. Mais tarde, a TDS deixa de ser aplicada apenas à percepção e às sensações, passando a ser utilizada também em áreas relacionadas com a aprendizagem e a retenção humanas, como a retenção verbal (Egan citado por Banks, 1970). Logo, áreas como a memória humana beneficiaram também da utilização da TDS.

Assim, aplicada à Psicofísica e, mais tarde, à Psicologia Cognitiva, a TDS toma como principal intuito a explicação da forma como os humanos discriminam os sinais do ruído e procura tornar possível a utilização de medidas para uma correcta quantificação desta capacidade discriminatória.

Segundo a TDS, a detecção de um determinado alvo depende da capacidade de distingui-lo do ruído envolvente. Esta detecção é feita através da **sensibilidade** (por vezes designada por discriminabilidade na literatura) à existência de determinados atributos associados ao alvo a detectar, cuja presença evidencia este último. No entanto, podem existir configurações semelhantes de atributos, com uma origem diferente da dos alvos a detectar, podendo levar a uma falsa detecção destes. Devido a esta situação, torna-se necessário considerar a capacidade de extracção das características específicas do alvo que permitem discriminá-lo em relação ao ruído, ou seja, o nível de **especificidade** na detecção (alguns autores designam por critério ou enviesamento). A adopção de níveis de especificidade conduz então a um enviesamento das respostas e subsequente redução das falsas identificações de alvos mas, por outro, à possível não-deteção de alvos. Assim, existem custos associados ao estabelecimento de um critério de especificidade, variando proporcionalmente em relação ao nível de discriminação adoptado.

Tendo em conta estes constrangimentos na relação entre sensibilidade e especificidade, é possível identificar quatro tipos de resultados em função do tipo de estímulo e da resposta dada, representada sob a forma de uma matriz na Figura 1.

| | | RESPOSTA | |
|----------|-------|-------------------|------------------------|
| | | “SIM” | “NÃO” |
| ESTÍMULO | ALVO | ACERTOS (HITS) | OMISSÕES |
| | RUÍDO | FALSOS ALARMES | REJEIÇÕES CORRECTAS |

Figura 1. Matriz de estímulo-resposta

Quando o alvo se encontra presente, e este é identificado, estamos na presença de um acerto (hit); no caso em que apenas existe ruído e há uma resposta no sentido da presença do alvo, ocorre um falso alarme. Podemos ainda considerar resultados em que há uma omissão, quando o alvo está presente e, contudo, não é detectado; e, por último, quando ocorre uma rejeição correcta, dada a ausência do alvo e a sua não detecção.

Tendo em conta estes possíveis resultados, e a forma como os níveis de sensibilidade e especificidade influenciam a detecção, torna-se então pertinente considerar o modo como a variação destes níveis determina a frequência dos quatro tipos de resultados da matriz.

Consideremos então estes pressupostos da TDS no contexto de uma tarefa de reconhecimento. Num teste de reconhecimento, o experimentador fornece o item alvo e a função do participante é recuperar o contexto de apresentação (ou não) deste. O objectivo da TDS neste tipo de tarefas é o de distinguir os processos mnésicos dos decisórios, e a prevalência destes no desempenho dos participantes experimentais.

Um dos tipos de teste de reconhecimento mais comuns é o chamado teste de reconhecimento do tipo sim/não (*yes-no recognition test*), no qual os participantes estudam uma lista de itens e, posteriormente, na fase de teste, são-lhes apresentados vários itens (estudados e não-estudados) sendo-lhes então pedido que decidam se cada um desses itens pertencia à lista estudada, respondendo “sim” ou “não”. Este tipo de teste de reconhecimento é especialmente útil na compreensão dos pressupostos TDS, já que permite um conjunto de manipulações, tanto ao nível da sensibilidade como da especificidade, explicitados mais adiante.

Consideremos então uma investigação hipotética, em que se procurou avaliar o desempenho de uma amostra de participantes no reconhecimento de uma categoria de estímulos K. No teste de reconhecimento, era perguntado aos sujeitos se cada um dos estímulos da fase de teste tinha sido previamente estudado ou não. Deste teste resulta uma tabela semelhante à matriz da Figura 1. Imaginemos então que, neste teste, foram obtidos os dados apresentados na Tabela 1.

Tabela 1

Frequência média de respostas, no hipotético teste de reconhecimento de estímulos da categoria K

| ESTÍMULO | RESPOSTA-AMOSTRA RECONHECIMENTO K | | | |
|----------|-----------------------------------|-------|-------|-----|
| | “SIM” | “NÃO” | TOTAL | |
| ANTIGO | | 50 | 50 | 100 |
| NOVO | | 15 | 85 | 100 |

Na Tabela 1, apenas dois resultados são relevantes para a análise dos níveis de sensibilidade e especificidade dos participantes experimentais, nomeadamente os acertos (hits) e os falsos alarmes, a partir de agora designados por H e F , respectivamente. Os outros dois resultados não são necessários para que se efectue a análise, já que podem ser obtidos através dos valores de H e F , dada a sua complementaridade em relação a estes, quando representados sob a forma de proporções, como se verifica na Tabela 2. Estas proporções podem ser encaradas como estimativas de probabilidade dada a possibilidade de representar tanto H como F através de probabilidades condicionais, designadamente $H=P(\text{SIM}/\text{ANTIGO})$ e $F=P(\text{SIM}/\text{NOVO})$. Através destas proporções é então possível encontrar medidas que avaliem a sensibilidade e especificidade destes resultados no teste de reconhecimento. As medidas de sensibilidade podem basear-se em pressupostos e tipos de dados distintos, tendo por isso vantagens e

limitações características. O conhecimento destas características permite escolher a medida mais adequada para cada estudo, e analisar criticamente os resultados apresentados na literatura.

Tabela 2

Proporção de respostas no hipotético teste de reconhecimento de estímulos da categoria K

| ESTÍMULO | RESPOSTA-AMOSTRA RECONHECIMENTO K | | |
|----------|-----------------------------------|-------|-------|
| | “SIM” | “NÃO” | TOTAL |
| ANTIGO | .50 | .50 | 1.00 |
| NOVO | .15 | .85 | 1.00 |

Medição da sensibilidade

Um método utilizado antes da introdução da TDS consistia na correcção dos resultados obtidos através da diferença entre a proporção de acertos (hits) e a proporção de falsos alarmes, ou seja, $H - F$. Apesar de se encontrar ultrapassado, este método permite perceber o princípio subjacente às medidas de sensibilidade, nomeadamente a necessidade de considerar os acertos e falsos alarmes conjuntamente. Consideremos então a amostra representada na Tabela 2, cujos resultados representam um desempenho de $H=.50$ e $F=.15$, correspondendo a sensibilidade a um valor de $.35$. Quanto maior for a proporção de acertos, maior será a sensibilidade; no entanto, quanto mais falsos alarmes se verificarem, menor deve ser a sensibilidade. Qualquer medida de sensibilidade tem de considerar esta relação entre acertos e falsos alarmes, aumentando com o aumento do primeiro, e diminuindo com o aumento do segundo. Porém, a mera diferença $H - F$ não permite caracterizar a verdadeira competência do observador em distinguir os alvos dos distractores, na medida em que a verdadeira competência corresponde à tendência central de uma população de potenciais desempenhos e não é possível caracterizar uma população a partir de uma única observação. Para que seja possível essa caracterização, torna-se necessário inferir pressupostos sobre os parâmetros dessa distribuição. Quer dizer, necessitamos de inferir a forma como a informação sobre os estímulos se encontra disponível para o observador (neste caso, informação disponível em memória). Segundo a TDS, a informação disponível é representada por uma única dimensão que quantifica as evidências (normalmente definida como dimensão de familiaridade), através da qual os estímulos se encontram distribuídos. Um pressuposto assumido tradicionalmente pela TDS é que esta distribuição é normal (Green & Swets, 1966). É com base nestes dois pressupostos que surge a medida de sensibilidade mais popular, o d' (lê-se “d prime” ou “d linha”), Esta medida parte do pressuposto que a quantidade de evidências produzidas tanto pelos alvos como pelo ruído, se distribuem sob a forma de duas curvas normais, representadas através de um eixo X de familiaridade e um eixo Y de densidade de probabilidade, como se pode observar na Figura 2.

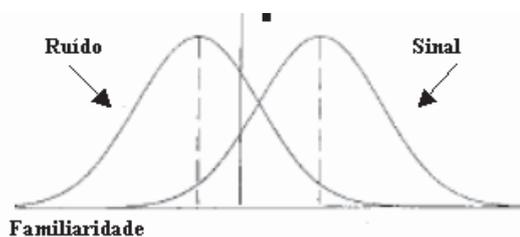


Figura 2. Distribuição normal dos alvos e do ruído, adaptado de Macmillan e Creelman (2005)

O cálculo de d' é bastante simples, já que é apenas a diferença entre o valor médio da distribuição dos alvos (μ_s) e da distribuição do ruído (μ_n) dividida pelo desvio padrão da distribuição do ruído (σ_n). Deste modo podemos expressar d' da seguinte forma:

$$d' = (\mu_s - \mu_n) / \sigma_n$$

O valor de d' permite então avaliar o nível de sensibilidade apresentado; enquanto o nível de acertos (hits) for superior ao de falsos alarmes, $d' > 0$; se d' for igual a 0, estamos perante uma sobreposição completa das duas distribuições, não existindo qualquer discriminação entre o alvo e o ruído e verificando-se um padrão de respostas aleatório. Quanto mais elevado for o valor de d' , maior será a sensibilidade.

No entanto este valor não pode ser calculado directamente, dado que apesar de se assumir a normalidade das distribuições, os seus parâmetros são desconhecidos, as distribuições apresentadas na Figura 2 são puramente teóricas. Porém, o facto de se assumir que as distribuições de sinal e ruído são normais, permite a utilização da distribuição normal reduzida (de média 0 e desvio padrão de 1) para a estimação de d' . Assim:

$$d' = zF - zH$$

O valor d' pode assim ser estimado manualmente, procurando-se nas tabelas da distribuição normal reduzida os valores z cujas probabilidades são equivalentes às proporções de acertos (hits) e de falsos alarmes; ou então através de soluções informáticas mais rápidas e cómodas, como as fórmulas de Excel, apresentadas no apêndice. Calculemos então o d' dos resultados do teste de reconhecimento:

$$H = .5 \text{ e } F = .15$$

$$d' = z(.5) - z(.15) = 1.036$$

Consideremos então que foi igualmente estudado o desempenho no reconhecimento de outras duas categorias de estímulos, a categoria M e N (apresentado na Tabela 3), e que queremos saber qual dos três métodos de estudo produz uma maior sensibilidade no reconhecimento de estímulos previamente estudados.

Tabela 3

Proporções das respostas num teste de reconhecimento das categorias M e N

| ESTÍMULO | RESPOSTA – AMOSTRA RECONHECIMENTO M | | |
|----------|-------------------------------------|-------|-------|
| | “SIM” | “NÃO” | TOTAL |
| ANTIGO | .85 | .15 | 1.00 |
| NOVO | .50 | .50 | 1.00 |
| ESTÍMULO | RESPOSTA – AMOSTRA RECONHECIMENTO N | | |
| | “SIM” | “NÃO” | TOTAL |
| ANTIGO | .85 | .15 | 1.00 |
| NOVO | .20 | .80 | 1.00 |

Se considerarmos apenas os acertos (hits) das amostras, iríamos concluir que tanto os métodos M e N são equivalentes, superiores aos do método K. No entanto, como já referimos, é necessário ter em conta não apenas os acertos (hits) mas também os falsos alarmes para que possamos avaliar a sensibilidade manifestada. No caso do método M, o aumento de acertos (hits) é acompanhado por um aumento de

falsos alarmes, o que levanta dúvidas em relação a uma sensibilidade superior à do método K. Calculemos então o d' desses resultados:

$$H=.85 \text{ e } F=.50$$

$$d'=z(.85)-z(.50)=1.036$$

Podemos então verificar que apesar do reconhecimento das categorias K e M levarem a valores diferentes ao nível dos falsos alarmes e dos acertos (hits), estas têm d' iguais, o que indica níveis de sensibilidade semelhantes. Comparemos então com o d' obtido com a categoria de estímulos N:

$$H=.85 \text{ e } F=.20$$

$$d'=z(.85)-z(.20)=1.87$$

Verificamos que o valor de d' obtido com a categoria N é superior ao encontrado na categoria M, indicando uma sensibilidade bastante elevada, apesar de ambas apresentarem uma proporção de acertos (hits) equivalente.

Se não tivesse sido utilizada a TDS na avaliação dos dados obtidos com estas três categorias, podia ter-se concluído erroneamente que tanto a categoria M como a N levam a performances semelhantes, superiores aos da categoria K. Na verdade, apenas com a categoria de estímulos N apresenta uma sensibilidade superior ($d'=1.87$), em comparação com as outras duas categorias que apresentam níveis de sensibilidade semelhantes ($d'=1.036$).

A medida de sensibilidade d' permite assim evitar erros de análise, possibilitando uma avaliação que vai além da superficialidade proporcionada pelos dados em bruto. O d' possui ainda a qualidade de não ser afectado pelos enviesamentos de respostas causados pelos critérios de especificidade, permitindo não apenas diferenciar os vários níveis de sensibilidade, mas também perceber os diferentes modos como o mesmo nível de sensibilidade pode ocorrer; o valor d' permite assim uma medida pura dos processos da resposta. Uma limitação importante no d' é o facto de não poder ser utilizada quando se verificam efeitos de tecto, isto porque quando o sujeito reconhece correctamente todos os estímulos antigos, o valor de d' é igual a ∞ (infinito). Este problema pode ser ultrapassado com a utilização de medidas de sensibilidade alternativas, como a alternativa “não-paramétrica” A' , ou então A_z , que tem como base a utilização de uma escala de avaliação. Ambas as medidas têm pressupostos e limitações específicas, pelo que serão descritas detalhadamente mais adiante.

Medidas de especificidade

Como ficou demonstrado anteriormente com as duas amostras, diferentes proporções de acertos (hits) e de falsos alarmes podem apresentar o mesmo nível de sensibilidade. As diferenças devem-se então a outro factor distinto, nomeadamente aos diferentes critérios de especificidade manifestados pelas amostras. A especificidade diz respeito ao limite de evidências utilizado para discriminar os vários estímulos. Este limite define assim um critério de reconhecimento, abaixo do qual os estímulos são rejeitados, e acima do qual são reconhecidos. Tal como se pode observar na Figura 3, o estabelecimento de um critério define não apenas a proporção de acertos (hits), mas também a prevalência de determinados tipos de erro. Quando o critério é mais conservador, como no caso de C1, há uma redução dos falsos alarmes, no entanto há um aumento de outro tipo de erro, a omissão. Pelo contrário, quando se estabelece um critério mais liberal como C2, há um aumento dos falsos alarmes acompanhado por uma diminuição das omissões. Dada a possibilidade deste tipo de enviesamentos, a medição da especificidade torna-se fundamental na compreensão dos padrões de resposta.

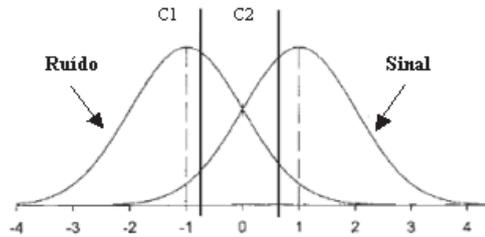


Figura 3. Diferentes estabelecimentos do critério de especificidade, adaptado de Macmillan & Creelman (2005)

Uma das maneiras de calcular o critério de especificidade é através da medida C , que utiliza a relação entre a intersecção das duas distribuições e as proporções de resposta manifestadas de forma a avaliar a natureza e magnitude do enviesamento. Quando não há qualquer enviesamento, $C=0$; no caso de um enviesamento causado por um critério mais liberal, $C<0$, ou então se estivermos perante um critério mais conservador, $C>0$:

$$C = .5 (zF + zH)$$

Como vimos anteriormente, as proporções de acertos (hits) não são indicativas da sensibilidade; as proporções de acertos (hits) nas amostras relativas às categorias K e M eram diferentes, no entanto traduziam o mesmo nível de sensibilidade. Esta situação leva a que se considere o estabelecimento de critérios de especificidade distintos para cada amostra, como pode ser demonstrado pelo cálculo de C :

| | | |
|---------------------------|---------------------------|---------------------------|
| Categoria K | Categoria M | Categoria N |
| $C = .5[z(.50) + z(.15)]$ | $C = .5[z(.85) + z(.50)]$ | $C = .5[z(.85) + z(.20)]$ |
| $C = .518$ | $C = -.518$ | $C = -.097$ |

Verifica-se então que as duas amostras apresentam enviesamentos de resposta tendencialmente opostos, resultantes do estabelecimento de critérios distintos. Na primeira amostra encontramos um critério ligeiramente conservador, enquanto que na segunda amostra verificamos exactamente o contrário, um critério de resposta mais liberal.

Uma derivação da medida C , é o C' , que pondera a especificidade em relação ao nível de sensibilidade manifestado. Esta medida decorre da necessidade de ter em conta d' , já que quanto maior este for, menor será o enviesamento provocado por um determinado valor C . Esta medida corrigida é calculada através da seguinte fórmula:

$$C' = C/d'$$

Outra das medidas utilizadas na avaliação dos critérios de especificidade é o β , ou coeficiente de verosimilhança, que se baseia na estimação a partir das densidades de probabilidade, tanto na distribuição de sinal (S), como na de ruído (N). O valor β é calculado da seguinte forma:

$$\beta = f(zH)/f(zF)$$

O critério neutro assume o valor 1; os enviesamentos liberais apresentam valores β menores que 1, enquanto que critérios mais conservadores tem valores β maiores do que 1. O cálculo de β pode no entanto ser representado de outro modo, utilizando-se logaritmos:

$$\ln(\beta) = \ln[z(H)z(F)]$$

No caso de $\ln(\beta)$, o critério é neutro para o valor 0, verificando-se um enviesamento liberal para valores negativos e um enviesamento mais conservador para valores positivos. Um enviesamento liberal significa que no ponto estabelecido, a probabilidade do sinal é inferior à probabilidade do ruído; pelo contrário, quando o critério estabelecido é conservador, a probabilidade do sinal é superior à do ruído. No caso em que o critério é neutro, as probabilidades de sinal e de ruído são equivalentes.

Tendo sido descritas três medidas de especificidade, C , C' e β , qual pode ser considerada a melhor ou mais adequada? Apesar de não existirem verdadeiras conclusões empíricas, vários autores (Lockhart & Murdock, 1970; Snodgrass & Corwin, 1988), consideram C a medida mais indicada, já que é aquela que é verdadeiramente independente da sensibilidade, permitindo uma avaliação pura dos processos decisórios e uma verdadeira distinção entre sensibilidade e enviesamento. Apesar da independência de C e da sua utilização em diversos modelos de memória como o SAM (Gillund & Shiffrin, 1984), o TODAM (Murdock, 1982) ou o MINERVA2 (Hintzman, 1988), modelos mais recentes como o ALT (Glanzer et al., 1993), o SLT (McClelland & Chappel, 1998) ou REM (Shiffrin & Steyvers, 1997) têm adoptado o β como medida de critério, dado que este permite explicar de uma forma mais simples alguns fenómenos mnésicos como [e.g., o efeito de espelho (Glanzer & Adams, 1985)].

A questão das diferentes medidas de critério reside no facto de que estas têm subjacentes processos decisórios completamente distintos, enquanto que C considera que apenas está em causa um processo baseado na familiaridade, β postula que os indivíduos calculam um coeficiente de verosimilhança de forma a aceitar ou rejeitar estímulos alvo (para uma discussão aprofundada deste tópico ver Morrel, Gaitan, e Wixted, 2002).

Curvas ROC

Como foi visto anteriormente, o mesmo nível de sensibilidade pode ser acompanhado por diferentes critérios de resposta, que alteram as proporções de acertos (hits) e falsos reconhecimentos dos participantes experimentais. Estas diferentes manifestações da mesma sensibilidade podem ser observadas nas curvas de isosensibilidade, normalmente designadas por curvas ROC (*Response Operating Characteristics*) que denotam as diferentes proporções de acertos (hits) e falsos alarmes contempladas num determinado valor d' , sendo a proporção de acertos (hits) representada pelo eixo X e a proporção de falsos alarmes pelo eixo Y, como se pode verificar na Figura 4. Estas curvas permitem assim compreender o modo como a variação do critério de especificidade afecta as respostas dadas e os erros contéudos, dado um valor de sensibilidade constante.

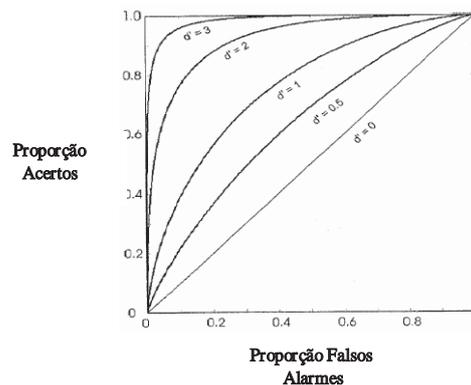


Figura 4. Exemplos de curvas de isosensibilidade, adaptado de Macmillan e Creelman (2005)

Outra vantagem das curvas ROC é a possibilidade de se mapear as respostas de participantes experimentais ao longo de diferentes níveis de especificidade, ou seja construir aquilo que se designa de curvas ROC observadas. A obtenção destes pontos necessita que os sujeitos alterem o seu critério de resposta, o que pode ser obtido directamente de duas formas (Macmillan & Creelman, 2005). A primeira forma é através da manipulação da proporção de palavras antigas e novas; quanto maior for a proporção de palavras antigas, mais liberal vai ser o critério, e vice-versa. A outra forma, alternativa à utilização das proporções, consiste na atribuição de uma pontuação ao desempenho dos sujeitos. Partindo do princípio que os participantes procuram maximizar os seus resultados, uma penalização nos falsos alarmes superior à bonificação obtida nos acertos, leva a que os sujeitos adoptem um critério mais conservador. Quando a bonificação é superior aos acertos, o critério deve ser mais liberal.

O problema destes dois métodos de manipulação do critério de resposta reside no facto de necessitarem que os sujeitos respondam a uma grande quantidade de estímulos em diversas condições de proporção ou pontuação. Esta limitação leva a que se opte por uma terceira alternativa – a utilização de níveis de confiança, que apesar de mais prática já que limita o número de condições, e consequentemente de respostas, implica a aceitação de um pressuposto adicional, nomeadamente que os níveis de confiança são equivalentes ao posicionamento do critério (ver Figura 5). A aceitação deste pressuposto permite inferir as proporções de acertos e falsos alarmes que os participantes manifestariam se estivessem a utilizar diferentes critérios de resposta.

Assim, a obtenção de curva ROC é feita através da utilização de um paradigma de avaliação, no qual se pede aos participantes experimentais que reconheçam estímulos, indicando numa escala o grau de confiança que têm das suas escolhas. Estas grelhas têm normalmente entre 6 a 10 graus de confiança, tendo como pólos um critério muito conservador (ex: claramente antigo) e um critério muito liberal (ex: claramente novo) (Murdock, 1965).

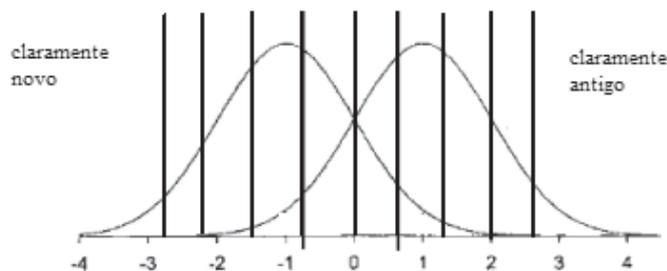


Figura 5. Segmentação das distribuições através de uma escala de avaliação, adaptado de Macmillan & Creelman (2005)

Os resultados obtidos são mapeados de forma cumulativa, iniciando-se pelo critério mais conservador até ao mais liberal. Os pontos encontrados permitem então estimar a curva correspondente a um determinado participante experimental ou amostra, e assim determinar d' .

Esta estimação é feita através da proporção de acertos (hits) e falsos alarmes encontrada no ponto de intersecção entre a diagonal menor (definida pela função $Y=1-X$) e a curva ROC observada, como se pode ver na Figura 6. Após a obtenção deste ponto na curva, o valor d' é calculado utilizando a fórmula já anteriormente utilizada. A utilização do ponto de intersecção deve-se ao facto deste denotar as proporções obtidas quando o critério de especificidade é nulo. Este método apresenta algumas limitações, no entanto permite obter uma estimativa de sensibilidade que não é influenciada pelos enviesamentos de resposta (Murdock, 1965).

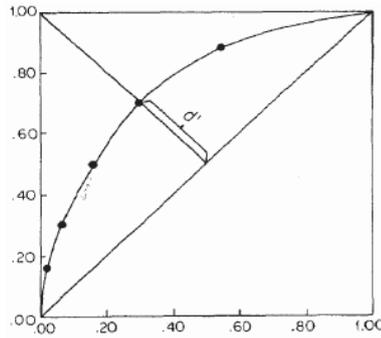


Figura 6. Exemplo de uma curva ROC observada, adaptado de Macmillan e Creelman (2005)

As curvas ROC trazem grandes vantagens ao estudo da memória de reconhecimento ao possibilitarem a análise de processos mnésicos distintos. Um desses casos é a distinção entre recordação explícita e a familiaridade dos estímulos, através da assimetria da curva e dos pontos de intersecção com os eixos (Rotello, Macmillan, & Van Tassel, 2000; Yonelinas, 1994).

Dado o pressuposto de normalidade das distribuições, é possível ainda transformar as proporções de acertos (hits) e falsos alarmes em coordenadas-z, obtendo-se desse modo uma zROC, definida pelas seguintes equações:

$$z(H) = -(C - \mu_s) / \sigma_s$$

$$z(F) = (C - \mu_n) / \sigma_n$$

Ao resolvermos estas equações em função do critério de especificidade (C), ficamos com:

$$z(H) = (\sigma_n / \sigma_s) z(F) + (\mu_s - \mu_n) / \sigma_s$$

Se considerarmos ainda que os desvios padrão das duas distribuições são equivalentes, a equação simplifica-se:

$$z(H) = z(F) + d'$$

A transformação da curva ROC leva a uma representação gráfica como a da Figura 7.

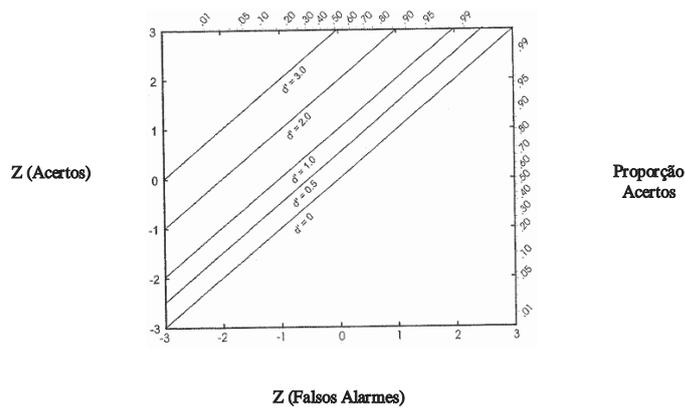


Figura 7. zROC para diferentes níveis de sensibilidade (d'), adaptado de Macmillan e Creelman (2005).

Apesar das ROC serem acessíveis na visualização dos pontos de intersecção, a sua transformação em zROC permite um melhor acesso a informação relativa às diferenças entre a distribuição de sinal e ruído, permitindo aferir a normalidade das distribuições através da linearidade da zROC (Green & Swets, 1966; no entanto esta método tem limitações, dado que existem distribuições não-normais que resultam numa linearidade zROC; ver Lockhart & Murdock, 1970), e determinar a diferença das médias das distribuições (d') através da distância vertical entre a linha z-ROC e a diagonal do acaso ($d'=0$).

Uma outra vantagem das zROC é o facto desta permitir a análise das variâncias de ambas as distribuições de uma forma mais precisa do que a que análise da simetria da curva ROC permite; sendo o declive da zROC a razão entre o desvio padrão do ruído e o desvio padrão do sinal medida pelo declive da linha zROC. Este último aspecto é de especial importância, já que importantes modelos de memória como o MINERVA2 (Hintzman, 1988), o modelo TODAM (Murdock, 1982) ou os modelos dualistas (Yonelinas, 1994) fazem previsões específicas relativamente à variância das distribuições, a sua igualdade e modificação consoante a força do sinal.

A medição da sensibilidade dos participantes baseada nas ROC apresenta vantagens consideráveis, no entanto a sua utilização implica que os participantes respondam com escalas de avaliação, o que pode comprometer a medição de outras variáveis como os tempos de resposta, já que não são equivalentes às verificadas com respostas dicotómicas sim/não. Em suma, a sua utilização deve ser ponderada, tendo em conta as vantagens e desvantagens.

Sensibilidade e comparação: Paradigmas de escolha forçada

A implementação da TDS em tarefas de reconhecimento foi até este momento abordada no contexto do paradigma de reconhecimento do tipo sim/não. Existem no entanto outras formas de avaliar a capacidade dos participantes experimentais reconhecerem informação previamente apresentada, como no caso do teste “*two-alternative-forced-choice*”, no qual em cada ensaio são apresentados simultaneamente dois estímulos. Ao contrário do paradigma de reconhecimento do tipo sim/não, que não distingue entre familiaridade inerente ao estímulo e a familiaridade causada pela sua apresentação prévia, o paradigma de escolha forçada possibilita a distinção destes através da apresentação simultânea de estímulos, manipulando-se a frequência tanto dos estímulos-sinal como dos estímulos-ruído. Neste caso, o participante experimental tem que escolher o sinal em cima ou em baixo. Esta é uma tarefa mais fácil que o reconhecimento do tipo sim/não já que os participantes experimentais não têm expectativas relativamente à apresentação do sinal em cima ou em baixo, pelo que esta é uma tarefa normalmente executada com o mínimo de enviesamento.

A TDS parte do pressuposto que os participantes experimentais estimam a familiaridade de cada um dos estímulos de forma independente, o que permite mapear os estímulos no espaço de decisão ortogonalmente, como se pode verificar na Figura 8.

Como se assume que o participante estima a familiaridade de cada item independentemente, obtêm-se as duas representações das distribuições de familiaridade presentes na Figura 8. Aplicando-se o teorema de Pitágoras a esta representação geométrica, mais precisamente ao triângulo formado pelas duas médias e pela origem, verifica-se que as duas médias das distribuições distam entre si $\sqrt{2}d'$, o que significa que o nível de sensibilidade dos indivíduos neste tipo de tarefa é superior em $\sqrt{2}$, em relação à tarefa de reconhecimento sim/não. Transpondo esta representação para o espaço unidimensional, obtêm-se uma distância semelhante, o que implica que a fórmula para calcular o d' seja compensada de modo a reflectir esta diferença na distância entre as médias das duas distribuições. Logo, a sensibilidade neste tipo de tarefa calcula-se então do seguinte modo:

$$d' = 1/\sqrt{2}[z(H) - z(F)]$$

Quanto à medida de especificidade, C , a sua fórmula de cálculo não se altera, já que não se costuma verificar grandes diferenças quanto ao enviesamento de resposta, entre esse tipo de tarefa e a tarefa de reconhecimento sim/não.

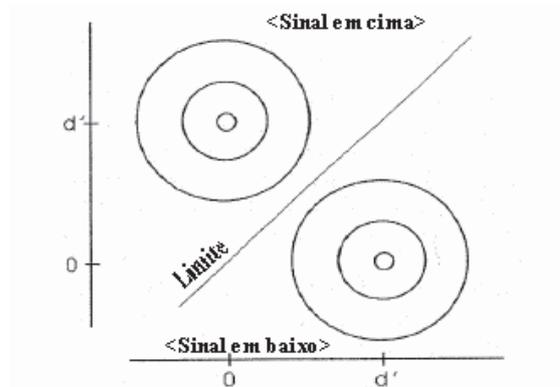


Figura 8. Representação bidimensional da tarefa de escolha forçada, adaptado de Macmillan e Creelman (2005)

Medidas alternativas: A proposta “não paramétrica”

Apesar da robustez das medidas paramétricas, alguns autores propõem a utilização de medidas que não tenham qualquer tipo de distribuição como pressuposto. Pollack e Norman (1964), Hodos (1970), Grier (1971) e Donaldson (1992) desenvolveram um conjunto de medidas não paramétricas, A' e B'' , e correspondentes fórmulas de computação, aproximadas de d' e C , respectivamente.

O princípio subjacente a estas medidas consiste na constatação de que a área debaixo da curva ROC numa tarefa de reconhecimento sim/não corresponde à proporção de acertos (hits) obtidos de forma não enviesada numa tarefa de escolha forçada. Tendo em conta que a análise do desempenho numa tarefa de escolha forçada não implica qualquer tipo de pressupostos, a medição da sensibilidade dos indivíduos através do cálculo da área debaixo da curva também não implicará qualquer tipo de pressupostos. Dentro desta lógica, e tendo em conta as propriedades da curva ROC, foi então proposto um processo de estimação que permitisse obter uma medida não-paramétrica da sensibilidade dos indivíduos, sendo apenas necessária uma única avaliação do desempenho, como o ponto (H,F) da Figura 9. Uma das características das ROC é o facto de terem um crescimento monótono acompanhado por um decréscimo do declive. Isto permite prever as áreas pelas quais a curva ROC pode passar ou não. Ao possuímos o ponto (H,F) , podemos delimitar duas áreas, $A1$ e $A2$, pelas quais a ROC tem obrigatoriamente que passar, e uma área I , e uma área S , que não são interceptadas pela curva ROC. Esta noção permitiu a Pollack e Norman (1964) criar uma fórmula que permite a estimação de A' :

$$A' = I + .5(A1 + A2)$$

A estimação do critério de especificidade (Hodos, 1970) foi igualmente desenvolvida:

$$B'' = (A1 - A2) / (A1 + A2 + 2S)$$

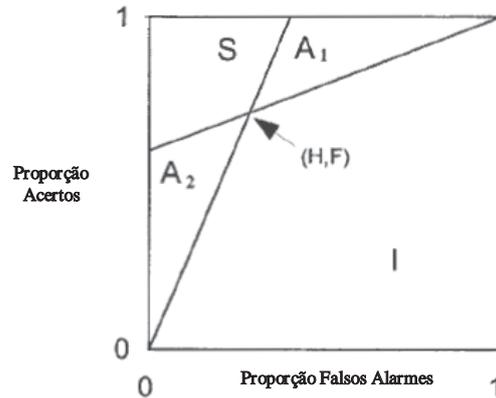


Figura 9. Áreas da curva ROC, adaptado de Macmillan e Creelman (2005)

As fórmulas de computação de A' e B'' foram igualmente definidas (Grier, 1971; Hodos, 1970):

Quando $H \leq F$,

$$A' = .5 + [(H-F)(1+H-F)]/[4H(1-F)]$$

$$B'' = [H(1-H) - F(1-F)]/[H(1-H) + F(1-F)]$$

Quando $H < F$,

$$A' = .5 + [(F-H)(1+F-H)]/[4F(1-H)]$$

$$B'' = [F(1-F) - H(1-H)]/[F(1-F) + H(1-H)]$$

A medida A' varia entre 0 e 1, sendo o valor .5 indicativo de reconhecimento aleatório; por outro lado, B'' varia entre -1 e 1, com 0 a denotar a ausência de enviesamento e valores positivos e negativos a indicar critérios liberais e conservadores, respectivamente. Estas medidas apresentam algumas vantagens, como a possibilidade de medir a sensibilidade de participantes experimentais com proporções de falsos alarmes ou de acertos (hits) de 0, algo que não é possível com d' ; para além disto apresentam poucas diferenças, estando altamente correlacionadas com as suas homólogas paramétricas. Contudo, existem autores que põem em causa a sua natureza não-paramétrica (Macmillan & Creelman, 2005). Diversos estudos têm demonstrado que de facto, A' assume muitos pressupostos acerca da distribuição subjacente ao desempenho dos participantes experimentais, assumindo diferentes tipos de distribuição consoante o nível de sensibilidade (Macmillan & Creelman, 1996).

Conclusão

Como vimos ao longo do artigo, a teoria de detecção de sinal não é apenas uma ferramenta útil na análise de dados, é uma ferramenta essencial para uma compreensão dos padrões de respostas dos participantes experimentais. A dissociação entre processos mnésicos e decisórios permite avaliar desempenhos e comparar diferentes teorias acerca da recuperação de informação em memória, possibilitando um debate aprofundado relativamente aos efeitos de diversas variáveis, como a comparação de diferentes propostas permite o desenvolvimento de novas teorias e modelos. Apesar de apenas ter sido abordada no contexto da literatura sobre memória de reconhecimento, a TDS pode ser aplicada a praticamente todos os

paradigmas que envolvam uma decisão dicotómica, sendo utilizado em investigações sobre memória implícita, e aplicado em tarefas como identificação perceptiva ou decisão lexical. Num estudo sobre memória implícita (Zeelenberg, Wagenmakers, & Raaijmakers, 2002), a TDS permitiu clarificar um debate nesta literatura, nomeadamente se a apresentação prévia de um estímulo aumentava a sensibilidade dos sujeitos em relação a estes em tarefas subsequentes, ou se estávamos perante um mero efeito de enviesamento de resposta. A TDS permitiu a estes autores afirmar que, para além de um enviesamento de resposta, existe um aumento da sensibilidade, ou seja, estão presentes dois efeitos. A utilização da TDS permite não apenas entender os contributos dos dois componentes no desempenho dos sujeitos, mas também rejeitar ideias antigas sobre os processos subjacentes a estes.

Para uma aprendizagem mais aprofundada da TDS em geral, recomendamos a consulta o livro de Macmillan e Creelman (2005). Em termos de artigos, Snodgrass e Corwin (1988) é uma referência clássica que faz uma introdução das diferentes medidas e apresenta uma aplicação da TDS num contexto clínico; o artigo de Verde, Macmillan e Rotello (2006) faz uma análise bastante interessante sobre a robustez das medidas; e a revisão de literatura feita por Wixted (2007) apresenta as vantagens da TDS enquanto modelo de memória.

Para finalizar, apresentamos uma lista que descreve os pontos principais sobre a TDS que devem ser extraídos do texto:

- a) A mera comparação dos acertos obtidos pelos sujeitos não é informativa, sendo necessário ter em conta os falsos alarmes.
- b) O desempenho dos sujeitos pode ser separado em duas componentes, uma componente mnésica que define a discriminabilidade entre os estímulos e o ruído, e uma componente de especificidade que estabelece um limite de evidências que define o critério de resposta.
- c) As representações gráficas ROC e zROC permitem não apenas a utilização de formas de medição alternativas mais robustas tendo em conta que assumem menos pressupostos do que as fórmulas, como o pressuposto de igualdade de variâncias, mas também permitem observar diferentes aspectos do desempenho como a intersecção dos eixos. Uma desvantagem das ROC é que implicam a utilização de escalas de avaliação, o que pode comprometer outros aspectos a ser estudados, como os tempos de resposta.
- d) Existem duas medidas de especificidade principais, C e β . Estas medidas assumem pressupostos fundamentalmente distintos, pelo que a sua utilização implica a aceitação dos mesmos.

Apêndice

Fórmulas de computação das medidas para folha de cálculo EXCEL

- 1) Introduzir as designações na primeira linha, uma designação por célula.
- 2) Introduzir as fórmulas correspondentes a cada designação nas células da linha dois.
- 3) Introduzir os valores nas células A-E, e o cálculo é automático

Designações (cada uma corresponde a uma célula da linha 1)

- A. Participante
- B. #acertos (hits)
- C. # omissões
- D. #falsos alarmes
- E. #rejeições correctas
- F. H

(cont. →)

(← cont.)

G. F
 H. z(H)
 I. z(F)
 J. d'
 K. c
 L. c'
 M. â
 N. lnâ
 O. A'
 P. B''
 Q. Pr
 R. Br

Fórmulas (correspondem às células da linha 2. substituir os números pela linha correspondente)

SE[C2>0;B2/(B2+C2);(B2-0,5)/(B2+C2)]

SE[D2>0;D2/(D2+E2);0,5/(D2+E2)]

INV.NORMP(F2)

INV.NORMP(G2)

H2-I2

0,5*(H2+I2)

L2/K2

EXP(K2*L2)

LN(N2)

SE[G2>F2;(0,5-(G2-F2)*(1+G2-F2)/(4*G2*(1-F2)))+(0,5+(F2-G2)*(1+F2-G2)/(4*F2*(1-G2)))]

SE[G2>F2;((F2*(1-F2))-(G2*(1-G2)))/((F2*(1-F2))+(G2*(1-G2)))+((G2*(1-G2))-(F2*(1-F2)))/((G2*(1-G2))+(F2*(1-F2)))]

F2-G2

G2/[1-(F2-G2)]

Referências

- Banks, W. P. (1970). Signal detection theory and human memory. *Psychological Bulletin*, 74, 81-99.
- Donaldson, W. (1992). Measuring recognition memory. *Journal of Experimental Psychology: General*, 121, 275-277.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, 91, 1-67.
- Glanzer, M., & Adams, J. K. (1985). The mirror effect in recognition memory. *Memory & Cognition*, 16, 5-16.
- Glanzer, M., Adams, J. K., Iverson, G. J., & Kim, K. (1993). The regularities of recognition memory. *Psychological Review*, 100, 546-567.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, 75, 424-429.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, 95, 528-551.
- Hodos, W. (1970). Nonparametric index of response bias for use in detection and recognition experiments. *Psychological Bulletin*, 74, 351-354.

- Lockhart, R. S., & Murdock, B. B., Jr. (1970). Memory and the theory of signal detection. *Psychological Bulletin*, 74, 100-109.
- Macmillan, N. A., & Creelman, C. D. (1996). Triangles in ROC space: History and theory of “non-parametric” measures of sensitivity and response bias. *Psychonomic Bulletin & Review*, 3, 164-170.
- Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory: A user's guide (2nd ed.)*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- McClelland, J. L., & Chappell, M. (1998). Familiarity breeds differentiation: A subjective-likelihood approach to the effects of experience in recognition memory. *Psychological Review*, 105, 724-760.
- Morrell, H. E. R., Gaitan, C., & Wixted, J. T. (2002). On the nature of the decision axis in signal-detection-based models of recognition memory. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 28, 1095-1110.
- Murdock, B. B. (1965). Signal detection theory and short-term memory. *Journal of Experimental Psychology*, 70, 443-447.
- Murdock, B. B. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review*, 89, 609-626.
- Pollack, I., & Norman, D.A. (1964). A non-parametric analysis of recognition experiments. *Psychonomic Science*, 1, 125-126.
- Rotello, C. M., Macmillan, N. A., & Van Tassel, G. (2000). Recall-to reject in recognition: Evidence from ROC curves. *Journal of Memory and Language*, 43, 67-88.
- Shiffrin, R. M., & Steyvers, M. (1997). A model for recognition memory: REM-retrieving effectively from memory. *Psychonomic Bulletin & Review*, 4, 145-166.
- Snodgrass, J. C., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117, 34-50.
- Verde, M. F., Macmillan, N. A., & Rotello, C. M. (2006). Measures of sensitivity based on a single hit rate and false-alarm rate: The accuracy, precision, and robustness of d' , Az , and A' . *Perception & Psychophysics*, 68, 643-654.
- Wixted, J. (2007) Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114, 152-176.
- Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 20, 1341-1354.
- Zeelenberg, R., Wagenmakers, E.-J., & Raaijmakers, J. G. W. (2002). Repetition priming in implicit memory tasks: Prior study causes enhanced discriminability, not only bias. *Journal of Experimental Psychology: General*, 131, 38-47.