



**Evidence for the contribution of a threshold retrieval process to semantic memory**

Journal:	<i>Quarterly Journal of Experimental Psychology</i>
Manuscript ID	QJE-STD 15-229.R1
Manuscript Type:	Standard Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Kempnich, Maria; University of St Andrews, School of Psychology & Neuroscience Urquhart, Josephine; University of St Andrews, School of Psychology & Neuroscience O'Connor, Akira; University of St Andrews, School of Psychology & Neuroscience Moulin, Chris; University of Bourgogne, , France, Laboratoire d'Etude de l'Apprentissage et du Développement
Keywords:	recollection, memory, semantic memory, word recognition

SCHOLARONE™  
Manuscripts

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

**Title**

Evidence for the contribution of a threshold retrieval process to semantic memory

**Authors**

Maria Kempnich\*, School of Psychology & Neuroscience, University of St. Andrews,  
Scotland

Josephine A. Urquhart\*, School of Psychology & Neuroscience, University of St. Andrews,  
Scotland

Akira R. O'Connor, School of Psychology & Neuroscience, University of St. Andrews,  
Scotland

Chris J.A. Moulin, LPNC CNRS 5105, Université Grenoble Alpes, France

\* Maria Kempnich and Josephine Urquhart made equally important contributions to this  
article and should be recognised as joint first authors.

**Corresponding Author**

Akira O'Connor, School of Psychology, University of St Andrews, St Mary's College, South  
Street, St Andrews, Fife, KY16 9JP, Scotland, UK. Email [aro2@st-andrews.ac.uk](mailto:aro2@st-andrews.ac.uk)

**Acknowledgement**

Maria Kempnich was supported by the University of St Andrews University Research  
Internship Placement Scheme and a Walker Trust Travel Scholarship.

Josephine Urquhart was supported by the Economic and Social Research Council 1+3  
Scheme.

**Disclosure Statement**

All authors declare that we have no financial interest and will receive no financial benefit  
from the direct applications of the research presented in this manuscript.

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

**ABSTRACT**

It is widely held that episodic retrieval can recruit two processes, a threshold context retrieval process (*recollection*) and a continuous signal strength process (*familiarity*). Conversely, and in spite of its importance for everyday memory, the processes recruited during semantic retrieval are less well specified. We developed a semantic task analogous to single-item episodic recognition to interrogate semantic recognition receiver operating characteristics (ROCs) for a marker of a threshold retrieval process. We then fit observed ROC points to three signal detection models: two models typically used in episodic recognition (unequal variance and dual process signal detection models) and a novel dual process *recollect-to-reject* (DP-RR) signal detection model that allows a threshold recollection process to aid both target identification and lure rejection. Given the nature of most semantic questions used here, we anticipated the DP-RR model would best fit the data obtained from our semantic task. In Experiment 1 (506 participants), we found evidence for a threshold retrieval process in semantic memory, with overall best fits to the DP-RR model. In Experiment 2 (316 participants), we found within-subjects estimates of episodic and semantic threshold retrieval to be uncorrelated, suggesting the relationship between the analogous memory processes is not straightforward. Our findings add weight to the proposal that semantic and episodic memory are served by similar dual process retrieval systems, though the relationship between the two threshold processes needs to be more fully elucidated.

**Keywords:** recollection, memory, semantic memory, word recognition

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 Cognitive and neuroscientific evidence indicates that two distinct retrieval processes play a  
4 role in episodic recognition: familiarity and recollection (for reviews, see Vilberg & Rugg,  
5 2008, and Yonelinas, 2002). Familiarity involves an interrogation of memory strength, a  
6 continuum along which *old* and *new* items (*targets* and *lures* respectively) elicit a simple  
7 memorial response. An episode recognised on the basis of familiarity alone will not yield any  
8 contextual evidence that it is old, merely the awareness that it relates to an experience from  
9 the past. Recollection on the other hand, provides context in the form of evidence retrieved  
10 to support the recognition judgement e.g. "I am certain that this is the road I travelled with  
11 my parents when we were on holiday as I remember stopping at that picnic area."

12  
13 Recollection is typically conceptualised as a 'threshold' process, not on a graded continuum  
14 like familiarity, but with contextual information either being successfully retrieved or not.  
15  
16

17  
18  
19  
20  
21 Recollection that occurs during recognition of a presented item typically indicates  
22 unambiguously that the item is a target (Yonelinas, 2001; though there are exceptions, e.g.  
23 Payne, Elie, Blackwell & Neuschatz, 1996). This is because such target-driven recollection  
24 brings to mind associative details encoded during study – details which it would not be  
25 possible to recollect were the item a lure. The absence of recollection during such trials  
26 however, does not necessarily signify that the currently judged item is a lure - just because  
27 one fails to retrieve supporting context, that does not rule out the possibility that an item is a  
28 target for which context was not encoded or is currently unavailable. Because episodic  
29 recognition tests tend to use cues that do not lead to identification of a single item, it is  
30 generally not possible to use recollection to aid rejection of a lure. For example, just because  
31 one recalls details associated with 'bush', that does not indicate that one can therefore  
32 not have encountered the currently presented 'lens'. Hence, in typical episodic recognition  
33 tasks, the threshold process of recollection aids only the identification of targets, and in so  
34 doing lends accompanying recognition judgements a high degree of subjective confidence.  
35  
36 The continuous process of familiarity on the other hand, can aid the identification of *both*  
37 lures and targets, and does so across the full range of subjective confidence.  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3  
4  
5 There remains some debate about the contribution of recollection to episodic retrieval (see  
6  
7 Dunn, 2004). For example Shiffrin and Steyvers' (1997) REM computational model  
8  
9 advances a number of previous models to predict ROC shapes without requiring this  
10  
11 process. Meanwhile, Mickes, Wixted & Wais (2009) posit that recollection is not an 'all-or-  
12  
13 none' high-threshold process, but that it is graded in much the way that we have described  
14  
15 familiarity. Nevertheless, the dual process nature of episodic retrieval has been largely  
16  
17 supported by results from a range of cognitive paradigms, including the 'remember-know'  
18  
19 (RK) procedure, process dissociation (PD), and receiver-operating characteristics (ROCs;  
20  
21 for a review see Yonelinas, 2002). Whilst the RK procedure has its origins in Tulving's  
22  
23 (1985) fractionation of declarative memory into episodic content which can be  
24  
25 "(R)emembered", and semantic content which is "(K)nown", it is now widely used to  
26  
27 differentiate remembered recollective content from known familiar content (e.g. Gardiner,  
28  
29 1988; Koen & Yonelinas, 2016). It requires participants to introspect on the subjective quality  
30  
31 of their episodic memory and approaches this from the position that the dual process  
32  
33 account is correct (participants are required to understand the qualitative distinction between  
34  
35 R and K responses before they can make them). PD does not require participants to  
36  
37 introspect on their retrieval processes, but asks them to make judgements in which they  
38  
39 either *include* all items they recognise or *exclude* some, which it is assumed can only be  
40  
41 done using the context brought to mind via recollection (Jacoby, 1991). Similarly, ROC  
42  
43 analyses make even fewer theoretical assumptions at the data collection stage and so are  
44  
45 less predisposed to favour dual process accounts when adjudicating between competing  
46  
47 accounts of episodic recognition memory. As outlined in detail below, ROC analyses fit  
48  
49 theoretically-specified curves to observed data on the assumption that the most accurate  
50  
51 account of the memory processes will provide the best fits overall (see Arndt & Reder, 2002,  
52  
53 for a lucid account of this method; also Yonelinas, 1994).  
54  
55

56  
57  
58 Semantic memory recognition processes underpin a variety of real world tasks and  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 behaviours: e.g. Does the word 'tortuous' mean twisty? Is this calf pain a result of plantar  
4 fasciitis? Is the brown wire in my plug live? Yet given the vast range of episodic recognition  
5 measures which could be repurposed for its investigation, it is surprising how little  
6 methodological crossover there is from episodic to semantic memory research. Interestingly,  
7 a number of noteworthy studies drawing on methodological crossover from episodic to  
8 semantic domains have used famous names as stimuli.

9  
10  
11  
12  
13  
14  
15  
16  
17 Bowles, Harlow, Meeking and Köhler (2012) conducted two experiments in which they used  
18 ROC fits to investigate the retrieval processes characterising the identification of famous  
19 compared to fictional names. In comparing models with different combinations of  
20 parameters, they found that ROC differences across their experiments could be best  
21 accounted for by model that varied two parameters, the displacement of the target  
22 distribution and either the standard deviation of the target distribution or the displacement of  
23 an additional lure distributing representing a higher baseline familiarity for famous names  
24 (target distribution displacement and standard deviation are explained in greater depth  
25 below). Bowles and colleagues found no benefit, beyond what would be expected from  
26 including an additional parameter, of including a parameter modelling recollection.

27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37 Waidergoren, Segalowicz and Gilboa (2012) used a different methodology, a semantic PD  
38 procedure, whereby participants classified people as famous or nonfamous and dead or  
39 alive. Contrary to Bowles and colleagues' (2012) ROC findings, Waidergoren et al found that  
40 their PD data could be best accounted for through the contribution of two processes: a signal  
41 strength-based process and a more effortful, contextually rich process—analogueous to  
42 episodic familiarity and recollection respectively. Data from this PD procedure suggests that  
43 dual processes may contribute to systems at play in semantic retrieval, and indeed, this is  
44 implicit in later work by Bowles and Köhler (2014). In a semantic task assessing recognition  
45 of famous names, Bowles and Köhler identified names which participants termed 'familiar-  
46 only'—names for which participants could not retrieve any additional context. The authors  
47 found that, although the 'familiar-only' status suggests no additional stimulus information  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 should be recovered, participants were able to perform above chance when given a forced-  
4  
5 choice semantic task (identifying the occupation associated with the name). Importantly to  
6  
7 their approach, in using terms aligned to episodic recognition research, Bowles and Köhler  
8  
9 suggest that a more complete semantic retrieval would involve a process episodic memory  
10  
11 researchers would deem analogous to recollection.

12  
13  
14  
15 If a dual process account generalised to all semantic memory retrieval, convergent evidence  
16  
17 for it could also be evident using ROC analyses. To test memory models against ROC data,  
18  
19 theoretically-determined ROC curves are fit to points plotted from old/new and confidence  
20  
21 responses to recognition tests, representing the cumulative proportion of targets ( $y$ -axis) and  
22  
23 lures ( $x$ -axis) given judgements from “high confidence old” (bottom-left of the ROC) through  
24  
25 to “high confidence new” (top-right). Episodic ROCs based on single-item recognition tasks  
26  
27 typically suggest a  $y$ -axis intercept above zero and an asymmetry about the diagonal joining  
28  
29 points  $[0, 1]$  and  $[1, 0]$  ( $y = 1 - x$ ; see Figure 1B). Episodic dual process theories  
30  
31 accommodate this ROC shape by incorporating the selective advantage to target  
32  
33 identification that threshold recollection offers. This takes the form of the parameter,  $R$ ,  
34  
35 specifying the degree to which the  $y$ -axis intercept falls above 0 i.e. the proportion of targets  
36  
37 which recollection unambiguously identifies as “old”, at no cost to the misidentification of  
38  
39 lures. As would be expected of parameter estimates from this method, convergence with RK  
40  
41 and PD methods is established (e.g. Koen & Yonelinas, 2016; Yonelinas, Dobbins,  
42  
43 Szymanski, Dhaliwal & King, 1996; Yonelinas, Kroll, Dobbins, Lazzara & Knight, 1998).  
44  
45 Extrapolating from these episodic memory findings to the current subject of investigation,  
46  
47 and based on Waidergoren and colleagues’ (2012) PD-derived evidence for a semantic  
48  
49 recollection-like process, it should be possible to find a corresponding marker of a threshold  
50  
51 retrieval process using ROCs constructed from on semantic data.

52  
53  
54  
55  
56 Two established models to which we fit ROCs obtained from semantic recognition are  
57  
58 standard models within the episodic recognition memory literature: the unequal variance  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 (UEV) signal detection model; and the previously described dual process (DP) signal  
4  
5 detection model. The UEV model is a single process model (i.e. does not assume the  
6  
7 existence of a recollection process) which accommodates typical episodic recognition ROCs  
8  
9 by allowing the target distribution dispersion to vary relative to the lure distribution dispersion  
10  
11 (Heathcote, 2003; Ratcliff, Sheu & Gronlund, 1992). Thus, it formalises the displacement of  
12  
13 the target distribution ( $d'$ ), and its standard deviation ( $\sigma$ ; see Figure 1A). When  $\sigma$  exceeds  
14  
15 the lure distribution standard deviation, the proportion of hits generally increases relative to  
16  
17 the proportion of false alarms. Under these circumstances, the ROC will extend from [0,0] to  
18  
19 [1,1], but display an asymmetry about the diagonal ( $y = 1 - x$ ; i.e. the dotted line that extends  
20  
21 from [0,1] to [1,0] in Figure 1A). When represented in z-space, the UEV model yields lines of  
22  
23 varying intercept and gradient. When fitting an UEV model ROC to data based on 6 degrees  
24  
25 of confidence, 7 parameters can vary: 5 criterion parameters (one for each confidence  
26  
27 boundary),  $d'$  and  $\sigma$ .

28  
29  
30  
31 The DP model accommodates typical episodic recognition ROCs by supplementing a target  
32  
33 distribution with the same dispersion as the lure distribution, with the recollection process  
34  
35 described previously (Yonelinas, 1994). It therefore formalises a  $d'$  parameter and a  
36  
37 recollection parameter ( $R$ ), with  $R$  representing the proportion of targets which are not  
38  
39 subject to the signal detection process, but are identified as high confidence hits via a  
40  
41 second, independent, threshold process (see Figure 1B). In a DP model ROC, the curve will  
42  
43 extend from [0, $R$ ], usually in the lower left quadrant, to [1,1] in the upper right quadrant,  
44  
45 displaying asymmetry about  $y = 1 - x$ . When represented in z-space, the DP model yields  
46  
47 lines with a curve in the lower aspect that corresponds to the magnitude of the  $R$  parameter.  
48  
49 Under the same conditions as the UEV model, the DP model also has 7 free parameters: 5  
50  
51 criterion parameters,  $d'$  and  $R$ .

52  
53  
54  
55  
56 The above models have different ways of accommodating ROCs asymmetry associated with  
57  
58 a higher proportion of high confidence target identifications than lure rejections.  
59  
60



## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 Nonetheless, they could not accommodate situation where ROC asymmetry was negated by  
4  
5 the additional contribution of recollection to the rejection of lures. Such a situation would be  
6  
7 unlikely to occur in standard episodic recognition tests (which is undoubtedly why an ROC  
8  
9 model that allows for this possibility has not gained traction) but is well described in  
10  
11 alternative tasks in which recollecting the target indicated by the cue can aid the rejection of  
12  
13 the lure with which the cue is presented (known as a recall-to-reject strategy; e.g. Rotello,  
14  
15 Macmillan & Van Tassel, 2000). For example, in episodic recognition, recollecting  
16  
17 information associated with studying 'glass' has no bearing on whether 'cadence' is a target  
18  
19 or a lure. This is because the question used to interrogate memory (implicitly, "This word  
20  
21 appeared on the previous list, true or false?") refers to as many probes as were studied. A  
22  
23 question with only one answer, on the other hand, affords the recollection aiding lure  
24  
25 rejection. Whilst this could be enacted in an episodic task, the most straightforward example  
26  
27 of this sort of question that is encountered on a day-to-day basis is in the domain of  
28  
29 semantic memory. The cue, "Who was the first person to run a mile in under 4 minutes?",  
30  
31 could be used to reject the lure "Eric Liddell" with high confidence, if the respondent recalled  
32  
33 the target answer Roger Bannister and recollected details consistent with this alternative  
34  
35 response. Thus, assuming that recollection can aid semantic memory retrieval, semantic  
36  
37 recognition tests of the form described would warrant a modified instantiation of the dual  
38  
39 process model for resulting ROC shapes.  
40  
41  
42

43  
44 The novel third model to which we fit ROCs obtained from semantic recognition was  
45  
46 therefore a dual process model modified to incorporate recollection-aided lure rejection. The  
47  
48 dual process recollect-to-reject (DP-RR) signal detection model, uses the same parameters  
49  
50 as the classic DP model but the  $R$  parameter is operationalised slightly differently (see  
51  
52 Formulae 1, adapted from Koen, Barrett, Harlow & Yonelinas, 2014). DP-RR ROCs show an  
53  
54  $x$ -axis intercept of  $1-R$  at  $y = 1$  (see Figure 1C). Thus, we propose that semantic retrieval  
55  
56 ROCs will, under the demands of this particular task, display evidence of a recollection-like  
57  
58 process that both parallels *and* deviates from that typically observed during episodic  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 recognition. In model parameter terms,  $R$  represents *both* the proportion of targets identified  
4 with high confidence at no cost to the misidentification of lures *and* the proportion of lures  
5 rejected with high confidence and at no cost to the misidentification of targets. It is important  
6 to note that this yoking of the two identification advantages to one parameter ensures that  
7 lure rejection *must* benefit from  $R$  as much as target identification does. The benefit of  
8 accommodating lure rejection simultaneously constrains the model, only allowing it to fit  
9 symmetric ROCs. Thus, in the DP-RR model ROC, the curve will extend from  $[0, R]$  to  $[1-R, 1]$   
10 with symmetry about  $y = 1 - x$ . When represented in z-space, the DP-RR model yields lines  
11 with curves in *both* the lower and upper aspects, which correspond to the magnitude of the  
12 R parameter. The number of free parameters, 7, remains identical to the classic DP model.

$$(Bi|target) = R + (1 - R)\Phi(c_k - d', 1)$$

$$(Bi|lure) = (1 - R)\Phi(c_k, 1)$$

23  
24  
25  
26  
27  
28  
29  
30 *Note:*  $Bi$  is the  $i$ th rating bin,  $(Bi|target)$  is the proportion of responses in each bin for  
31 targets,  $(Bi|lure)$  is the proportion of responses in each bin for lures,  $R$  is the  
32 recollection parameter,  $\Phi$  is the Gaussian cumulative distribution function,  $c_k$  is the  
33  $k$ th criterion placement parameter (where the maximum value of  $k$  is equal to  $i - 1$ ),  
34 and  $d'$  is the familiarity parameter representing the distance between target and lure  
35 distributions.  
36  
37  
38  
39  
40  
41  
42

[Formulae 1]

43  
44  
45  
46  
47 All three models can fully accommodate equal variance single process signal detection  
48 ROCs, which are symmetrical about  $y = 1 - x$  and extend from the origin to  $[1, 1]$ , by setting  $\sigma$   
49 = 1 or  $R = 0$ .  
50  
51  
52

53  
54  
55 (Figure 1 about here)  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 The possibility of analogous threshold contributions to semantic and episodic retrieval raises  
4 the question of how these processes relate to one another. A core quality of some definitions  
5 of episodic recollection is its capacity to stimulate mental time-travel – the re-living of past  
6 experiences (Suddendorf & Corballis, 1997, Tulving, 1985). In contrast, the process of  
7 transferring episodic content to semantic memory has been characterised as the stripping  
8 away of this episodic specificity (e.g. going from remembering your teacher telling you that  
9 the English were defeated by the Scots at the Battle of Bannockburn, to knowing that this is  
10 the case without being able to bring to mind how you know; Conway, Gardiner, Perfect,  
11 Anderson & Cohen, 1997; also Klein, 2013). Thus, there are inconsistencies in at least the  
12 subjective qualities of recollection and the putative semantic recollection-like process which  
13 indicate a range of possible points of divergence. It may be that the threshold retrieval  
14 process is unitary, acting on different memory stores, which recover qualitatively different  
15 retrieval content. Or it may be that these two processes are entirely dissociable – supported  
16 by different neural structures and comparable only in their eventual effects on subjective  
17 confidence ratings (c.f. Waidergoren et al, 2012). An obvious exploratory step is to use a  
18 within-subjects approach which matches the methods used to estimate the two retrieval  
19 processes as closely as possible, to establish whether there is any correspondence in the  
20 degree to which semantic and episodic memories are supported by them.  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40

41 Thus, two questions motivated our research. First, we were interested in whether the  
42 semantic recollection-like process reported by Waidergoren et al (2012) would be evident as  
43 a threshold process in ROC curves fit to semantic memory retrieval. Second, we wanted to  
44 establish whether any semantic threshold process would show a clear relationship with  
45 episodic recollection parameters, recovered in an analogous manner. Our starting point was  
46 to collect confidence judgements obtained during a single-item semantic recognition task  
47 and use these to construct ROCs. To this end, in Experiment 1 we developed and employed  
48 a recognition task to assess the semantic retrieval processes recruited to answer general  
49 knowledge questions. Our hypothesis was that a recollection-like threshold process would  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 be evident in semantic ROCs. In Experiment 2, we capitalised on the identification of a  
4 semantic threshold retrieval process by running within-subjects comparisons of semantic  
5 and episodic recognition memory.  
6  
7  
8  
9

**EXPERIMENT 1****METHODS****Stimuli**

10  
11  
12  
13  
14  
15  
16  
17 With the aim of constructing ROCs from semantic retrieval data, we developed a test of  
18 recognition memory for semantic materials which has the same constraints and procedures  
19 as a classic episodic memory recognition task. In episodic tasks there is usually a binary  
20 decision, old or new, based on 6 degrees of confidence, ranging from certain old to certain  
21 new. Although one tends not to explicitly probe a question in episodic memory tasks, a  
22 stimulus word is shown and each trial is essentially of the form: "This word appeared on the  
23 previous list, true or false?" (e.g. Donaldson, 1996; c.f. Mill & O'Connor, 2014). Thus, to  
24 parallel this design, our semantic questions were a true/false format with the same 6  
25 confidence points: sure true to sure false. Standard procedure is to probe participants'  
26 memories with equal numbers of targets and lures, so we generated a set of questions and  
27 provided a true answer for half of the trials and a false answer for the other half.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40

41 Stimuli were 500 general knowledge questions generated by the experimenters, covering a  
42 wide subject range (see Supplementary Materials A for a complete list). In generating  
43 stimuli, we prioritised open-ended questions which allowed the use of a recall strategy as a  
44 viable way in which the assessed knowledge could be retrieved e.g. "Which country is  
45 associated with haggis?" For each question, we generated one incorrect answer (lure;  
46 "Wales") as well as the correct answer (target; "Scotland"). During each memory test, a  
47 random sample of 60 questions was selected for each participant. Questions were  
48 presented alongside either a target answer (30 trials) or a lure answer (30 trials). The  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 experiment was programmed using JavaScript and presented to participants via their  
4  
5 internet browsers.  
6  
7

**Participants**

8  
9  
10  
11 506 participants (259 men, 239 women, 8 did not disclose sex) completed the online  
12  
13 experiment. Participants were recruited via links to the experiment posted on the laboratory  
14  
15 website and social networking sites (e.g. Facebook and Twitter). Informed consent was  
16  
17 obtained in accordance with the University Teaching and Research Ethics Committee at the  
18  
19 University of St Andrews. Of the 500 who disclosed their age, mean reported age was 26.9  
20  
21 years (SD = 10.1). We found no differences in sensitivity (equal variance [EV]  $d'$ ) between  
22  
23 the 414 participants who reported that English was their first language ( $M = 1.06$ , 95% CIs =  
24  
25 [1.02, 1.09]) and the 83 participants who reported that it was not ( $M = 1.03$  [0.95, 1.10])<sup>1</sup>.  
26  
27 Similarly we found no differences in bias (EV  $c$ ) over the two groups ( $M = -.097$  [-.120, -.075]  
28  
29 and  $M = -.104$  [-.154, -.054] respectively). We therefore included all participants in  
30  
31 subsequent analyses, collapsing across reported first language.  
32  
33

**Design and Procedure**

34  
35  
36  
37 Onscreen instructions were followed by a single test block. A 0.5 s fixation cross preceded  
38  
39 each self-paced test trial. During each test trial, the fixation was replaced by a question  
40  
41 presented in 250% font size at the top of the screen, above an answer (either a target or a  
42  
43 lure) presented in 200% font size (relative font sizes were used to allow for scalable text  
44  
45 sizes according to display size/resolution). Below the answer were six response buttons,  
46

---

47  
48 <sup>1</sup> All preliminary signal detection parameter analyses were conducted on sensitivity ( $d'$ ) and  
49  
50 bias ( $c$ ) parameters derived from the assumptions of an equal variance signal detection  
51  
52 model (Green & Swets, 1966; Macmillan & Creelman, 2005). In calculating these  
53  
54 parameters, we applied a correction for errorless responding as detailed in Snodgrass and  
55  
56 Corwin (1988).  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 from left to right: false sure 1, false probably 2, false guess 3, true guess 4, true probably 5,  
4 true sure 6. If using a keyboard, participants responded by pressing the corresponding  
5 number key. If using a touchscreen device, participants responded by pressing the  
6 appropriate button on the screen. After a response had been rendered, the fixation period  
7 before the next trial was initiated. Across the entire test, 30 target trials and 30 lure trials  
8 were presented with the allocation of question target/lure status being randomised (e.g. the  
9 question “The Roman numeral D represents what number?” was equally likely to be  
10 presented alongside the target “500” as the lure “50”). This ensured that the format of the  
11 test would not lend itself to the adoption of a bias towards either ‘true’ or ‘false’ responding.  
12 Once the 60th trial had been responded to, participants were taken to a debrief page, which  
13 also provided feedback on their accuracy and confidence.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26

**RESULTS****Item Selection**

27  
28  
29  
30  
31 To acknowledge that participants could look up question answers in a separate tab/window,  
32 we excluded responses with response times exceeding 10 s from all subsequent analyses.  
33 We excluded 5.26 [4.70, 5.82] responses per participants, which reduced mean response  
34 times from 5.76 s [5.57 s, 5.94 s] to 4.91 s [4.82, 5.00].  
35  
36  
37  
38  
39  
40

**Model Fits**

41  
42 To fit each participant to each model, we used standard maximum likelihood estimation  
43 (MLE) procedures operationalised in the ROC Toolbox for Matlab (Koen et al, 2014;  
44 [https://github.com/jdkoen/roc\\_toolbox/releases](https://github.com/jdkoen/roc_toolbox/releases); we also compared model fits between the  
45 DP-RR model and the equal variance signal detection model in Supplementary Materials B).  
46 This yielded successful fits to all models in 502 participants (99.2%) of the sample. In the  
47 additional question-selective analyses detailed below, successful fits were obtained to all  
48 models in 500 participants (98.8%).  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 Figure 2A plots the mean observed data against predicted ROCs derived from the mean  
4 parameters output by the MLE fitting procedures. The data points are approximately  
5 symmetrical about  $y = 1 - x$  and the proximity to the boundaries in the lower left and upper  
6 right of the plot are not consistent with an ROC that tends to [0,0] and [1,1].  
7  
8  
9

10  
11  
12  
13 For a number of items, participants did not use as much of the response scale as for others.  
14 Specifically, these items appeared to be very easy (answered with high confidence as both  
15 targets and lures), with the potential that they were biasing the shape of the ROC towards  
16 showing markers of a recollection-like process. We therefore conducted an additional ROC-  
17 fitting procedure on responses to a subsample of questions which encouraged more  
18 complete use of the response scale. We eliminated 89 items in which 4 or more of the 12  
19 possible response bins were unused (sure false to sure true [6] across the item's use as a  
20 target and a lure [2]). This question exclusion had the effect of decreasing overall EV  $d'$  from  
21 1.08 [1.05, 1.12] to 0.79 [0.76, 0.84] but had no effect on EV  $c$ , which was -.123 [-.145, -  
22 .101] before the exclusion and -.125 [-.149, -.100] after. The excluded items were thus  
23 confirmed disproportionate contributors of high sensitivity responses. Figure 2B plots the  
24 question-selective data against predicted ROCs derived from the mean parameters output  
25 by the MLE fitting procedures. These data still show symmetry about  $y = 1 - x$ , and although  
26 the reduced sensitivity is evident as a decreased area under the curve, the first and last  
27 ROC points remain inconsistent with a curve that tends to [0,0] and [1,1].  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45

46 Across both the full dataset and the question-selective subset, the DP-RR model curve  
47 shows the most consistent overlap with all of the observed ROC points. The UEV and DP  
48 models both show an overshoot of points in the middle range, with this particularly prominent  
49 for the UEV model fit to the full dataset (Figure 2A). The UEV and DP models also show an  
50 undershoot in the upper right as they both tend to [1, 1]. (The descriptives summarising the  
51 parameters used to generate each ROC curve are shown in Table 1.) The zROCs for both  
52 datasets are also best accommodated by the DP-RR model. The curves at both upper and  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 lower aspects of the line are accommodated by the DP-RR model, but cannot be  
4 accommodated by the linear UEV model or the single curve of the DP model.  
5  
6  
7

8  
9 Given the close correspondence between the full dataset and the question-selective subset,  
10 and the risk of biasing our results towards best fitting the DP-RR model, we use only  
11 question-selective data in future analyses.  
12  
13  
14

15  
16  
17 (Figure 2 about here)

18  
19 (Table 1 about here)  
20  
21  
22

23 ROC curves and zROC lines based on averaged model parameters do not necessarily best  
24 represent the fit/misfit of the models on a per participant basis. We therefore illustrate each  
25 participant's model misfit (estimated ROC point position relative to the observed position) for  
26 each ROC criterion point in Figure 3. Examination of the scatterplots reveals that the DP-RR  
27 model is noisier in its fits than the UEV and DP models, though the overall mean misfit,  
28 reflecting the averaged model fits, shows less systematic bias (mean misfit point closer to  
29 the origin). To formally assess the absolute misfit at each point, we calculated Euclidean  
30 misfit distances (shown in Figure 4) and entered these values into a 3 (model: UEV; DP; DP-  
31 RR) x 5 (ROC point) ANOVA. As suggested by the scatterplots, there was a main effect of  
32 model,  $F(2, 1010) = 52.57, p < .001, \eta_p^2 = .094$ , with the DP-RR model showing greatest  
33 absolute misfit ( $M = .044, [.042, .047]$ ) relative to the DP ( $M = .037, [.036, .039]$ ) and UEV ( $M$   
34  $= .035, [.033, .036]$ ) models. (There was also a significant main effect of ROC point,  $F(4,$   
35  $2020) = 107.01, p < .001, \eta_p^2 = .175$ , and a significant interaction,  $F(8, 4040) = 117.22, p <$   
36  $.001, \eta_p^2 = .188$ .) These two illustrations suggest that, whilst the DP-RR model's shape  
37 appears to best accommodate group data from the semantic memory task, misfit on the  
38 participant level may be higher for the novel model compared to the two established models.  
39  
40 We now turn to analysis of the fit statistics to assess whether misfit is indeed greater for the  
41 DP-RR model than the others.  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

(Figure 3 about here)

(Figure 4 about here)

Log likelihood ( $LL$ ) parameters represent the probability of the data given the parameter estimates. They are always negative, with more positive (closer to 0) values indicating better model fits. We calculated  $LL$  for each participant (see Table 1 for descriptives) and entered the parameters for the question-selective dataset into a three-way (model: UEV, DP, DP-RR) repeated measures ANOVA (identical analysis of the full dataset yielded revealed the same effects). There was a significant difference in  $LL$  according to model,  $F(2,998) = 69.82$ ,  $p < .001$ ,  $\eta_p^2 = .123$ , with the DP-RR model yielding *better* (more positive) fits than the other two models. This advantage in the fit statistics was borne out in the proportion of participants for whom the DP-RR model was best fitting, compared to the UEV and DP models: .538, .318 and .144 respectively,  $\chi^2(2) = 116.96$ ,  $p < .001$ . It should be noted that the UEV and DP models both typically account for ~99.9% of variance in standard episodic ROCs (e.g. O'Connor, Guhl, Cox & Dobbins, 2011), meaning that these DP-RR advantages are noteworthy.

The majority of analyses support the DP-RR model as best characterising semantic memory retrieval within these recognition tests. The proximity of points to the upper and left bounds in the ROC, alternatively represented by the curve at upper and lower aspects of the zROC are not well accommodated by either the UEV or DP models (Figure 2). Conversely, the DP-RR model, which affords recollection-like benefit to both the high confidence identification of targets and the high confidence rejection of lures, best estimates the parameters characterising retrieval here. Although fits to the DP-RR model are noisier than to competing models, they result in less systematic deviation from the observed data (Figures 3 and 4) and as a result yield significantly better fit parameters than the UEV and DP models. In sum, Experiment 1 demonstrates that semantic retrieval does indeed display an ROC marker of a

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 recollection-like threshold process (albeit operationalized slightly differently to episodic  
4 recollection), and we proceed to Experiment 2 with the aim of using the DP-RR model to  
5 best recover semantic retrieval threshold parameters to for within-subjects comparison with  
6 episodic retrieval threshold parameters.  
7  
8  
9

**EXPERIMENT 2**

10  
11  
12  
13  
14  
15 We next looked for indication of whether semantic and episodic threshold retrieval share the  
16 same underlying process – whether the semantic threshold retrieval process is comparable  
17 to recollection from to episodic memory. To this end, we administered the semantic  
18 recognition test from Experiment 1 alongside an analogous episodic recognition task. A  
19 correlation between the recovered semantic and episodic threshold parameters could  
20 indicate some association which could be explored with further analyses. Experiment 2 also  
21 allowed us to attempt to replicate the findings from Experiment 1 using an independent  
22 sample. Thus, we first compared recognition fits to the three models, before using the model  
23 parameters recovered from these fits to search for an association in threshold parameters  
24 across tasks.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

**METHODS****Stimuli**

37  
38  
39  
40  
41  
42 As in Experiment 1, Experiment 2 was programmed using JavaScript and presented to  
43 participants via their internet browsers.  
44  
45

46  
47  
48 *Semantic Test.* The 411 general knowledge questions used in the question-selective subset  
49 of Experiment 1 were used in the semantic test. For each participant, a random sample of 50  
50 semantic questions (25 targets, 25 lures) was drawn from this question pool.  
51  
52  
53

54  
55  
56 *Episodic Test.* A word pool comprising 2200 singular, common nouns from the English  
57 Lexicon Project (Balota et al., 2007) was used in the episodic test. A Hyperspace Analogue  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 to Language (HAL) frequency high-pass cut-off of 7.70 served to exclude highly distinctive  
4 items (final word list characteristics: mean HAL frequency = 8.98, mean word length = 7.24,  
5 mean number of syllables = 2.43). For each participant, a random sample of 100 episodic  
6 memory questions was administered in a single study-test block. Fifty words presented at  
7 both study and test (targets) and 50 words were presented only at test (lures).  
8  
9  
10  
11  
12  
13  
14

**Participants**

15  
16  
17 316 participants (115 men, 187 women, 14 did not disclose sex) completed the online  
18 experiment. Recruitment and ethics procedures were identical to those in Experiment 1. Of  
19 the 301 who disclosed their age, mean reported age was 27.7 years (SD = 11.8). Once  
20 again, we found no differences in EV  $d'$  between the 240 participants who reported that  
21 English was their first language and the 62 participants who reported that it was not  
22 (semantic: first language  $M = .684$  [.624, .744], second language  $M = .692$  [.581, .803];  
23 episodic: first language  $M = 2.00$  [1.87, 2.13], second language  $M = 2.10$  [1.92, 2.27]).  
24 Similarly we found no differences in EV  $c$  over the two memory tests (semantic: first  
25 language  $M = -.146$  [-.183, -.108], second language  $M = -.178$  [-.256, -.100]; episodic: first  
26 language  $M = .061$  [.018, .103], second language  $M = .125$  [.022, .228]). We included all  
27 participants in subsequent analyses.  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

**Design and Procedure**

42  
43 Onscreen instructions were followed by the episodic study-test block, and then the semantic  
44 test block. In the episodic study phase, a 0.5 s fixation cross preceded each self-paced  
45 study trial (RT  $M = 2.30$  s). During each study trial, the fixation was replaced by a word  
46 presented in 850% font size in the middle of the screen. Below the word was a task cue  
47 “syllables?” (220% font), and six response buttons, 1 to 6+. Participants could use a  
48 keyboard or touchscreen (where available) to respond. After a response had been rendered,  
49 the fixation period before the next trial was initiated. After 50 study trials, the episodic test  
50 phase was initiated. A 0.5 s fixation again preceded each self-paced test trial. Test words  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

(850% font) were presented in the middle of the screen, above the task cue “recognition?” (220% font). The six responses buttons were, from left to right: new sure 1, new probably 2, new guess 3, old guess 4, old probably 5, old sure 6. Once the 100th trial had been responded to, participants were taken to the self-paced semantic test, which was a slightly shorter version of that described for Experiment 1 (50 rather than 60 trials). The semantic test was followed by a debrief, which provided feedback on participant accuracy and confidence across the two tests.

**RESULTS****Item Selection**

We excluded responses with response times exceeding 10s. In the semantic test, we excluded 4.97 [4.33, 5.60] responses per participants, which reduced mean response times from 6.16 s [5.91 s, 6.41 s] to 5.18 s [5.06 s, 5.30 s]. For the episodic test, although participants would not have been able to look up the answers online, for consistency we applied the same response exclusion criteria, excluding 0.67 [0.50, 0.85] responses per participant, which reduced overall response times from 2.10 s [2.01 s, 2.18 s] to 2.00 s [1.94 s, 2.07 s].

The overall semantic test EV  $d'$  of .688 [.630, .745] was lower than in the question selective subset of Experiment 1, suggesting the 411 semantic items used did not result in excessively high sensitivity responses. (Mean EV  $c$  was -.176 [-.212, -.140]). We therefore proceeded with the analyses without excluding any additional semantic items.

**Receiver Operating Characteristics and Model Fits**

We first addressed how the competing models from Experiment 1 would fit the data obtained in each memory test. We anticipated that the DP-RR model would best fit semantic retrieval responses, whereas one of the established models for episodic recognition would best fit episodic retrieval responses.

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3  
4  
5 We used an analytic procedure identical to that used in Experiment 1. Successful fits were  
6  
7 obtained in 309 participants (98.1%) for the semantic task and in 280 participants (88.9%)  
8  
9 for the episodic task. The difference here is likely driven by the difference in  $d'$  over the two  
10  
11 tasks. Mean EV  $d'$  for the semantic task, 0.70 [0.64, 0.76] was far lower than for the episodic  
12  
13 task, 1.99 [1.88, 2.10]. Although this sensitivity disparity is potentially problematic when  
14  
15 adjudicating model fits, we restrict the bulk of our analyses below to sensitivity-matched  
16  
17 subsamples.  
18

19  
20  
21 Figure 5A uses the whole sample to plot the average observed ROC points against average  
22  
23 predicted ROC curves for semantic and episodic tasks. Once again, for the semantic task,  
24  
25 the points are approximately symmetrical about  $y = 1 - x$  and inconsistent with an ROC that  
26  
27 tends to [1, 1]. In contrast, the episodic task ROC shows points that could be from an ROC  
28  
29 that tends to [1, 1]. To best ensure that subsequent findings were attributable to the tasks  
30  
31 themselves and not differences in sensitivity associated with each task outlined above, for  
32  
33 the remainder of this section, we report analyses applied to subsamples whose sensitivity  
34  
35 were comparable (though descriptive for the full sample are shown in Table 2 and all  
36  
37 patterns of significance were almost identical across full and sensitivity-matched  
38  
39 subsamples). For the semantic task, we used a high-pass filter of 0.875, applied to EV  $d'$ ,  
40  
41 yielding 109 participants with mean EV  $d'$  of 1.24 [1.18, 1.30]. For the episodic task, we used  
42  
43 a low-pass filter of 1.925, applied to the EV  $d'$  estimate, yielding 108 participants with mean  
44  
45 EV  $d'$  of 1.26 [1.07, 1.44]. Figure 5B plots the average observed ROC points and average  
46  
47 predicted ROCs from fits to these subsamples. Once again, the proximity of semantic ROC  
48  
49 points to the top boundary, suggests a better fit for the DP-RR model to the semantic data,  
50  
51 though the best-fitting model for the episodic data is less obvious.  
52  
53  
54

55  
56 (Figure 5 about here)  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3  
4  
5 We first examine analyses of data from the semantic recognition task. Whilst the DP-RR  
6  
7 model appears to best fit the data on a group-averaged level, the misfit scatterplot in Figure  
8  
9 6 once again shows that the DP-RR fits are noisier than those for the competing models,  
10  
11 though with less systematic deviation (as demonstrated by reduced deviation from [0, 0] in  
12  
13 the mean plots). Absolute misfit distances (Figure 8 Panel A) in a 3 (model) x 5 (ROC point)  
14  
15 ANOVA demonstrated a main effect of model,  $F(2, 216) = 19.79, p < .001, \eta_p^2 = .155$ , with  
16  
17 DP-RR showing greatest absolute misfit ( $M = .039, [.035, .043]$ ) relative to the DP ( $M = .034,$   
18  
19  $[.031, .037]$ ) and UEV ( $M = .028, [.025, .030]$ ) models. (There was also a significant main  
20  
21 effect of ROC point,  $F(4, 432) = 17.65, p < .001, \eta_p^2 = .140$ , and a significant interaction,  $F(8,$   
22  
23  $864) = 13.29, p < .001, \eta_p^2 = .110$ .)  
24  
25

26  
27 (Figure 6 about here)  
28  
29  
30  
31

32 In spite of the misfit distances, *LL* fit statistics (Table 2) again showed that the semantic data  
33  
34 within the sensitivity-matched sample are best accommodated by the DP-RR model.  
35  
36 Entered into a repeated measures ANOVA, a significant effect of model on *LL* was found,  
37  
38  $F(2,216) = 10.24, p < .001, \eta_p^2 = .087$ , with the DP-RR model yielding better fits than the  
39  
40 other two models. This advantage in the fit statistics was borne out in the proportion of  
41  
42 participants for whom the DP-RR model was best fitting, compared to the UEV and DP  
43  
44 models: .495, .385 and .119 respectively,  $\chi^2(2) = 24.46, p < .001$ . Thus, these data from the  
45  
46 semantic task in Experiment 2 replicated those of Experiment 1 in ROC shape, misfit  
47  
48 distance and model fit.  
49  
50

51  
52 (Table 2 about here)  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 Moving to the episodic recognition task, group-averaged ROC fits appear more ambiguous  
4 as to the best-fitting model, but the misfit scatterplot in Figure 7 shows that the DP-RR fits  
5 are again noisier than those for the competing models. This time, and in contrast to the  
6 means plots in Figure 6, Figure 7 shows that the DP-RR model yields more systematic misfit  
7 (greater mean deviation from [0, 0]) than the competing models. Absolute misfit distances  
8 (Figure 8 Panel B) in a 3 (model) x 5 (ROC point) ANOVA yielded a main effect of model,  
9  $F(2, 214) = 19.48, p < .001, \eta_p^2 = .154$ , with DP-RR showing greatest absolute misfit ( $M =$   
10  $.024, [.020, .027]$ ) relative to the DP ( $M = .017, [.016, .019]$ ) and UEV models ( $M = .016,$   
11  $[.014, .018]$ ). (There was also a significant main effect of ROC point,  $F(4, 428) = 35.10, p <$   
12  $.001, \eta_p^2 = .247$ , and a significant interaction,  $F(8, 856) = 18.89, p < .001, \eta_p^2 = .150$ .)

23  
24  
25 (Figure 7 about here)

26  
27  
28 (Figure 8 about here)

29  
30  
31  
32  
33 Consistent with the more ambiguous ROC fits, *LL* fit statistics were less clear-cut for the  
34 episodic data. There was no significant effect of model on *LL* in a repeated-measures  
35 ANOVA,  $F(2,214) = 1.03, p = .360, \eta_p^2 = .009$ , and although the UEV model yielded  
36 numerically better fits than its competitors, the *LL* values are all very close. Overall, the DP-  
37 RR model's advantage in fitting semantic data is lost when applied to episodic data, with the  
38 UEV model fitting a greater proportion of participants best, compared to the DP-RR and DP  
39 models: .491, .287 and .222 respectively,  $\chi^2(2) = 12.72, p = .002$ . These new data from the  
40 episodic task in Experiment 2 demonstrate that the DP-RR model is not a generally better-  
41 fitting model for all recognition tasks, but is specific in its advantage for certain tasks – in this  
42 case to data obtained from the semantic recognition tasks of Experiments 1 and 2.

43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55 Having established that DP-RR-recovered threshold process parameters are reliable and  
56 specific in quantifying responses to semantic recognition memory responding, we now  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 proceed to test the association between within-subjects semantic and episodic threshold  
4 process estimates using the two models that provide such an estimate: the novel DP-RR  
5 model within semantic data, and the established DP models within episodic data.  
6  
7  
8  
9

**Association between Semantic and Episodic Threshold Process Estimates**

10  
11  
12 As a preliminary check, we first examined the full sample correlations between EV  $d'$  and  $c$   
13 parameters and mean confidence responses to hits and correct rejections. Within episodic  
14 and semantic tasks, we anticipated that: i)  $d'$  would be positively correlated with both hit and  
15 correct rejection confidence; and ii) the two confidence means would be positively  
16 correlated. We also anticipated that: iii) each of the four parameters would be positively  
17 correlated with its cross-domain equivalent. Table 3 (shaded cells) shows that these  
18 predictions were largely supported by the data, but there was no correlation across domains  
19 between the EV  $d'$  or the  $c$  parameters.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30

31 (Table 3 about here)  
32  
33  
34

35 We then proceeded to the main correlation analysis testing the association between the  
36 threshold parameters ( $R$ ) estimated for semantic and episodic tasks. To do this we identified  
37 all participants for whom both the DP-RR model had recovered a fit for their semantic task  
38 responses and the DP model had recovered a fit for their episodic task responses. This left  
39 278 participants (88.3% of the original full sample). In this subsample, mean semantic EV  $d'$   
40 was 0.69 [0.63, 0.76] and mean episodic EV  $d'$  was 1.99 [1.88, 2.10].  
41  
42  
43  
44  
45  
46  
47  
48

49 As a preliminary check, we found no correlation between the equal variance  $d'$  parameters  
50 across the two tasks,  $r(276) = .031$ , 95% CIs based on 10000 bootstrapping samples = [-  
51 .066, .121],  $p = .609$ . As in the full sample, in the subsample there was no significant  
52 relationship between participants' abilities to discriminate true from false general knowledge  
53 questions in the semantic test and their abilities to recognise old from new words in the  
54  
55  
56  
57  
58  
59  
60



## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 episodic test. We proceeded to test the correlation between semantic DP-RR and episodic  
4 DP *R* parameters, as an association between the two (such as one driven by a reliance on  
5 the same retrieval process) could be hidden by a divergence in other processes contributing  
6 to overall sensitivity. No association whatsoever was observed between semantic and  
7 episodic *R* parameters,  $r(276) = -.001 [-.122, .120]$ ,  $p = .987$ , (see Figure 9; there was also  
8 no correlation between respective DP-RR and DP *d'* parameters recovered using the same  
9 fitting procedures  $r[276] = .085 [-.048, .215]$ ,  $p = .156$ ). In case a true association between  
10 threshold parameters was masked by *R* parameters estimated to be 0, we removed all  
11 participants with DP-RR or DP *R* parameters less than 0.01 and ran a second correlation,  
12 also showing no relationship,  $r(153) = .065 [-.099, .225]$ ,  $p = .421$ . Thus, we found no  
13 association between the two threshold processes we propose contribute to semantic and  
14 episodic retrieval.  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

29 (Figure 9 about here)  
30  
31  
32

**DISCUSSION**

33  
34  
35 We used a novel single-item semantic recognition task to find evidence for a recollection-like  
36 threshold retrieval process within semantic memory. In Experiment 1 we found that the DP-  
37 RR model, a variation of the dual process model that allows recollection to contribute to both  
38 the identification of targets and the rejection of lures, provides better fits to the observed  
39 data than the unequal variance single process model and the standard dual process model.  
40 In Experiment 2 found the DP-RR model fitting advantage was specific to our semantic test,  
41 and does not persist in the standard single-item episodic recognition test. Finally, we used  
42 the threshold parameters recovered using these fitting procedures and found there to be no  
43 association between the relative magnitudes of the threshold parameters recovered in  
44 semantic and episodic recognition.  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 Our experimental task recovered ROCs indicating the presence of a process analogous to  
4 episodic recollection in semantic memory. This finding is consistent with that of Waidergoren  
5 and colleagues' (2012) study using a semantic process dissociation task, but used an  
6 independent experimental paradigm to recover this convergent result. Specific to our  
7 procedure, the ROC model that best fit semantic recognition responding incorporated a  
8 recollection parameter necessary for dual process accounts of episodic memory (Yonelinas,  
9 Aly, Wang & Koen, 2010). Using the same terms as are used to justify similar patterns in  
10 episodic tasks, our results show that semantic memory retrieval includes a process by which  
11 we are able to identify answers to questions assessing knowledge as unambiguously true.  
12 This process likely draws on the recovery of contextualising information e.g. "I know with  
13 high certainty that Canberra is the capital of Australia because i) the capital isn't Sydney, ii)  
14 Canberra is in a territory called the Australian Capital Territory, and iii) I remember seeing  
15 the Aboriginal Tent Embassy outside Parliament House when I visited Canberra." This third  
16 contextual point is an example where episodic recollection can provide context for semantic  
17 knowledge, but we argue that the semantic threshold process is not necessarily dependent  
18 upon episodic content: we can imagine that a number of different sources of information can  
19 bring to bear on recognition decisions. Nonetheless, this raises a potentially important  
20 consideration for those looking to extend this finding further. Whereas the mappings of  
21 responses to components of dual process episodic memory models are considered fuzzy  
22 because of the lack of process purity (high confidence target recognition can be justified by  
23 *both* recollection and familiarity; Wixted, 2007), there is in this case a lack of 'system purity'  
24 in the putative dual process semantic memory system. Episodic and semantic content may  
25 well be used to justify semantic recollection-like retrieval.

26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52 This proposed system impurity revisits a similar crossover in the origins of the episodic dual  
53 process account. Tulving (1985) originally intended the RK procedure to differentiate  
54 episodic material (R responses) from semantic material (K responses). Indeed, such a  
55 taxonomy without consideration of more recent work, would have led us to hypothesise that  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 semantic retrieval should recruit only a single process, familiarity. Importantly though,  
4  
5 Tulving's system distinction became a process distinction as episodic recognition  
6  
7 researchers began to consider these responses qualitatively different, but both within the  
8  
9 episodic memory system (e.g. Gardiner, 1988). Taken together, these system and process  
10  
11 impurities point to flexible memory arrangements within which content consistent with  
12  
13 response justification, regardless of domain, can be amalgamated. As regards our  
14  
15 examination of semantic memory processes, a RK procedure requiring participants to give  
16  
17 justifications could provide a better indication of the breakdown of episodic and semantic  
18  
19 justifications. We can also envisage scenarios in which non-memorial reasoning and  
20  
21 problem-solving processes might contribute to responses, and could therefore be  
22  
23 considered influential of a memory decision-making system drawing on a range of cognitive  
24  
25 processes. In the first instance though, it is clear that Tulving's taxonomy continues to  
26  
27 require revision, with mounting evidence that content and processes which would previously  
28  
29 have been considered exclusively episodic, weighing in on semantic retrieval too.  
30  
31

32  
33 A consistent finding across our two experiments is the overall advantage the DP-RR model  
34  
35 showed in accommodating the semantic ROC data. Although both dual process models  
36  
37 incorporate the threshold process aiding target detection, the key advantage for the DP-RR  
38  
39 model is its ability to model the upper intercept as  $[1-R, 1]$ . This intercept replaces the  $[1, 1]$   
40  
41 intercept observed in episodic recognition, and is indicative of a high-certainty lure rejection  
42  
43 process unique to these particular semantic materials. We propose that this results from the  
44  
45 mutual exclusivity of our semantic questions, which afford the following reasoning: "If the  
46  
47 answer to the question is Ahab then it cannot also be Ishmael". Such recall-to-reject  
48  
49 strategies can apply in episodic memory too, as discussed earlier (e.g. Rotello et al. 2000).  
50  
51 These apply especially where one has to differentiate between two familiar stimuli. But, in  
52  
53 typical item-by-item tests of episodic recognition memory, the reasoning "'table' was not a  
54  
55 studied word, because 'apple' was" does not hold. Thus, recollection-like retrieval  
56  
57 contributes to semantic recognition responses but, given the mutual exclusivity of target and  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 lure options for each semantic question, its utility here is two-fold, to confirm targets and to  
4  
5 reject lures.  
6  
7

8  
9 We suggest that this deviation from what would be expected based on standard episodic  
10  
11 ROCs is driven by the nature of the semantic task used. If we had used a semantic task for  
12  
13 which correct answers were not mutually exclusive, we would likely have observed an  
14  
15 asymmetric benefit to targets, and not lures, akin to that observed in episodic recognition.  
16  
17 More generally, in considering how to refine the present findings, we believe it is important to  
18  
19 disentangle the question format from the memory store interrogated. If it is taken that  
20  
21 recollection operates within both semantic and episodic systems, the mutual exclusivity of  
22  
23 the cue-probe combination should be what determines whether the DP or the DP-RR best  
24  
25 fits data obtained from either system. We contend that responses to general knowledge  
26  
27 questions which have more than one answer (e.g. "The following word was invented by  
28  
29 Shakespeare:" where any number of words could be correct) would be more likely to be best  
30  
31 fit by the DP rather than the DP-RR model (though this also depends on the recollective  
32  
33 affordance of the lure probe). Alternatively, an episodic paired-associate recognition task in  
34  
35 which retrieval of the target would rule out the lure, would be best fit by the DP-RR task.  
36  
37 Such tasks would go some way towards demonstrating the format-dependent utility of a  
38  
39 threshold retrieval process in all forms of memory decision-making.  
40  
41  
42

43  
44 Having found that the overall pattern of results points towards the DP-RR model best fitting  
45  
46 the semantic data, it is curious that the same model produces greater absolute misfit  
47  
48 distances than its competitors. An insight into why this might be can be found from  
49  
50 examination of the correlations between absolute misfit distances across criterion points.  
51  
52 There are 10 correlations (C1-C2, C1-C3 ... C4-C5), over which the Experiment 1 grand  
53  
54 average correlation  $r$  is .532 [.452, .612] for the DP-RR model, .173 [.075, .271] for DP and  
55  
56 .288 [.136, .441] for UEV. That the DP-RR model shows greatest coherence in its misfit could  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 result from it providing fair fits across all points, rather than good fits for most, compared to  
4  
5 an excellent/bad fit for the remaining one or two points.  
6  
7

8  
9 Relatedly, a modification of the DP-RR model which would have improved the absolute  
10  
11 misfit as well as the overall fits, would be to remove the yoking of target identification- and  
12  
13 lure rejection-aiding recollective processes. Whilst the questions/cues remain identical  
14  
15 across target/lure trials, the probes themselves have the potential to elicit differing  
16  
17 recollective response—in addition to recall and recollection cued by the question, a target  
18  
19 probe may in itself trigger recollection of context that confirms its status, whilst a lure probe  
20  
21 may trigger context that disconfirms it as the correct response. Operationalising this in a  
22  
23 signal detection model straightforward (indeed, it is provided as a model within Koen et al.'s  
24  
25 2014 ROC toolbox), but it would necessitate an additional parameter relative to the two  
26  
27 established models. We wanted to avoid this in the first instance—there are issues with  
28  
29 using fit statistics that penalise for additional parameters, see Supplementary Material B—  
30  
31 and the *LL* fits suggest that in spite of this constraint, the DP-RR model is largely superior to  
32  
33 its competitors in its fits. Both issues remain to be fully elucidated, but should not detract  
34  
35 from the overall superiority of the DP-RR model, even in its current form, when it comes to  
36  
37 fitting the semantic data.  
38  
39

40  
41 A criticism of the presented data, which led to knock-on effects for our analyses, concerns  
42  
43 the sensitivity discrepancy between the episodic and semantic tests of Experiment 2.  
44  
45 Episodic task performance was far higher than semantic performance, meaning that any  
46  
47 differences in fits across task could have been attributed to the models' capacities to fit at  
48  
49 differing sensitivities, rather than their capacities to fit based on the tasks themselves. Our  
50  
51 solution here was to restrict our sample to subsamples with matched sensitivities, and while  
52  
53 this overcame the problem of differential sensitivity, it removed the within-subjects strength  
54  
55 of the Experiment 2 design. It would have been preferable to have been able to use the  
56  
57 same samples for both sets of fits, and this is a design consideration to be borne in mind  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 both for future studies and the interpretation of the correlations between  $R$  parameters  
4  
5 derived from these two tests.  
6  
7

8  
9 We found no association between the contributions of semantic and episodic threshold  
10  
11 parameters to recognition performance in each domain. Significant  $R$ -parameter correlations  
12  
13 estimated using three different recognition tasks (albeit within the same episodic memory  
14  
15 system), have previously been found within subjects (e.g. Koen & Yonelinas, 2016),  
16  
17 suggesting that this approach yields correlations when performance depends on the same  
18  
19 underlying process. Thus, it would be tempting to argue that the absence of a relationship  
20  
21 across our tasks indicates that the two threshold processes identified are not unitary. The  
22  
23 opposite case would have been far easier to make if a significant correlation had been found  
24  
25 (though with the caveat that third variables can underpin shared variance in cognition), but it  
26  
27 is not possible to argue convincingly that the absence of a correlation indicates the absence  
28  
29 of a relationship.  
30  
31

32  
33 A number of factors, not least the previously discussed sensitivity discrepancy across tests,  
34  
35 may have contributed to the present results. The varied nature of the semantic task,  
36  
37 compared to the episodic task, may have played a role in masking any association between  
38  
39  $R$  parameters. There are vast individual differences in the near limitless domains of semantic  
40  
41 knowledge. While the within-subjects tendency to retrieve context to support semantic  
42  
43 judgements could be static, if Participant A is only able to draw on very little contextual  
44  
45 material to support rejection of a pop music lure, but Participant B's interest in the subject  
46  
47 furnishes them with far contextual material, it follows that Participant B will show greater  
48  
49 evidence of recollection than Participant A, regardless of their individual tendencies. We can  
50  
51 also return to discussion of the system purity of semantic and episodic retrieval to explain  
52  
53 any lack of an association. Some semantic questions may have been easier if the  
54  
55 participant had reasoned the target as true or the lure as false. For the episodic task, some  
56  
57 participants may have been better able to sustain attention to the study list than others,  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 thereby enhancing encoding. Thus, reasoning and attention abilities may contaminate the  
4 estimates of  $R$ , further diluting the chances of finding an association. Greater experimental  
5 control could go some way towards eliminating these confounds. For example, there is the  
6 intriguing possibility of having participants study artificial semantic material under controlled  
7 conditions, which is then tested both episodically and semantically. However, such a  
8 procedure lay outwith the realms of the current set of online experiments, and the question  
9 of whether the absence of an association indicates the true absence of a mechanistic  
10 relationship remains to be answered.  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20

21 As a starting point for theoretical discussions, our results suggest that semantic and episodic  
22 memory systems recruit multiple retrieval processes which show parallels in function. More  
23 intriguingly, they open the door to the possibilities of both convergence and divergence  
24 within the same threshold process, applied to retrieval from different memory stores. In  
25 applying established episodic memory analysis techniques to the study of semantic memory,  
26 we have produced further evidence to suggest we should rethink widely held assumptions of  
27 how we bring knowledge to mind.  
28  
29  
30  
31  
32  
33  
34  
35  
36

**REFERENCES**

- 37  
38  
39 Arndt, J., & Reder, L.M. (2002). Word frequency and receiver operating characteristic curves  
40 in recognition memory: evidence for a dual-process interpretation. *Journal of*  
41 *Experimental Psychology: Learning, Memory, and Cognition*, 28(5), 830-842.  
42 <http://dx.doi.org/10.1037/0278-7393.28.5.830>  
43  
44  
45  
46  
47 Balota, D.A., Yap, M.J., Cortese, M.J., Hutchison, K.A., Kessler, B., Loftis, B., et al (2007).  
48 The English lexicon project. *Behavior Research Methods*, 39, 445–459.  
49 <http://dx.doi.org/10.3758/BF03193014>  
50  
51  
52  
53  
54 Bowles, B., Harlow, I.M., Meeking, M.M. & Köhler, S. (2012). Discriminating famous from  
55 fictional names based on lifetime experience: evidence in support of a signal-detection  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

- 1  
2  
3 model based on finite mixture distributions. *Journal of Experimental Psychology:*  
4 *Learning, Memory, and Cognition*, 38, 78-91. <http://dx.doi.org/10.1037/a0025198>  
5  
6  
7 Bowles, B. & Köhler, S. (2014). Availability of semantic knowledge in familiar-only  
8  
9 experiences for names. *Journal of Experimental Psychology: Learning, Memory, and*  
10 *Cognition*, 40, 724-737. <http://dx.doi.org/10.1037/a0035584>  
11  
12  
13 Conway, M.A., Gardiner, J.M., Perfect, T.J., Anderson, S.J. & Cohen, G.M. (1997). Changes  
14  
15 in Memory Awareness During Learning: The Acquisition of Knowledge by Psychology  
16  
17 Undergraduates. *Journal of Experimental Psychology: General*, 4, 393-413.  
18  
19 <http://dx.doi.org/10.1037/0096-3445.126.4.393>  
20  
21 Donaldson, W. (1996). The role of decision processes in remembering and knowing.  
22  
23 *Memory & Cognition*, 24(4), 523-533. <http://dx.doi.org/10.3758/BF03200940>  
24  
25  
26 Dunn J.C. (2004) Remember-know: A matter of confidence. *Psychological Review*, 111,  
27  
28 524-542. <http://dx.doi.org/10.1037/0033-295X.111.2.524>  
29  
30 Gardiner, J.M. (1988). Functional aspects of recollective experience. *Memory & Cognition*,  
31  
32 16, 309-313. <http://dx.doi.org/10.3758/BF03197041>  
33  
34 Green, D.M., & Swets, J.A. (1966). *Signal detection theory and psychophysics*. New York:  
35  
36 Wiley.  
37  
38 Heathcote, A. (2003). Item recognition memory and the receiver operating characteristic.  
39  
40 *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 1210-  
41  
42 1230. <http://dx.doi.org/10.1037/0278-7393.29.6.1210>  
43  
44 Jacoby, L.L. (1991). A Process dissociation framework: Separating automatic from  
45  
46 intentional uses of memory. *Journal of Memory and Language*, 30, 513-541.  
47  
48 [http://dx.doi.org/10.1016/0749-596X\(91\)90025-F](http://dx.doi.org/10.1016/0749-596X(91)90025-F)  
49  
50 Klein, S.B. (2013). Making the case that episodic recollection is attributable to operations  
51  
52 occurring at retrieval rather than to content stored in a dedicated subsystem of long-  
53  
54 term memory. *Frontiers in Behavioral Neuroscience*, 7:3.  
55  
56 <http://dx.doi.org/10.3389/fnbeh.2013.00003>  
57  
58  
59  
60



## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

- 1  
2  
3 Koen, J.D., Barrett, F.S., Harlow, I.M. & Yonelinas, A.P. (2014). *ROC Toolbox Manual*.  
4  
5 Davis: The Regents of the University of California.  
6  
7 Koen, J.D. & Yonelinas, A.P. (2016). Recollection, not familiarity, decreases in healthy  
8  
9 ageing: Converging evidence from four estimation methods. *Memory*, 24, 75-88.  
10  
11 <http://dx.doi.org/10.1080/09658211.2014.985590>  
12  
13 Macmillan, N.A., & Creelman, C.D. (2005). *Detection theory: A user's guide (2nd ed.)*. New  
14  
15 York: Lawrence Erlbaum.  
16  
17 Mickes, L., Wais, P.E. & Wixted, J.T. (2009) Recollection is a continuous process.  
18  
19 *Psychological Science*, 20, 509-515. <http://dx.doi.org/10.1111/j.1467->  
20  
21 [9280.2009.02324.x](http://dx.doi.org/10.1111/j.1467-9280.2009.02324.x)  
22  
23 Mill, R.D., & O'Connor, A.R. (2014). Question format shifts bias away from the emphasised  
24  
25 response in tests of recognition memory. *Consciousness and Cognition*, 30, 91-104.  
26  
27 <http://dx.doi.org/10.1016/j.concog.2014.09.006>  
28  
29 O'Connor, A.R., Guhl, E.N., Cox, J.C. & Dobbins, I.G. (2011). Some memories are odder  
30  
31 than others: Judgments of episodic oddity violate known decision rules. *Journal of*  
32  
33 *Memory and Language*, 64, 299-315. <http://dx.doi.org/10.1016/j.jml.2011.02.001>  
34  
35 Payne, D.G., Elie, C.J., Blackwell, J.M. & Neuschatz, J.S. (1996). Memory illusions:  
36  
37 Recalling, recognizing, and recollecting events that never occurred. *Journal of Memory*  
38  
39 *& Language*, 35, 261–285. <http://dx.doi.org/10.1006/jmla.1996.0015>  
40  
41 Ratcliff, R., Sheu, C.-F., & Gronlund, S.D. (1992). Testing global memory models using ROC  
42  
43 curves. *Psychological Review*, 99, 518–535. <http://dx.doi.org/10.1037/0033->  
44  
45 [295X.99.3.518](http://dx.doi.org/10.1037/0033-295X.99.3.518)  
46  
47 Rotello, C.M., Macmillan, N.A., & Van Tassel, G. (2000). Recall-to-reject in recognition:  
48  
49 Evidence from ROC curves. *Journal of Memory and Language*, 43(1), 67-88.  
50  
51 <http://dx.doi.org/10.1006/jmla.1999.2701>  
52  
53 Shiffrin, R.M. & Steyvers, M. (1997). A model for recognition memory: REM—retrieving  
54  
55 effectively from memory. *Psychonomic Bulletin & Review*, 4, 145-166.  
56  
57 <http://dx.doi.org/10.3758/BF03209391>  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

- 1  
2  
3 Snodgrass, J.G., & Corwin, J. (1988). Pragmatics of measuring recognition memory:  
4  
5 Applications to dementia and amnesia. *Journal of Experimental Psychology: General*,  
6  
7 117(1), 34–50. <http://dx.doi.org/10.1037/0096-3445.117.1.34>  
8
- 9 Suddendorf, T., & Corballis, M.C. (1997). Mental time travel and the evolution of the human  
10  
11 mind. *Genetic, social, and general psychology monographs*, 123(2), 133-167.  
12
- 13 Tulving, E. (1985). Memory and consciousness. *Canadian Psychology/Psychologie*  
14  
15 *Canadienne*, 26(1), 1-12. <http://dx.doi.org/10.1037/h0080017>  
16
- 17 Vilberg, K.L. & Rugg, M.D. (2008). Memory retrieval and the parietal cortex: a review of  
18  
19 evidence from a dual-process perspective. *Neuropsychologia*, 46, 1787-99.  
20  
21 <http://dx.doi.org/10.1016/j.neuropsychologia.2008.01.004>  
22
- 23 Waidergoren, S., Segalowicz, J. & Gilboa, A. (2012). Semantic memory recognition is  
24  
25 supported by intrinsic recollection-like processes: “The butcher on the bus” revisited.  
26  
27 *Neuropsychologia*, 50, 3573-3587.  
28  
29 <http://dx.doi.org/10.1016/j.neuropsychologia.2012.09.040>  
30
- 31 Wixted, J.T. (2007). Dual-process theory and signal-detection theory of recognition memory.  
32  
33 *Psychological Review*, 114, 152-176. <http://dx.doi.org/10.1037/0033-295X.114.1.152>  
34
- 35 Yonelinas, A.P. (1994). Receiver-operating characteristics in recognition memory: evidence  
36  
37 for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and*  
38  
39 *Cognition*, 20(6), 1341-1354. <http://dx.doi.org/10.1037/0278-7393.20.6.1341>  
40
- 41 Yonelinas, A.P. (2001). Components of episodic memory: The contribution of recollection  
42  
43 and familiarity. *The Philosophical Transactions of the Royal Society of London Series*  
44  
45 *B*, 356(1413), 1363-1374. <http://dx.doi.org/10.1098/rstb.2001.0939>  
46
- 47 Yonelinas, A.P. (2002). The nature of recollection and familiarity: A review of 30 years of  
48  
49 research. *Journal of memory and language*, 46(3), 441-517.  
50  
51 <http://dx.doi.org/10.1006/jmla.2002.2864>  
52
- 53 Yonelinas, A.P., Aly, M., Wang, W., & Koen, J.D. (2010). Recollection and familiarity:  
54  
55 examining controversial assumptions and new directions. *Hippocampus*, 20(11), 1178-  
56  
57 1194. <http://dx.doi.org/10.1002/hipo.20864>  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 Yonelinas, A.P., Dobbins, I., Szymanski, M.D., Dhaliwal, H.S., & King, L. (1996). Signal-  
4  
5 detection, threshold, and dual-process models of recognition memory: ROCs and  
6  
7 conscious recollection. *Consciousness and cognition*, 5(4), 418-441.

8  
9 <http://dx.doi.org/10.1006/ccog.1996.0026>

10  
11 Yonelinas, A.P., Kroll, N.E., Dobbins, I., Lazzara, M., & Knight, R.T. (1998). Recollection and  
12  
13 familiarity deficits in amnesia: convergence of remember-know, process dissociation,  
14  
15 and receiver operating characteristic data. *Neuropsychology*, 12(3), 323-339.

16  
17 <http://dx.doi.org/10.1037/0894-4105.12.3.323>  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

## TABLES WITH CAPTIONS

**Table 1: Experiment 1 model parameters and fit statistics for the competing signal detection models**

Dataset, Model	Model parameters			Fit statistic
	$d'$	$\sigma$	$R$	$LL$
<b>Experiment 1</b>				
<b>Full</b> ( $n = 502$ , data from 500 question items)				
UEV	1.98 [0.76, 3.22]	1.59 [0.55, 2.63]	-	-80.52 [-81.62, -79.41]
DP	1.01 [0.96, 1.05]	-	.213 [.197, .229]	-80.41 [-81.51, -79.31]
DP-RR	0.43 [0.38, 0.49]	-	0.278 [0.264, 0.292]	-79.69 [-80.78, -78.59]
<b>Question-selective</b> ( $n = 500$ , data from 411 question items)				
UEV	1.04 [0.99, 1.09]	1.06 [1.03, 1.10]	-	-68.44 [-69.46, -67.42]
DP	0.75 [0.71, 0.79]	-	0.16 [0.15, 0.18]	-68.40 [-69.41, -67.39]
DP-RR	0.33 [0.27, 0.38]	-	0.20 [0.19, 0.21]	-67.90 [-68.92, -66.90]

**Note:** Means are shown above 95% Confidence Intervals in brackets.  $LL$  represent log

likelihood parameters summarising each model's goodness of fit. A higher  $LL$  value indicates a better model fit.

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

**Table 2: Experiment 2 model parameters and fit statistics for the competing signal detection models within full and sensitivity-matched samples**

Dataset, Model	Model parameters			Fit statistic
	$d'$	$\sigma$	$R$	$LL$
<b>Experiment 2 Full Samples</b>				
<b>Semantic (<math>n = 309</math>, EV <math>d' = 0.85</math> [0.80, 0.91])</b>				
UEV	0.89 [0.83, 0.96]	1.07 [1.03, 1.12]	-	-67.59 [-68.93, -66.25]
DP	0.63 [0.58, 0.69]	-	0.14 [0.13, 0.16]	-67.52 [-68.85, -66.18]
DP-RR	0.33 [0.26, 0.39]	-	0.15 [0.14, 0.17]	-67.13 [-68.45, -65.81]
<b>Episodic (<math>n = 280</math>, EV <math>d' = 2.02</math> [1.91, 2.12])</b>				
UEV	2.44 [2.25, 2.62]	1.37 [1.28, 1.46]	-	-107.41 [-110.70, -104.13]
DP	1.65 [1.54, 1.76]	-	.307 [.275, .339]	-107.71 [-110.97, -104.45]
DP-RR	1.43 [1.29, 1.56]	-	.221 [.190, .252]	-107.89 [-111.18, -104.61]
<b>Experiment 2 Sensitivity-matched subsamples</b>				
<b>Semantic (<math>n = 116</math>, EV <math>d' = 1.35</math> [1.30, 1.40])</b>				
UEV	1.33 [1.23, 1.45]	0.98 [0.89, 1.08]	-	-66.18 [-68.09, -64.27]
DP	1.11 [1.04, 1.17]	-	.156 [.122, .190]	-66.33 [-68.24, -64.42]
DP-RR	0.70 [0.59, 0.82]	-	.207 [.177, .585]	-65.93 [-67.81, -64.05]
<b>Episodic (<math>n = 115</math>, EV <math>d' = 1.32</math> [1.14, 1.50])</b>				
UEV	1.41 [1.17, 1.64]	1.22 [1.15, 1.30]	-	-123.96 [-128.67, -119.27]
DP	1.06 [0.86, 1.26]	-	.219 [.180, .257]	-124.20 [-128.75, -119.65]
DP-RR	0.80 [0.60, 1.00]	-	.184 [.146, .222]	-124.17 [-128.86, -119.47]

**Note:** Means are shown above 95% Confidence Intervals in brackets. EV  $d'$  represents the equal variance signal detection sensitivity parameter (see Footnote 1). A higher  $LL$  value indicates a better model fit.

THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

Table 3: Correlations (Pearson’s r) between Basic Descriptives for Episodic and Semantic Tasks in Experiment 2

Task	Parameter	Episodic			Semantic			
		EV c	H conf.	CR conf.	EV d'	EV c	H conf.	CR conf.
Episodic	EV d'	.163** [.042, .272]	.406*** [.245, .546]	.405*** [.252, .533]	.081 [-.010, .168]	-.083 [-.176, .013]	-.003 [-.106, .102]	.043 [-.064, .156]
	EV c	-	-.192** [-.307, -.069]	.159** [.040, .267]	.108 [.002, .209]	.098 [-.017, .213]	.005 [-.097, .103]	-.007 [-.118, .096]
	H conf.	-	-	.685*** [.616, .746]	-.140* [-.251, -.033]	-.096 [-.219, .031]	.158** [.046, .262]	.216*** [.111, .320]
	CR conf.	-	-	-	-.132* [-.233, -.030]	-.094 [-.211, .028]	.159** [.050, .267]	.192*** [.084, .299]
Semantic	EV d'	-	-	-	-	.092 [-.027, .210]	.201*** [.082, .316]	.221*** [.110, .145]
	EV c	-	-	-	-	-	.061 [-.067, .185]	.021 [-.100, .145]
	H conf.	-	-	-	-	-	-	.555*** [.457, .640]

**Note:** Pearson’s correlation coefficients are shown above 95% Confidence Intervals in brackets (based on 10000 bootstrapping samples). All *ns* = 316. EV *d'* and EV *c* represents the equal variance signal detection sensitivity and bias parameters (see Footnote 1). *H conf.* and *CR conf.* represent the confidence (from 1-guess, to 3-sure) to hits and correct rejections. \*\*\* denotes correlation significant at the .001 level, \*\* at the .01 level and \* at the .05 level. Shaded cells denote correlations for which we made specific predictions. All predictions were supported.

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

**FIGURE TITLES AND CAPTIONS****Figure 1: Hypothetical ROCs from three competing signal detection models**

Across all models,  $d'$  is held constant at 1. The left-hand-side of panels shows ROCs, the right-hand-side shows zROCs. Panel A: The unequal variance (UEV) signal detection model shows the effect of increasing  $\sigma$  (lighter to darker lines) on ROC curve asymmetry about the diagonal  $y = 1 - x$  (the dotted line that extends from  $[0, 1]$  to  $[1, 0]$ ). This manifests as an alteration to the gradient and intercept of the zROC lines. Panel B: The dual process (DP) signal detection model shows the effects of increasing  $R$  on ROC curve asymmetry via alteration in the y axis intercept. This manifests as an increasing curve in the lower aspect of the zROC line. Panel C: the dual process recollect-to-reject (DP-RR) signal detection model shows the effects of increasing  $R$  on both ROC intercepts, with no change in curve symmetry. This manifests as an introduction of two curves, in the lower and upper aspects of the zROC.

**Figure 2: Experiment 1 observed data and mean fits from competing signal detection models.**

The left-hand-side of panels shows ROCs, the right-hand-side shows zROCs. The five points (shown at the intersection of their respective 95% CIs on the ROC plots) represent the mean observed ROC points. The three lines are constructed using the mean parameters recovered when participants are fit individually, with the UEV model shown as a dotted line, the DP model shown as a dashed line and the DP-RR model shown as a solid line. Panel A shows data and fits from the full dataset. Panel B shows data and fits from the question-selective subset.

**Figure 3: Experiment 1 misfit scatter plots**

Scatterplots showing target and lure discrepancies between the predicted and observed criterion points. The criterion points C1-C5 are the boundaries between “sure true”/“probably true” (C1) through to “probably false”/“sure false” (C5). The first three columns show misfit

## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

1  
2  
3 for each participant within the DP-RR (black points), DP (yellow points), and UEV (blue  
4 points) signal detection models. The ‘mean’ column plots the average misfit from each  
5 model simultaneously (error bars represent 95% CIs), using the same colour coding. Target  
6 and lure discrepancy are in  $d'$  units.  
7  
8  
9

**Figure 4: Experiment 1 absolute misfit**

10  
11  
12  
13  
14  
15 The mean Euclidean distance between predicted and observed criterion points for the three  
16 competing signal detection models. The criterion points C1-C5 are the boundaries between  
17 “sure true”/“probably true” (C1) through to “probably false”/“sure false” (C5). Error bars  
18 represent 95% CIs. Misfit distance is in  $d'$  units.  
19  
20  
21  
22  
23  
24

**Figure 5: Experiment 2 observed data and mean fits to semantic and episodic data from competing signal detection models.**

25  
26  
27  
28  
29 The five points at the intersections of their 95% CIs represent the means observed ROC  
30 points. The lines are constructed using the mean parameters recovered for the UEV, DP and  
31 DP-RR signal detection models. Panel A shows fits to the full sample of participants. Panel  
32 B shows fits to the sensitivity-matched subsamples. The left-hand-side of panels shows  
33 semantic data fits, the right-hand-side shows episodic data fits.  
34  
35  
36  
37  
38  
39  
40  
41

**Figure 6: Experiment 2 semantic data misfit scatter plots**

42  
43 Scatterplots showing target and lure discrepancies between the predicted and observed  
44 criterion points for semantic data from Experiment 2. Scatter plots vary according to criterion  
45 point (C1-C5; vertical axis) and model (horizontal axis). The ‘mean’ column plots the  
46 average misfit from each model simultaneously (error bars represent 95% CIs), using the  
47 same colour coding. Target and lure discrepancy are in  $d'$  units.  
48  
49  
50  
51  
52  
53  
54  
55

**Figure 7: Experiment 2 episodic data misfit scatter plots**



## THRESHOLD RETRIEVAL IN SEMANTIC MEMORY

Scatterplots showing target and lure discrepancies between the predicted and observed criterion points for episodic data from Experiment 2. Scatter plots vary according to criterion point (C1-C5; vertical axis) and model (horizontal axis). The 'mean' column plots the average misfit from each model simultaneously (error bars represent 95% CIs), using the same colour coding. Target and lure discrepancy are in  $d'$  units.

**Figure 8: Experiment 2 absolute misfit to semantic and episodic data**

The mean Euclidean distance between predicted and observed criterion points for the three competing signal detection models. Panel A shows mean absolute misfit for semantic data and Panel B for episodic data. Error bars represent 95% CIs. Misfit distance is in  $d'$  units.

**Figure 9: Within-Subjects Semantic and Episodic  $R$  parameters.**

Semantic task  $R$  parameters were recovered using the DP-RR model, episodic  $R$  parameters using the DP model.

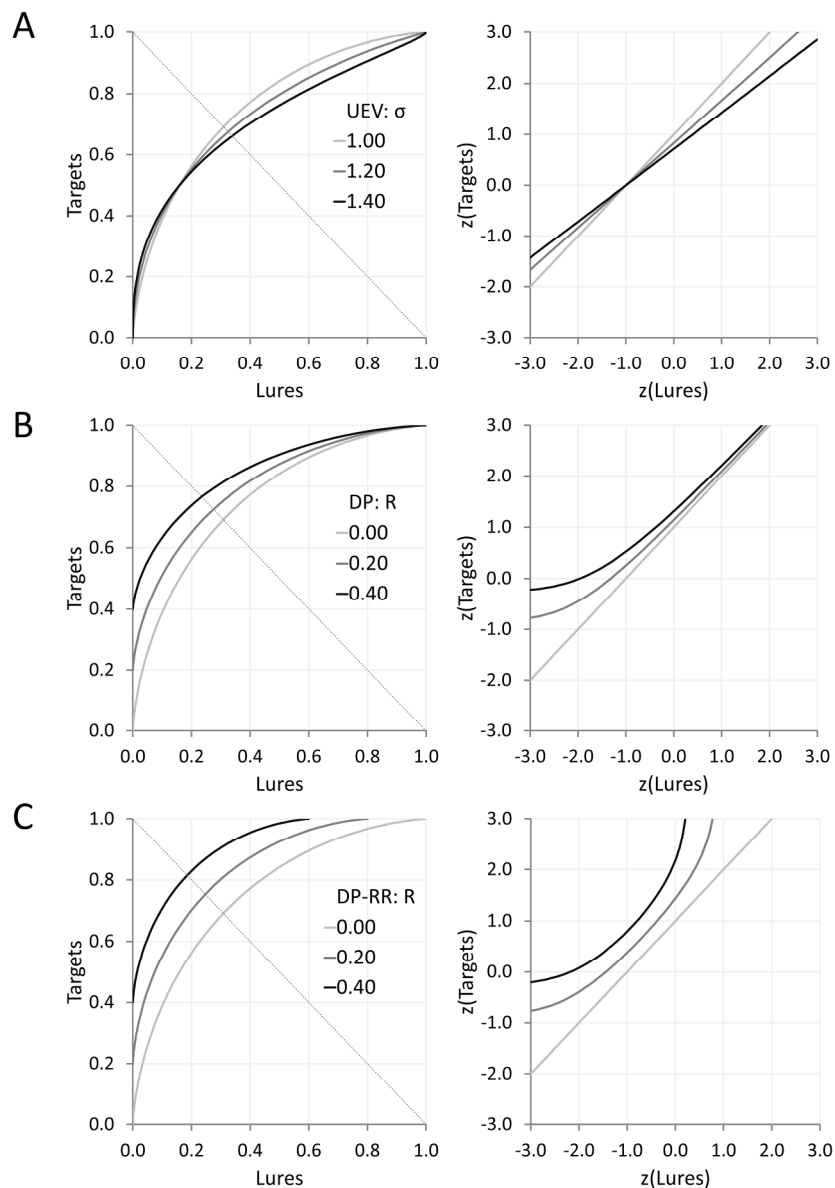


Figure 1: Hypothetical ROCs from three competing signal detection models

Across all models,  $d'$  is held constant at 1. The left-hand-side of panels shows ROCs, the right-hand-side shows zROCs. Panel A: The unequal variance (UEV) signal detection model shows the effect of increasing  $\sigma$  (lighter to darker lines) on ROC curve asymmetry about the diagonal  $y = 1 - x$  (the dotted line that extends from  $[0,1]$  to  $[1,0]$ ). This manifests as an alteration to the gradient and intercept of the zROC lines. Panel B: The dual process (DP) signal detection model shows the effects of increasing  $R$  on ROC curve asymmetry via alteration in the  $y$  axis intercept. This manifests as an increasing curve in the lower aspect of the zROC line. Panel C: the dual process recollect-to-reject (DP-RR) signal detection model shows the effects of increasing  $R$  on both ROC intercepts, with no change in curve symmetry. This manifests as an introduction of two curves, in the lower and upper aspects of the zROC.

190x267mm (300 x 300 DPI)

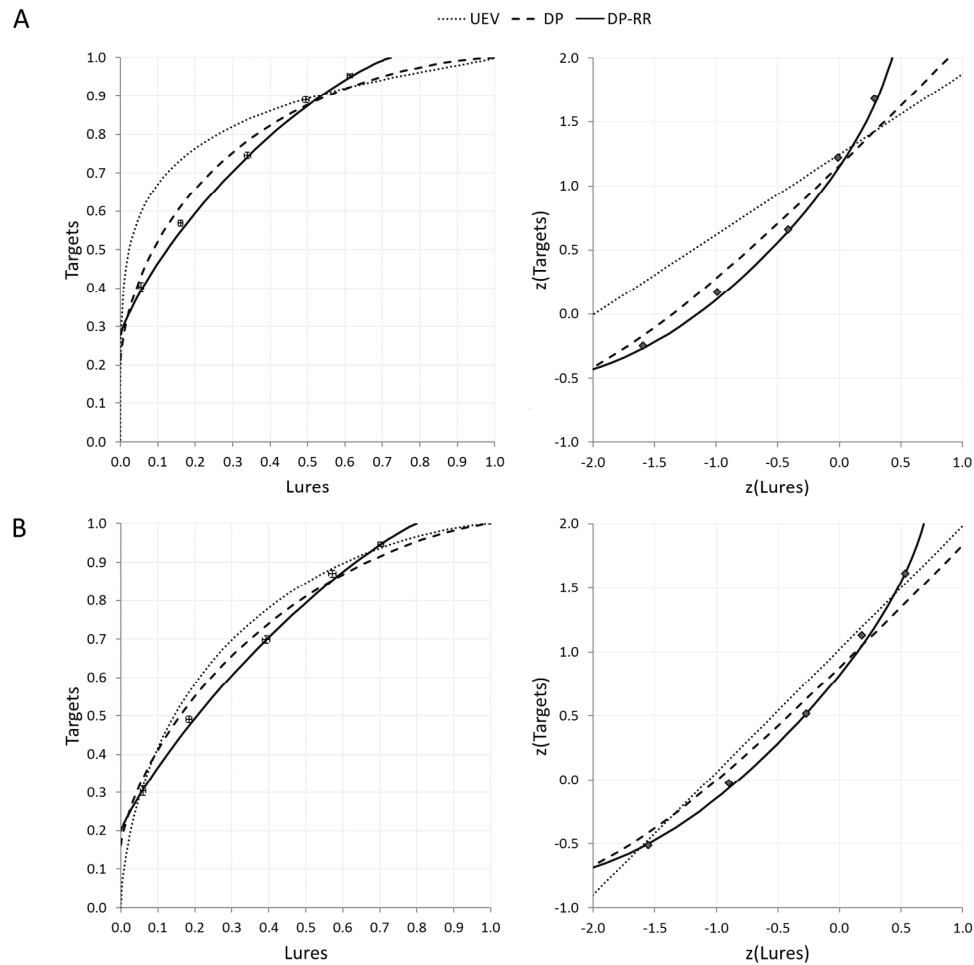


Figure 2: Experiment 1 observed data and mean fits from competing signal detection models. The left-hand-side of panels shows ROCs, the right-hand-side shows zROCs. The five points (shown at the intersection of their respective 95% CIs on the ROC plots) represent the mean observed ROC points. The three lines are constructed using the mean parameters recovered when participants are fit individually, with the UEV model shown as a dotted line, the DP model shown as a dashed line and the DP-RR model shown as a solid line. Panel A shows data and fits from the full dataset. Panel B shows data and fits from the question-selective subset.

218x212mm (220 x 220 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

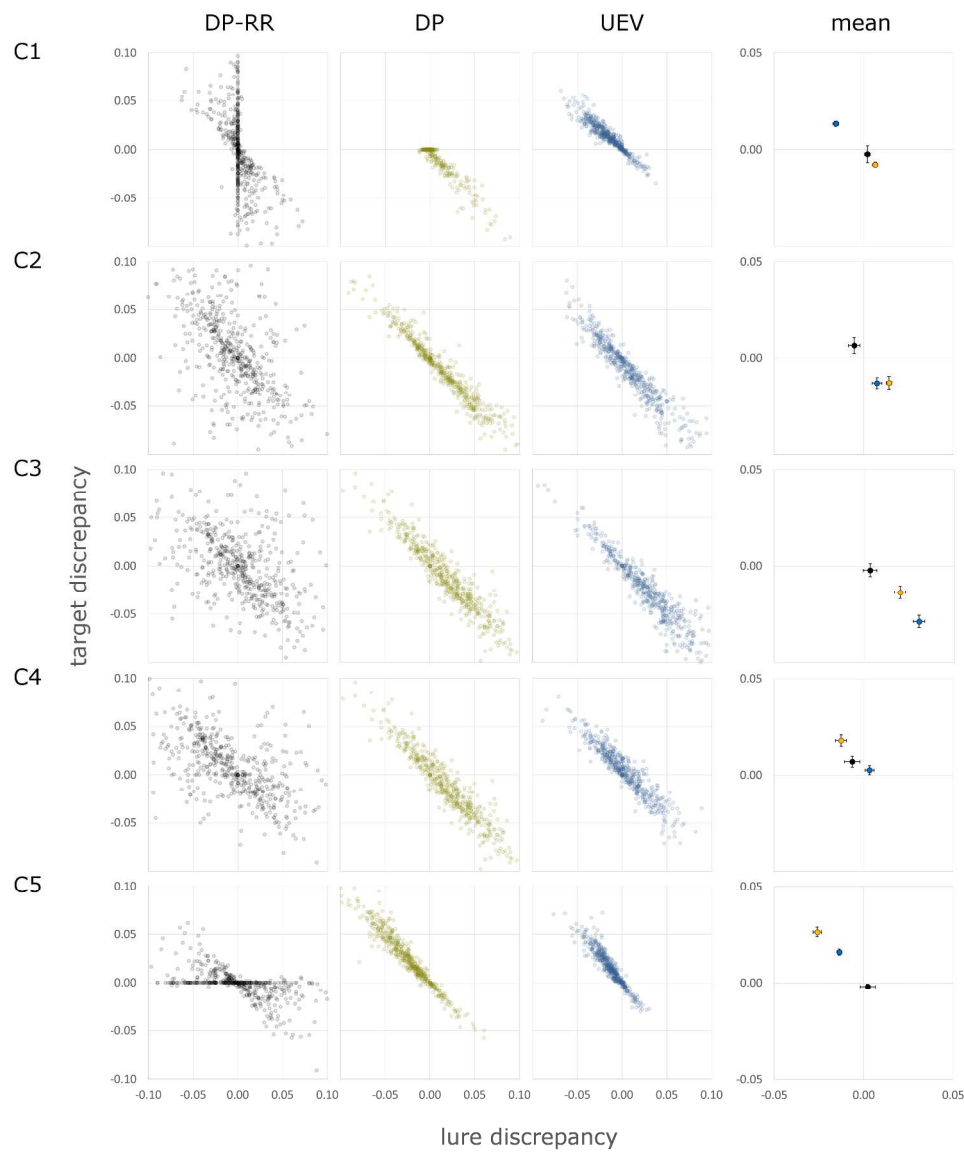


Figure 3: Experiment 1 misfit scatter plots  
 Scatterplots showing target and lure discrepancies between the predicted and observed criterion points. The criterion points C1-C5 are the boundaries between "sure true"/"probably true" (C1) through to "probably false"/"sure false" (C5). The first three columns show misfit for each participant within the DP-RR (black points), DP (yellow points), and UEV (blue points) signal detection models. The 'mean' column plots the average misfit from each model simultaneously (error bars represent 95% CIs), using the same colour coding. Target and lure discrepancy are in  $d'$  units.  
 250x297mm (300 x 300 DPI)

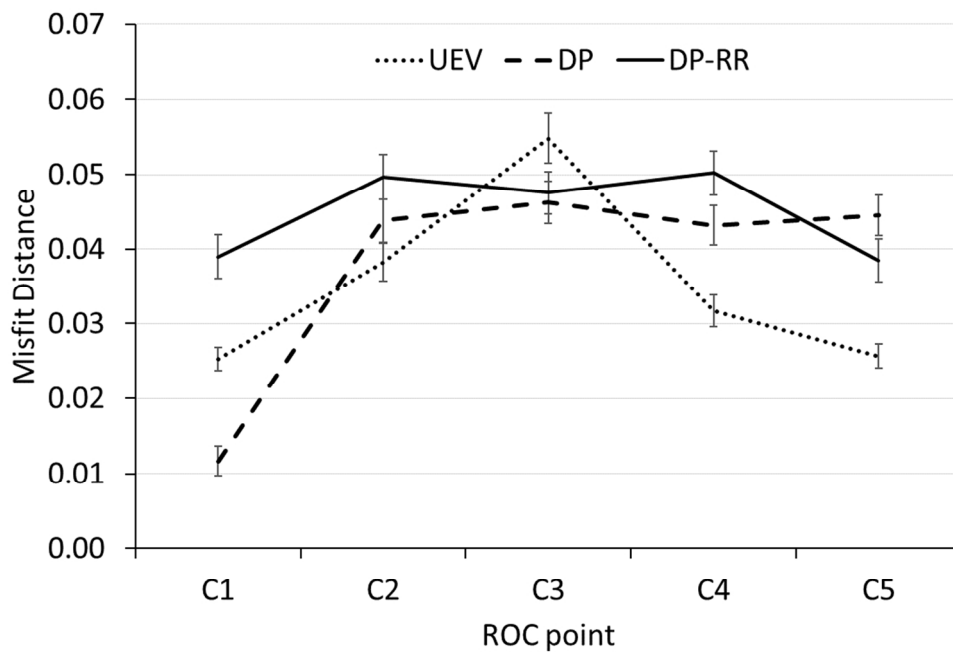


Figure 4: Experiment 1 absolute misfit

The mean Euclidean distance between predicted and observed criterion points for the three competing signal detection models. The criterion points C1-C5 are the boundaries between "sure true"/"probably true" (C1) through to "probably false"/"sure false" (C5). Error bars represent 95% CIs. Misfit distance is in  $d'$  units.  
408x277mm (72 x 72 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

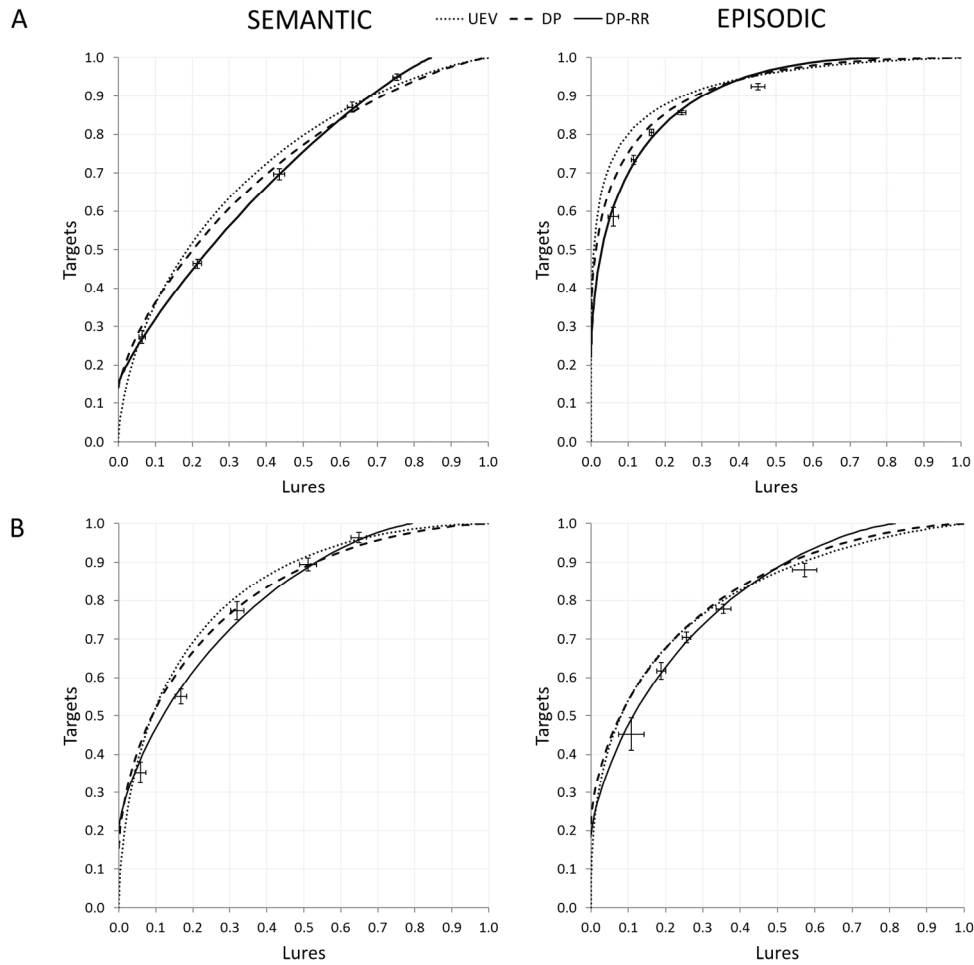


Figure 5: Experiment 2 observed data and mean fits to semantic and episodic data from competing signal detection models.

The five points at the intersections of their 95% CIs represent the means observed ROC points. The lines are constructed using the mean parameters recovered for the UEV, DP and DP-RR signal detection models.

Panel A shows fits to the full sample of participants. Panel B shows fits to the sensitivity-matched subsamples. The left-hand-side of panels shows semantic data fits, the right-hand-side shows episodic data fits.

218x212mm (220 x 220 DPI)

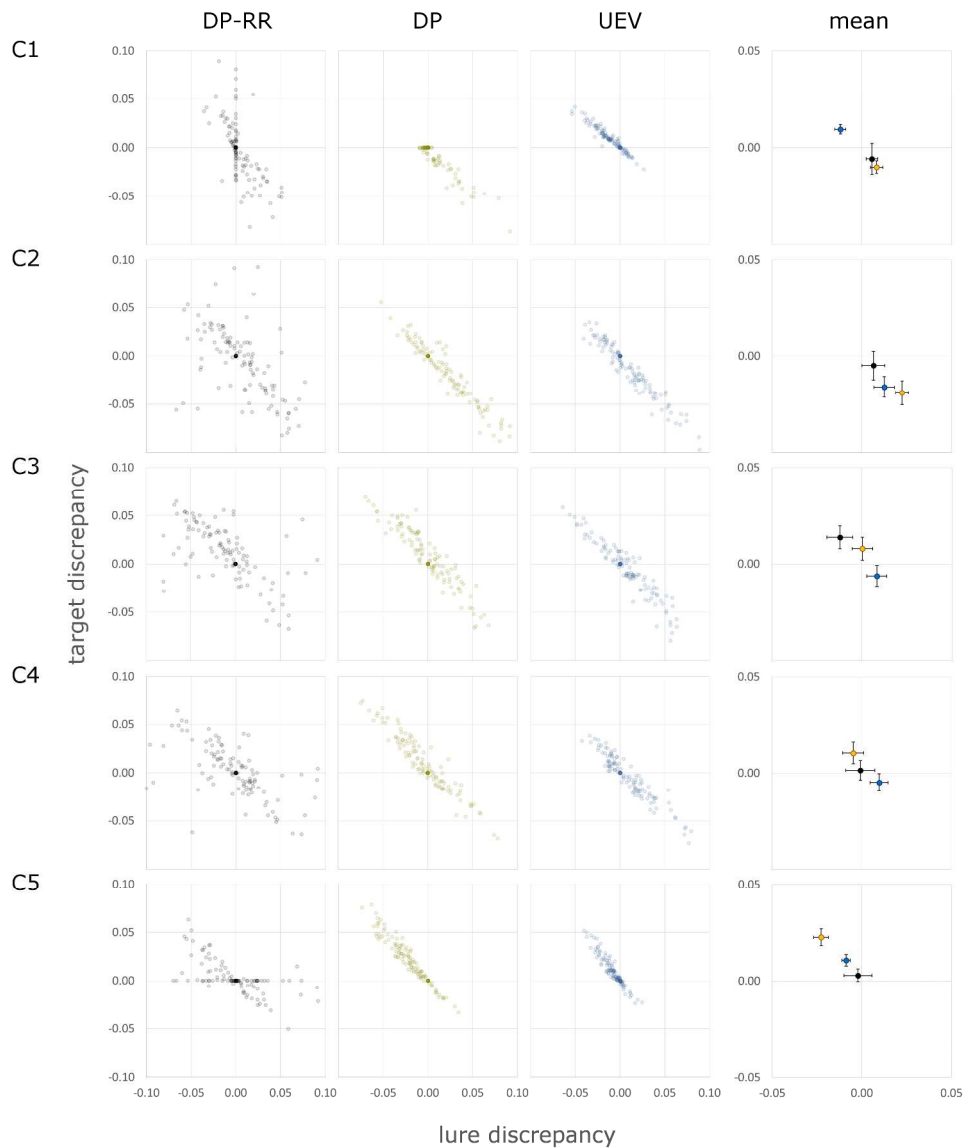


Figure 6: Experiment 2 semantic data misfit scatter plots  
 Scatterplots showing target and lure discrepancies between the predicted and observed criterion points for semantic data from Experiment 2. Scatter plots vary according to criterion point (C1-C5; vertical axis) and model (horizontal axis). The 'mean' column plots the average misfit from each model simultaneously (error bars represent 95% CIs), using the same colour coding. Target and lure discrepancy are in  $d'$  units.  
 250x297mm (300 x 300 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

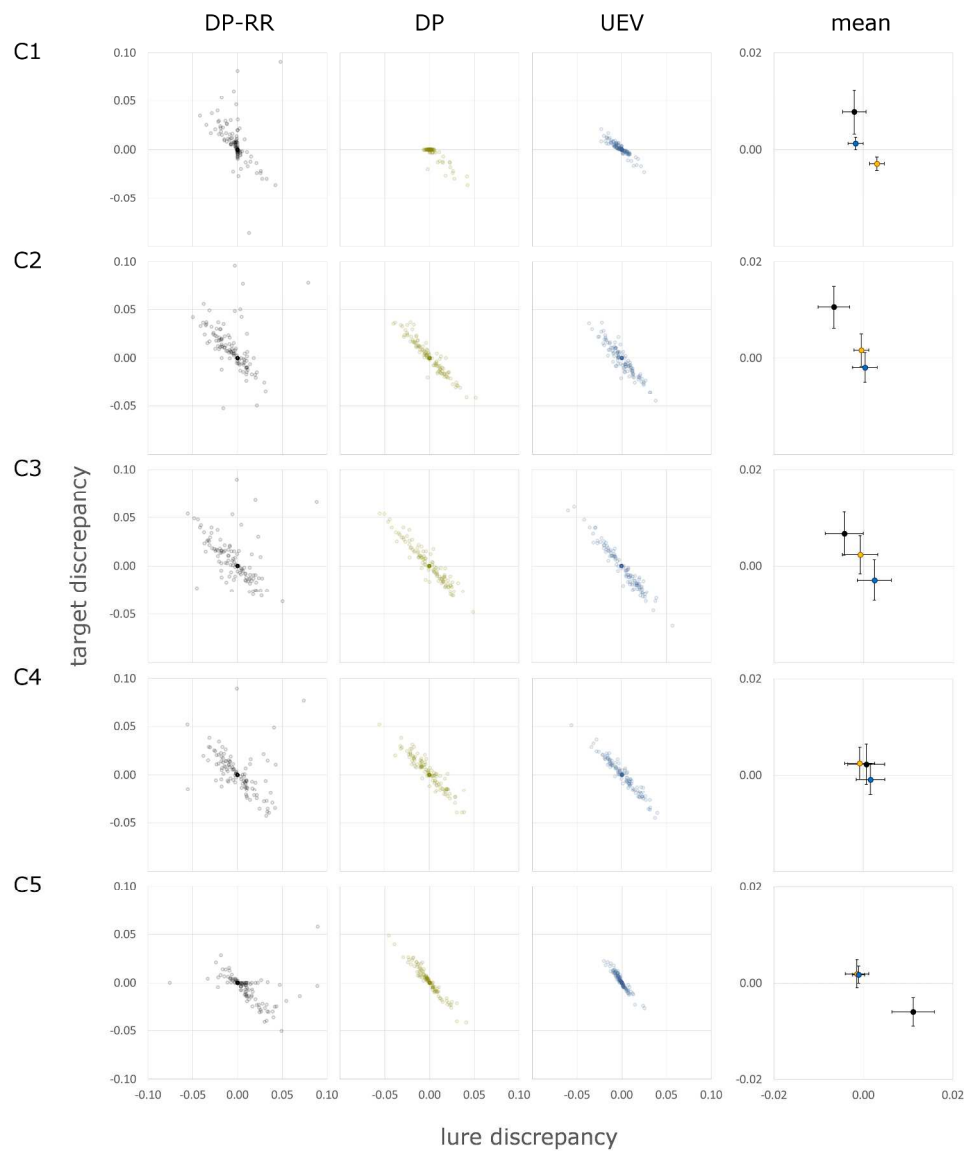


Figure 7: Experiment 2 episodic data misfit scatter plots  
 Scatterplots showing target and lure discrepancies between the predicted and observed criterion points for episodic data from Experiment 2. Scatter plots vary according to criterion point (C1-C5; vertical axis) and model (horizontal axis). The 'mean' column plots the average misfit from each model simultaneously (error bars represent 95% CIs), using the same colour coding. Target and lure discrepancy are in  $d'$  units.  
 250x297mm (300 x 300 DPI)



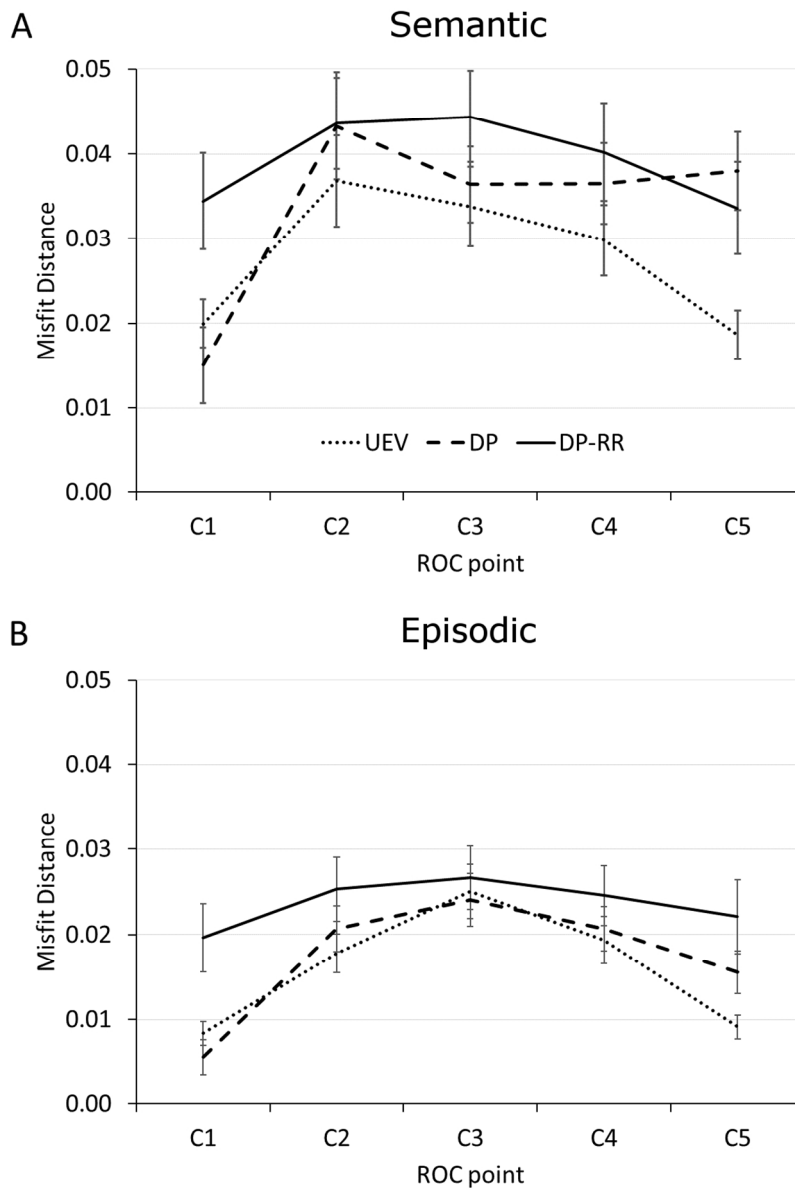


Figure 8: Experiment 2 absolute misfit to semantic and episodic data  
 The mean Euclidean distance between predicted and observed criterion points for the three competing signal detection models. Panel A shows mean absolute misfit for semantic data and Panel B for episodic data. Error bars represent 95% CIs. Misfit distance is in  $d'$  units.  
 135x198mm (220 x 220 DPI)

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

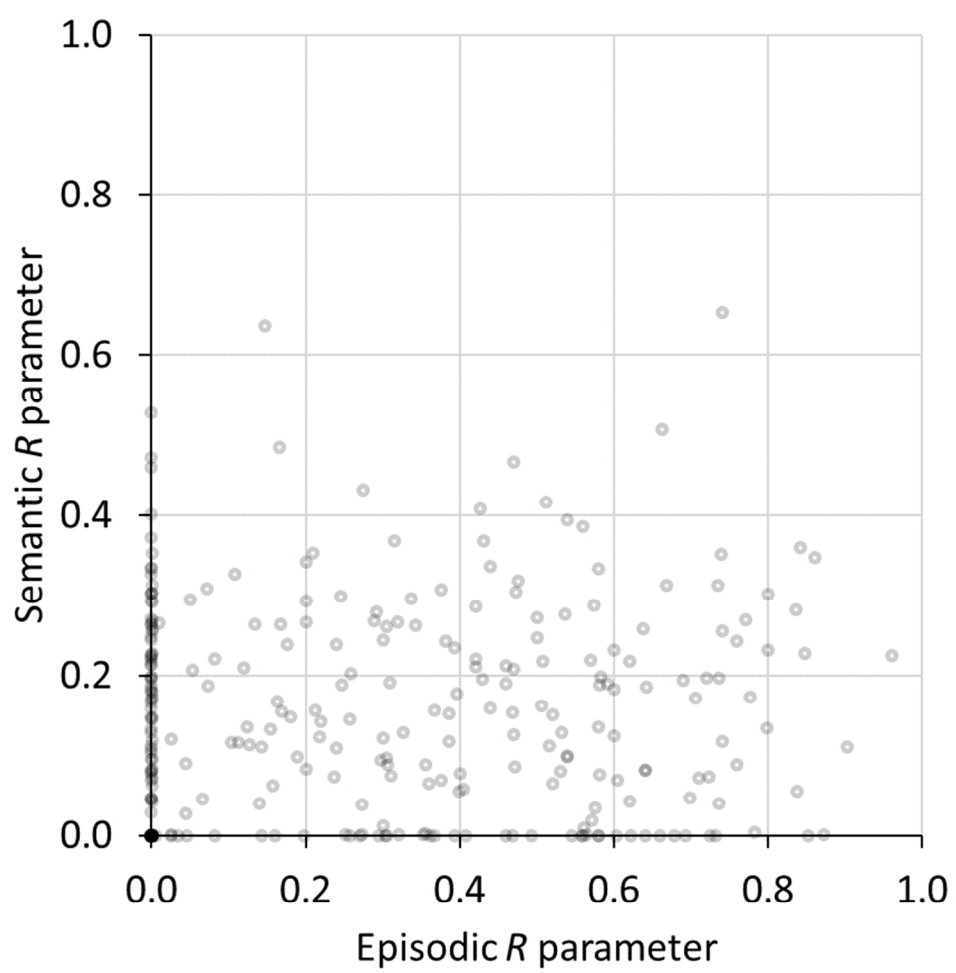


Figure 9: Within-Subjects Semantic and Episodic R parameters.  
Semantic task R parameters were recovered using the DP-RR model, episodic R parameters using the DP model.  
316x317mm (72 x 72 DPI)

### Supplementary Materials A: Semantic Recognition Memory Questions

#	Question	Target	Lure	Exc.
1	Who won a record 8 Grammy Awards in 1984?	Michael Jackson	Bryan Adams	
2	Pan Troglodyte is the Latin name for which creature?	Chimpanzee	Gorilla	
3	Which US state has the nickname 'The Natural State'?	Arkansas	Washington	
4	What is Luke Skywalker's uncle's name?	Owen	Ben	
5	The cricketer, Mudhsuden Singh Panesar, goes by which nickname?	Monty	Spinner	
6	What nationality is film director Baz Luhrmann?	Australian	Canadian	
7	Who wrote the Harry Potter series of books?	J.K. Rowling	J.R.R. Tolkien	X
8	What is the common name for the compound C <sub>6</sub> H <sub>12</sub> O <sub>6</sub> ?	Glucose	Kerosene	
9	Port-au-Prince is the capital of which country?	Haiti	Morocco	
10	When did Bill Clinton first become president of the US?	1993	1990	
11	James Bond first drove an Aston Martin which film?	Goldfinger	Octopussy	
12	What is the first name of English footballer Wayne Rooney's wife?	Coleen	Victoria	
13	What is the common name of Sodium Borate?	Borax	Table Salt	
14	Who has the nickname The Austrian Oak?	Arnold Schwarzenegger	Christoph Waltz	
15	When did Euro coins and banknotes enter circulation?	2002	1998	
16	Antibiotics were first produced from what?	Fungi	Sand	
17	Which Swedish car manufacturer started as aeroplane company?	SAAB	Volvo	
18	What nationality was the painter Gustav Klimt?	Austrian	German	
19	Who shouts "You shall not pass" in the film The Fellowship of the Ring?	Gandalf	Frodo Baggins	X
20	Who sailed from Bristol to North America in 1497?	John Cabot	Christopher Columbus	
21	Where were the 2008 Summer Olympics held?	Beijing	Sydney	X
22	In what film was there a character called Tyler Durden?	Fight Club	American Beauty	X
23	Prospero and Miranda are characters in which play?	The Tempest	A Midsummer Night's Dream	
24	Who wrote the novel 'The Catcher in the Rye'?	J.D. Salinger	Truman Capote	
25	Where is the headquarters of the European Central Bank?	Frankfurt	Strasbourg	
26	According to the UNWTO (2012) rankings, which country is the most popular tourist destination?	France	Great Britain	
27	In which year did Casablanca win an Oscar for best picture?	1944	1946	
28	Laughing gas is a compound of oxygen and which other gas?	Nitrogen	Helium	
29	What was America's first National Park?	Yellowstone	Everglades	
30	What is the national bird emblem of the United States?	Eagle	Owl	X
31	Who wrote the 'Zombie Survival Guide'?	Max Brooks	Richard Matheson	
32	How many 'toes' does a camel have on each foot?	2	3	
33	What is the capital of Belarus?	Minsk	Belgrade	
34	How many players are there in a soccer team?	11	10	
35	Where was the actor Will Smith born?	Philadelphia	Miami	X

Supplementary Materials A: Semantic Recognition Memory Questions

36	Who directed the movie This is Spinal Tap?	Rob Reiner	Terry Gilliam	
37	Who said "Freedom is indivisible, and when one man is enslaved, all are not free"?	Nelson Mandela	Martin Luther King	
38	How many national parks are there in the United Kingdom?	15	12	
39	How many legs does an ant have?	6	4	
40	An early Apple computer was given which girl's name?	Lisa	Sara	
41	Who was the 40th President of the United States?	Ronald Reagan	George Bush	
42	Which Finnish heavy metal music group won the Eurovision song contest in 2006?	Lordi	Axewitch	
43	Which sportsman's nickname was 'The Refrigerator'?	William Perry	Ted Washington	
44	Which French footballer also stars in a Ken Loach film set in Manchester?	Eric Cantona	Zinedine Zidane	
45	In Norse mythology, who is the god of thunder?	Thor	Odin	
46	Ringo Starr replaced which drummer in The Beatles?	Pete Best	Keith Moon	
47	What is the French for 'brain'?	Cerveau	Creneau	
48	The Rasmus and Nightwish are both bands form which country?	Finland	Germany	
49	What was Butch Cassidy's real name?	Robert Leroy Parker	Gerald Cassidy	
50	What was Darth Vader's name before he turned to the dark side?	Anakin Skywalker	Count Dooku	
51	Who played the lead role in the film Castaway?	Tom Hanks	Kevin Costner	
52	Which city is the capital of Spain?	Madrid	Barcelona	
53	Where does the Scottish Parliament sit?	Holyrood	Westminster	
54	Which computer company did Steve Jobs start after leaving Apple?	Next	Dell	
55	Which English King was beheaded in 1649?	Charles I	Charles II	
56	Which of Dickens' novels was partly set in the US?	Martin Chuzzlewit	Little Dorrit	
57	What is the foodstuff which gave its name to the 'pantry'?	Bread	Bacon	
58	How many times have humans landed on the moon?	Six	Five	
59	In which year were the Manchester United football team involved in the Munich Air Crash?	1958	1966	
60	Who wrote the novel 'Robinson Crusoe'?	Daniel Defoe	Jonathan Swift	
61	The Concorde aircraft was a collaboration between the UK and which country?	France	Germany	
62	Which newspaper was the first to carry a crossword puzzle?	New York World	The Times	
63	From where might you browse and borrow books?	Library	Theatre	X
64	What is Avogadro's number approximately equal to?	6.02E+23	1.66E-27	
65	By what process do plants produce oxygen?	Photosynthesis	Respiration	X
66	The distance of a marathon race is	26.2miles (42.2km)	28.8miles (46.4km)	
67	What is the name of the supercontinent that existed from approximately 510 to 180 million years ago?	Pangea	Cambria	
68	Who invented the television?	John Logie Baird	Tim Berners-Lee	
69	How long does it take for the moon to orbit the Earth?	27 days	27 hours	
70	How long is the normal length of play in a game of rugby?	80 minutes	90 minutes	
71	Who killed John Lennon?	Mark Chapman	Jack Ruby	
72	What part of a plant is the potato an example of?	Tuber	Rhizome	

Supplementary Materials A: Semantic Recognition Memory Questions

73	Which member of the Beatles married Yoko Ono?	John Lennon	Paul McCartney	
74	How many states is Germany made up of?	16	20	
75	Regina George is a character in which film?	Mean Girls	John Tucker Must Die	X
76	How many hours are in a day?	24	25	X
77	When did Neil Armstrong first land on the moon?	20th July 1969	20th August 1969	
78	With whom does Anna Karenina commit adultery ?	Vronsky	Oblonsky	
79	Where is Michelangelo's statue David?	Florence	Rome	
80	When is Martin Luther King Jr's Day?	the 3rd Monday of January	the 3rd Thursday of November	
81	If you're a bird, I'm a bird is a quote from which film?	The Notebook	(500) Days of Summer	
82	What is the captain's name in Moby Dick?	Ahab	Ishmael	
83	At which pole do penguins live?	South Pole	North Pole	
84	Who was the third president of the United States of America?	Thomas Jefferson	John Adams	
85	The Berlin Wall fell in which year?	1989	1991	
86	2Pac represented which coast during the hip hop rivalry of the 1990s?	West Coast	East Coast	
87	ABBA won the Eurovision Song Contest in which year?	1974	1980	
88	The chemical symbol Au represents what element in the periodic table?	Gold	Silver	
89	Luna Lovegood was sorted into which house in Harry Potter?	Ravenclaw	Hufflepuff	
90	Apart from white and red, what colour is featured in the flag of the Russian federation?	Blue	Green	
91	What colour do you get from mixing red and blue?	Purple	Green	X
92	How many months in the Gregorian calendar have 31 days?	7	6	
93	Please sir, I want some more. is a quote from which book?	Oliver Twist	Huckleberry Finn	X
94	What is the capital of Ghana?	Accra	Abidjan	
95	Organic chemistry concerns the study of matter that contains which element?	Carbon	Oxygen	
96	Apple's spreadsheet package is called	Numbers	Calc	
97	Mains electricity in the USA is transmitted as what sort of current?	Alternating current (AC)	Direct current (DC)	
98	What Roman numeral represents the number five?	V	X	X
99	What was the first Disney feature film?	Snow White and The Seven Dwarfs	Bambi	
100	What does NASA stand for?	National Aeronautics and Space Administration	National Air and Space Academy	
101	What is the rabbit called in Beatrix Potter's tales?	Peter	James	
102	When did Gutenberg introduce the printing press?	1450s	1540s	
103	Sushi is a delicacy from which country?	Japan	China	X
104	What is Sherlock Holmes' partner called?	Watson	Walton	X
105	How many rings are on the Olympic flag?	5	6	X
106	Where can you go to see old paintings?	Museum	Zoo	X

Supplementary Materials A: Semantic Recognition Memory Questions

107	Foie gras is made from what organ?	Liver	Heart	
108	Which country is associated with haggis?	Scotland	Wales	
109	Where is Pope Francis from?	Argentina	Brazil	
110	Which of the world's longest rivers flows through Kenya?	Nile	Amazon	
111	Fermat's Last Theorem is closely associated with which scientific subject?	Mathematics	Psychology	
112	An important concept in game theory, what concept helps predict outcomes when there are several decision-makers?	Nash Equilibrium	Mensch Equilibrium	
113	What adjective does F. Scott Fitzgerald use in the title of a book to describe its protagonist Jay Gatsby?	Great	Grand	
114	How many letters are in the English alphabet?	26	25	X
115	When was the attack on Pearl Harbour?	7th December 1941	7th December 1945	
116	In what year was Queen Elizabeth II born?	1926	1925	
117	How many players are in a basketball team?	5	7	
118	What season follows spring?	Summer	Winter	X
119	If written out in full, how many zeros are there in a googol?	100	1000	
120	Someone who will not wait happily could be described as:	Impatient	Attentive	X
121	Which member of the bear family eats only bamboo?	Panda	Polar bear	X
122	What is the introduction of material to a person's immune system to prevent disease is called?	Vaccination	Examination	X
123	In the Gregorian calendar, which month only has 28 days?	February	April	X
124	Practical jokes are traditionally celebrated on which day of the year?	1st April	25th December	X
125	What is Latin for beard?	Barba	Bucca	
126	What is the medical name for a belly button?	Umbilicus	Labrum	
127	What is the currency of the Czech Republic?	Koruna	Ruble	
128	Ansel Adams, Willard Van Dyke and Edward Weston founded which photographic collective?	Group f/64	Focus Group	
129	How many chambers is a cow's stomach divided into?	4	2	
130	Which word, encountered in computer security, literally means 'the part of a plant containing grains'?	Shibboleth	Passcode	
131	What is the third letter of the English alphabet?	C	D	X
132	London lies on which river?	The Thames	The Severn	X
133	In which country is the city of Berlin?	Germany	Austria	X
134	In the Muppets, what nationality is the Chef?	Swedish	French	
135	How many legs are there on the Isle of Man flag?	3	2	
136	Which of the two cities, Paris or London, lies the furthest north?	London	Paris	
137	Which of the two actors, Bill Murray or Ben Stiller is the older?	Bill Murray	Ben Stiller	
138	Who directed 'The Life Aquatic'?	Wes Anderson	Michel Gondry	
139	Morrissey was the lead singer of which Manchester band?	The Smiths	Joy Division	
140	The film 'Man on the Moon' is about which American comedian?	Andy Kaufman	John Belushi	
141	What meat is traditionally used in a Shepherd's Pie?	Lamb	Beef	
142	In the Pixar animation, what kind of fish is 'Nemo'?	Clown Fish	Gold Fish	X

Supplementary Materials A: Semantic Recognition Memory Questions

143	Who served as the president of Iraq between 1979 and 2003?	Saddam Hussein	Osama Bin Laden	
144	Which mineral is often referred to as fool's gold?	Pyrite	Bauxite	
145	On which continent is the country Guyana?	South America	Africa	
146	What is a plant with a trunk, branches and leaves usually referred to as?	Tree	Bush	X
147	Wool is most commonly derived from the coat of which animal?	Sheep	Cow	X
148	The Wimbledon tennis tournament takes place in which English city?	London	Liverpool	
149	The Americas Cup is a competition involving which mode of transport?	Boat	Car	
150	Stella Artois is a beer from which country?	Belgium	France	
151	Who wrote the book 'Charlie and the Chocolate Factory'?	Roald Dahl	Enid Blyton	
152	What is the international dialling code for Brazil?	55	54	
153	Which band sang 'Losing My Religion'?	R.E.M.	Radiohead	
154	Who starred as the cross-dressing lead in 'Tootsie'?	Dustin Hoffman	Al Pacino	
155	Who wrote the book 'The Corrections'?	Jonathan Franzen	David Foster Wallace	
156	Which band released the 1977 album 'Rumours'?	Fleetwood Mac	The Eagles	
157	The daffodil is the national flower of which country?	Wales	The Netherlands	
158	The Guardian newspaper was founded in which English city?	Manchester	London	
159	Microsoft has headquarters in which US city?	Redmond	Cupertino	
160	What is the symbol for the chemical element potassium?	K	Na	X
161	The onion is the most widely cultivated species in which plant genus?	Allium	Salvia	
162	The 1904 World's Fair took place in which city?	St Louis	Chicago	
163	What is the boiling point of water in Fahrenheit?	212	232	
164	The drug heroin can be derived from which plant?	Poppy	Hemp	
165	In which country can the Great Pyramid of Giza be found?	Egypt	Turkey	X
166	Lego originated in which country?	Denmark	Germany	
167	Sake is an alcoholic drink originating in which country?	Japan	China	X
168	What term describes the fear of nakedness?	Gymnophobia	Gynophobia	
169	A 9V battery produces what sort of current?	DC	AC	
170	Light is emitted and absorbed in packets called:	Photons	Protons	
171	Who won an Oscar for their acting role in the film '12 Years A Slave'?	Lupita Nyong'o	Chiwitel Ejiofor	
172	Actor Jon Hamm was made famous by his role in which US television series?	Mad Men	Breaking Bad	
173	Who won the 2013 World Series?	Boston Red Sox	New York Yankees	
174	How many days are there in a Gregorian leap year?	366	367	
175	What is the past participle of 'to sit'?	Sat	Sitting	X
176	What is the plural of 'child'?	Children	Childish	X
177	How many old pence were there in a British Shilling?	12	10	
178	'Dinosaur' derives from the Greek words meaning what?	Terrible lizard	Eating beast	
179	On a clear, bright, cloudless day, what colour is the sky?	Blue	White	X
180	In the world of music, what does CD stand for?	Compact Disc	Computer Data	X

Supplementary Materials A: Semantic Recognition Memory Questions

181	A dozen is a grouping of how many?	12	10	X
182	In what year did the micro blogging site Twitter start?	2006	2008	
183	Stanley Kubrick is most well known for being what?	Film Director	Painter	X
184	Who was the first man to set foot on the moon?	Neil Armstrong	Buzz Aldrin	
185	What decade is often described as 'swinging'?	1960s	1970s	
186	What colour is traditionally worn to funerals?	Black	Red	X
187	In Western cultures, shaking someone's hand is usually a form of:	Greeting	Insult	X
188	The device on which you hang your clothes to dry outside is called a:	Washing line	Drying cord	
189	When was smoking in pubs banned in the United Kingdom?	2007	2008	
190	In what year was Justin Bieber born?	1994	1993	
191	Someone who comes from Bristol is called what?	A Bristolian	A Bristain	
192	What is the process which returns water to the atmosphere from the earth's surface?	Evaporation	Transpiration	
193	When fruit is ready to be picked and eaten, it is described as:	Ripe	Rested	X
194	Great Britain is on which continent?	Europe	North America	X
195	In which way would you usually turn a tap to turn it on?	Clockwise	Anti-clockwise	
196	On what side of the road should one drive in France?	The right	The left	
197	'Jaws' was a film about what sort of animal?	A shark	A crocodile	X
198	Who is facebook's CEO?	Mark Zuckerberg	Sheryl Sandberg	
199	Some young women were referred to as Flappers in which decade?	The Roaring Twenties	The Booming Fifties	
200	Holding one's hair back with a hair tie is referred to as a _____ tail.	Pony	Horse	X
201	The members of the lowest caste in India are referred to as?	The Untouchables	The Invisibles	
202	Who exclaims "Out, damned spot!" in the play Macbeth?	Lady Macbeth	Macbeth	
203	What reptiles complete the film title '_____ on a plane'?	Snakes	Turtles	X
204	In 'Game of Thrones' the Stark family's motto is?	Winter is coming.	The lark is calling.	
205	Which e-book reader was designed by Amazon?	Kindle	Kobo	X
206	How many states make up the United States of America?	50	52	
207	What magnitude of handicap is indicative of a good golf player?	Low	High	
208	What pattern consists of horizontal and vertical stripes in multiple colours?	Tartan	Polka Dots	
209	How many complete novels did Jane Austen write?	6	5	
210	The last German Emperor was called?	Wilhelm II	Bismarck	
211	Who is creative director for Chanel?	Karl Lagerfeld	Marc Jacobs	
212	The island of Formentera lies in which island group?	Balearic	Canary	
213	In music, what does the abbreviation DJ stand for?	Disc Jockey	Dancing Jockey	X
214	Snoop Dogg released a reggae album under which alias?	Snoop Lion	Snoop Tiger	
215	Wine is typically made from which fruit?	Grapes	Oranges	X
216	'Zumba' describes what activity?	A dance fitness program	A weight lifting routine	X
217	Proteins consist of chains of:	Amino acids	Fatty acids	X



Supplementary Materials A: Semantic Recognition Memory Questions

218	Who was the ancient Greek god of the underworld?	Hades	Athene	X
219	Which apes are more closely related to humans, bonobos or gorillas?	Bonobos	Gorillas	
220	'Suzi & the 7 D <sup>3</sup> sseldorf' is the name given to:	A nail polish	A film	
221	What does considering a task 'a piece of cake' imply about it?	It is easily accomplished	It is just one of many	X
222	Which university still operating today is the oldest in the world?	The University of Bologna	The University of Oxford	
223	Who won the FIFA World Cup 1998?	France	Brazil	
224	In what city was 'eau de cologne' invented?	Cologne	Paris	
225	In which country would you be most likely to find a Kibbutz?	Israel	Slovenia	
226	Which city is the largest in Western Asia?	Tehran	Baghdad	
227	'Bridget Jones's Diary' is based on which 19th century novel?	Pride and Prejudice	Madame Bovary	
228	Who delivered the shortest acceptance speech at an Oscars ceremony?	Alfred Hitchcock	Joe Pesci	
229	What is the speed of light?	299 792 458 m / s	399 792 458 m / s	
230	What is the Earth's gravitational field strength in m / s <sup>2</sup> ?	9.81	19.81	
231	What is the capital of Canada?	Ottawa	Vancouver	
232	Which of Goethe's works inspired a series of suicides?	The Sorrows of Young Werther	Faust	
233	Water makes up what percentage of a typical cucumber?	90-98%	50-60%	
234	On what does Sleeping Beauty prick her hand before falling into a deep sleep?	A Spindle	A Rose	
235	Which city is considered the birthplace of the stock market?	Amsterdam	London	
236	In 'The Wind in the Willows', what hobby is Mr. Toad obsessed with?	Motoring	Betting	
237	What colour are the eyebrows of the Wood Grouse?	Red	White	
238	What is a hexameter?	A classic metre in literature	A classic measure of length	
239	From what event did the unification of the German language develop?	Publication of Martin Luther's Bible Translation	Ludwig S <sup>3</sup> tterlin's Writing Reform	
240	Which monarch is described as the Sun King?	Louis XIV	Louis XVI	
241	What did Mr. Spock from Star Trek often call unknown but intriguing things?	Fascinating	Interesting	
242	How do the Beatles' lyrics "Ob-la-di, ob-la-da, life goes on" continue?	Bra	La	
243	Where did Napoleon I die?	Saint Helena	Elba	
244	How does the poem with the first line "I eat my peas with" continue?	Honey	Potatoes	
245	How are spectacle lenses for shortsighted people shaped?	Concave	Convex	
246	The Supremes urge their listeners to "stop in the name of" what?	Love	Law	
247	In 'The Adventures of Asterix', why is Obelix never allowed any magic potion?	He fell into a cauldron full of the potion as a child.	He is too overweight.	
248	In a Roman legion, what usually consisted of two	Maniples	Cohors	

Supplementary Materials A: Semantic Recognition Memory Questions

	centuriae?			
249	In what city can the original international prototype of the metre be found?	Paris	London	
250	How do honey bees communicate where to find blossoms to each other?	By performing a waggle dance	Through a distinct buzzing	
251	In 'The Life of Brian', The Pythons' song recommends which life philosophy?	Always look on the bright side of life.	Hey Babe, take a walk on the wild side.	
252	Who was King Arthur's mentor?	The Wizard Merlin	Sir Lancelot	
253	Where did Florence Nightingale, also known as 'The Lady with the Lamp', receive medical training?	D <sup>3</sup> sseldorf	Perth	
254	Zeus, disguised as a bull, was ridden on by which mythological character?	Europa	Aphrodite	
255	In which club were the Beatles first spotted by their manager Brian Epstein?	Cavern Club	Tavern Club	
256	For what did Janis Joplin ask God for in a song recorded in 1970?	A Mercedes Benz	World Peace	
257	Which actress has won the most Academy Awards with a total of four Oscars?	Katharine Hepburn	Audrey Hepburn	
258	What is the nickname of the New York state maximum security prison situated in Ossining?	Sing Sing	The Rock	
259	In 'David Copperfield', Betsey Trotwood constantly chases away which animals from her lawn?	Donkeys	Bees	X
260	According to the bible, which sea did Jesus walk on?	Sea of Galilee	Dead Sea	
261	Which of Mozart's operas is based on the legends of Don Juan?	Don Giovanni	The Marriage of Figaro	
262	Johann Sebastian Bach composed the Toccata in D minor for what instrument?	The Organ	The Violin	
263	What is the only Spice Girls film called?	Spice World	Girl Power	
264	What event led to an uproar at the 2003 MTV Video Music Awards?	Britney Spears' and Madonna's kiss	Lady Gaga's meat dress	
265	What is Rihanna's nickname?	RiRi	RiHa	
266	Who is the only politician to have held the position of Prime Minister of France twice?	Jacques Chirac	Nicolas Sarkozy	
267	Who is the biological father of Heidi Klum's firstborn child?	Flavio Briatore	Seal	
268	In 'The Sound of Music', what country is the von Trapp family from?	Austria	Germany	
269	In what year was Julie Andrews born?	1935	1945	
270	What was Victoria Beckham's maiden name?	Adams	Miller	
271	What mythical beast is Scotland's national animal?	Unicorn	Dragon	
272	Which fashion designer duo was found guilty of attempting tax evasion?	Dolce & Gabbana	Victor & Rolf	
273	What single was Madonna's first number one hit in the US?	Like a Virgin	Material Girl	
274	In 'Winnie-the-Pooh', which animal claims to possess a higher intellect than all others?	Owl	Piglet	
275	What was the name of the first non-human ape who learned to communicate using American sign language?	Washoe	Koko	
276	'A Beautiful Mind' is based on the life of which scientist?	John Forbes Nash, Jr.	Richard P. Feynman	

Supplementary Materials A: Semantic Recognition Memory Questions

277	In what city was Albert Einstein born?	Ulm	Princeton	
278	When were bananas first supplied in the UK?	1888	1945	
279	The chemical element with atomic number 110 was named after which German city?	Darmstadt	Wuppertal	
280	What chemical compound makes many plants green?	Chlorophyll	Cellulose	X
281	How heavy is an Ostrich's egg?	Approx. 1.4 kg / 3 lb	Approx. 3 kg / 6.6 lb	
282	How many national flags have ten or more colours?	4	1	
283	Frankly, my dear, I don't give a damn is a line from which film?	Gone With The Wind	The Color Purple	
284	In what year was the first iPod sold?	2001	2004	
285	What does 'tulipomania' refer to?	A period where tulips were traded at extraordinary prices	A mental disorder where individuals are obsessed with tulips	
286	What is a Jigglypuff?	A Pokemon	A type of pastry	
287	Atheophobia is the fear of:	Atheists	Ducks	
288	Which artist took the most expensive photograph to date, Rhein II?	Andreas Gursky	Jeff Wall	X
289	What is K-pop?	A South Korean music genre	A sweet	
290	Who is the eldest member of the Backstreet Boys?	Kevin Richardson	Chris Kirkpatrick	
291	The Birkin bag is from which fashion brand?	Hermès	Louis Vuitton	
292	How does 'Gossip Girl' sign off her messages?	xoxo	xxx	
293	What do young Swedish girls traditionally do during Midsummer's Eve?	Pick seven flowers and jump over seven hedges in silence	Sing an ancient song in a circle with candles in their hands	
294	IKEA-manufactured BILLY is:	A bookshelf	A desk	
295	What is the official language of Andorra?	Catalan	Spanish	
296	Andy Warhol famously painted soup cans from which company?	Campbell's	Heinz	
297	In which city did the first Starbucks open in 1971?	Seattle	San Francisco	
298	In which two countries can you find more sheep than people?	Scotland & New Zealand	Ireland & Poland	
299	What colour usually is a tin of Nivea creme?	Blue	Orange	
300	How much alcohol by volume is typically in Sake?	15%	40%	
301	In what year did UK TSB Bank split off from Lloyds TSB?	2013	1995	
302	What instrument is featured on the coat of arms of Ireland?	A harp	A flute	
303	What animal is featured on the back of the Greek 1Ç coin?	Owl	Bull	
304	From which country do Claddagh rings originate?	Ireland	Scotland	
305	In 'Jane Eyre', where does Mr Rochester live?	Thornfield Hall	Thrushcross Grange	
306	In what village in Yorkshire can you visit the Brontë parsonage?	Haworth	Thornton	
307	What did U.S. President Theodore Roosevelt send around the world to demonstrate power?	The Great White Fleet	Surveillance Aircraft	
308	According to hieroglyphics, what does the Egyptian goddess Isis' name mean?	Throne	Mother	

Supplementary Materials A: Semantic Recognition Memory Questions

309	According to legend, where was King Arthur's sword Excalibur forged?	Avalon	Camelot	
310	The Fourier transform is used to transform signals between which domains?	Time and Frequency	Amplitude and Phase	X
311	What is Stanford University's motto?	"Die Luft der Freiheit weht" - The wind of freedom blows.	"Aien Aristeuein" - Ever To Excel	
312	Which college did Barack Obama first attend before transferring to Columbia University?	Occidental College	University of Dayton	
313	What is gluten?	A protein composite	A sugar	
314	Who was awarded the first Nobel Prize in Physics in 1901?	Wilhelm Röntgen	Max Planck	
315	What was Queen Victoria's opinion on make-up?	She publicly declared it vulgar and improper.	She advocated its use and praised its effects.	
316	During which years was alcohol prohibited in the United States of America?	1920-1933	1918-1928	
317	What does the Japanese word 'origami' translate as?	Folding of paper	Creating beautiful things	
318	Which martial art is taught in the Israeli military?	Krav Maga	Karate	
319	Who produced the soundtrack for the film 'Despicable Me'?	Pharrell Williams	Jay-Z	
320	Which individual has won the most Tony Awards, at 21?	Harold Prince	Tommy Tune	
321	Which film is currently ranked as the highest-grossing animated film of all time?	Frozen	The Lion King	
322	Where are Porsche's headquarters located?	Zuffenhausen, Stuttgart	Untert <sup>3</sup> rkheim, Stuttgart	
323	Bentley, Bugatti, Lamborghini and Porsche all belong to which automotive company?	Volkswagen AG	General Motors	
324	What is a Furby?	An electronic robotic toy	A furry frisbee	X
325	At what age did Kate Moss begin to model?	14	16	
326	People of what religious faith celebrate Diwali?	Hinduism	Buddhism	
327	In what century did the Brothers Grimm publish their first collection of fairy tales?	19th century	17th century	
328	What is 'Bananagrams'?	A word game	A sugary drink	
329	What is the value of the letter X in the English language version of Scrabble?	8	10	
330	From which country does the ice cream brand Häagen-Dazs originate?	USA	Denmark	
331	In what year did Anna Wintour become editor-in-chief of American Vogue?	1988	1999	
332	What is 86 degrees Fahrenheit in degrees Celsius?	30	20	
333	In the Opening Ceremony of the 2012 Summer Olympic Games, how did Queen Elizabeth II supposedly arrive?	By jumping out of a helicopter	In a speedboat driven up the Thames	
334	In the US sitcom Friends, what is the name of their often visited cafe?	Central Perk	Insomnia Cafe	X
335	What is the chemical formula for water?	H2O	O2	X

Supplementary Materials A: Semantic Recognition Memory Questions

336	What is Jordan's official language?	Arabic	French	
337	For how many years had Nelson Mandela been imprisoned before his release?	27	12	
338	Out of 'wheelbarrow' and 'umbrella', which is a playing piece in the game monopoly?	Wheelbarrow	Umbrella	
339	The first astrological sign in the Zodiac, Aries, is represented by which animal?	Ram	Bull	
340	Which originally Californian fast food chain lets customers order 'Animal Style fries' from their secret menu?	In-N-Out Burger	Five Guys	
341	Who sang 'The Final Countdown'?	Europe	David Bowie	
342	What does the Dormouse in 'Alice in Wonderland' struggle with?	Staying awake	Figuring out its birthday	
343	What is South Park character Cartman's first name?	Eric	Stan	
344	Who is the only character in the TV series 'Family Guy' with an English accent?	Stewie	Peter	
345	In which TV series did a whole season turn out to be a character's dream?	Dallas	Bonanza	
346	Who is Jenna Marbles?	A Youtube personality	A singer	
347	Who is the world's largest zip manufacturer?	YKK group	ZIP group	
348	How are Huey, Dewey, and Louie Duck related to Donald Duck?	They are his nephews	They are his sons	
349	In his 1967 single, Louis Armstrong sang about what?	A Wonderful World	Spanish Eyes	
350	In the musical 'West Side Story', what is the female main character called?	Maria	Bess	
351	What do mosses use to reproduce?	Spores	Seeds	
352	Shakespeare's Juliet lived in which city?	Verona	Venice	
353	What did J.F. Kennedy proclaim in German when visiting Berlin in 1963?	"Ich bin ein Berliner."	"Ich bin ein Amerikaner."	
354	Which artist wrote the words "Ceci n'est pas une pipe". (This is not a pipe) in his painting of a pipe?	Magritte	Matisse	
355	What comparison by Irina Dunn is a popular feminist slogan?	"A woman needs a man like a fish needs a bicycle."	"A woman needs a man like a pig needs wings."	
356	What grain is porridge traditionally made from?	Oats	Wheat	
357	What did the Greek philosopher Diogenes of Sinope say about sunlight on meeting Alexander the Great?	"Stand out of my sunlight."	"Sunlight is the best disinfectant"	
358	Which member of the British royal family said "I'd like to be a queen of people's hearts"?	Diana, Princess of Wales	Camilla, Duchess of Cornwall	
359	What is the English translation of the national motto of France?	Freedom, Equality, Brotherhood	Peace, Order, and Good Government	
360	Which philosopher said "I know that I know nothing?"	Socrates	Nietzsche	
361	On what continent are the highest mountains in the world?	Asia	Africa	
362	Which parent carries the fertilised eggs in seahorses?	Father	Mother	
363	What is a cat thought to show when twitching its tail from side to side?	Aggression	Joy	
364	What is a dog thought to show when wagging its tail?	Joy or friendliness	Aggression	X
365	What is the capital of France?	Paris	Madrid	X
366	By process of brood parasitisation, the reed warbler is	Common cuckoo	Black-headed duck	

Supplementary Materials A: Semantic Recognition Memory Questions

	known to rear the young of which other species?			
367	In Scotland, on what day of the year is St Andrew's Day celebrated?	30th November	17th March	
368	What did Saint Martin of Tours share with a poor man?	His coat	His bread	
369	According to the Rolling Stones' song, Mick Jagger can't get no ... what?	Satisfaction	Happiness	X
370	Out of 'The Wall' and 'The Mall', which is an album by Pink Floyd?	The Wall	The Mall	X
371	What book was published by Adolf Hitler in 1925?	Mein Kampf	My New Order	X
372	With which slogan did Barack Obama win his first presidential election?	Yes, we can.	Here I am.	X
373	What is a PB&J?	A peanut butter and jelly sandwich	A Peruvian Airline	X
374	What is the fast mouse from 'The Looney Tunes Show' called?	Speedy Gonzales	Roadrunner	
375	What are all three of 'the warrior', 'downward-facing dog' and 'the plough'?	Yoga poses	Constellations	
376	What were the names given to the atomic bombs detonated over Nagasaki and Hiroshima?	Fat Man and Little Boy	Fat Boy and Little Girl	
377	What animal is on the logo of clothing shop Brooks Brothers?	A sheep	A fox	
378	What is 3 x 3?	9	8	X
379	Into which part of the body is a needle inserted during a lumbar puncture?	Spine	Ear	
380	Which of the members of Monty Python also wrote the sitcom 'Fawlty Towers'?	John Cleese	Michael Palin	
381	Which British television provider does not screen any advertisements?	BBC	ITV	
382	Who won the Nobel Prize for literature in 2005?	Harold Pinter	José Saramago	
383	In what year did Albert Einstein win a Nobel prize?	1921	1935	
384	Who wrote 'Adventures of Huckleberry Finn'?	Mark Twain	Charles Dickens	X
385	When did the American Civil War begin?	1861	1865	
386	What is the first name of the Argentinian short-story writer Borges?	Jorge	Gregoire	
387	What is the largest city by population in Switzerland?	Zurich	Geneva	
388	Water from the river Rhone in Switzerland will eventually end up in which sea?	The Mediterranean	The North Sea	
389	The world's first ever international soccer match was held in 1872 between which two teams?	England and Scotland	Brazil and Uruguay	
390	Who sang and recorded the words "You may say I'm a dreamer, but I'm not the only one..." in 1971?	John Lennon	Paul McCartney	
391	What does to 'Xerox' something mean?	To copy it	To destroy it	
392	Which American president gave his name to a dam on the Colorado river?	Hoover	Harrison	X
393	In which corner of an envelope are postage stamps typically placed?	Top right	Top left	X
394	In which direction do you read Hebrew text?	From right to left	From left to right	
395	On what side of the road should one drive in Japan?	Left	Right	
396	How many centimetres are there in a metre?	100	1000	X
397	Fred, Daphne, Velma and Shaggy are characters from	Scooby-Doo	The Simpsons	X

Supplementary Materials A: Semantic Recognition Memory Questions

	which cartoon series?			
398	Which TV series was set in the Stone Age town of Bedrock?	The Flintstones	Stig Of The Dump	X
399	Pall Mall, Viceroy and Woodbine are all brands of what?	Cigarette	Chocolate	
400	In which county in England is the ancient monument Stonehenge?	Wiltshire	Somerset	
401	What material is composed of about 75% silicon dioxide, with sodium oxide, sodium carbonate and lime?	Glass	Coal	
402	Which of the two metals, Cobalt or Silver is magnetic?	Cobalt	Silver	
403	Which fictional device contained a 'flux capacitor'?	The time-travelling car in Back To The Future	The Death Star in Star Wars	
404	In France, Dark Vador is the name given to which fictional character?	Darth Vader	Evil Emperor Zurg	
405	Which company introduced the term 'action figure' in 1964 with their G.I. Joe toy?	Hasbro	Mattel	
406	Which psychologists proposed the Working Memory model in 1974?	Baddeley and Hitch	Craik and Tulving	
407	Who was responsible for the influential 1798 'An Essay on the Principle of Population'?	Malthus	Adam Smith	
408	Whose law describes the force required to compress or extend a spring by a certain distance?	Hooke	Boyle	
409	In his experimental law, Boyle showed that what decreases as the volume of a gas increases?	Pressure	The passing of time	
410	What colour on a traffic light indicates that you may go?	Green	Red	X
411	'Fractal' was a word coined by which Polish born mathematician?	Mandelbrot	Kac	
412	In the traditional colour model, what is the complementary colour of orange?	Blue	Red	
413	What nationality was the painter Renoir?	French	Dutch	
414	What is the past participle of 'to sell'?	Sold	Selled	X
415	Udgang, Izlaz, & Salada are all signs which indicate what?	An Exit	An Entrance	
416	Of the Peanuts characters, who plays the piano?	Schroeder	Linus	
417	In which film does the character played by Marilyn Monroe avoid wearing glasses?	How To Marry A Millionaire	Some Like It Hot	
418	What is the last letter of the Greek alphabet?	Omega	Sigma	
419	Which light wavelengths does chlorophyll absorb?	Red and blue wavelengths	Green and blue wavelengths	
420	'Hur mycket kostar det?' is a question in which language?	Swedish	German	
421	Escoffier's dish of peaches, raspberry sauce and ice cream is known as what?	Peach Melba	Peach Sundae	
422	In most western cultures, how many meals is it customary to eat per day?	3	9	X
423	Which Alfred Hitchcock film was the most profitable?	Psycho	The Birds	
424	In the epic poem 'Nibelungenlied', what does Siegfried bathe in to become invulnerable?	Dragon blood	Unicorn blood	
425	Which novel inspired the idiom 'tilting at windmills'?	Don Quixote	Krabat	
426	Who in 1999 was the first US winner of 'Who Wants to be a Millionaire'?	John Carpenter	Kim Hunt	
427	What has been described by experts as 'culturally valued	General Knowledge	Intelligence	

Supplementary Materials A: Semantic Recognition Memory Questions

	knowledge communicated by a range of non-specialist media'?			
428	Which British quiz show features a spot-lit black chair from which contestants answer specialist and general knowledge questions?	Mastermind	EggHeads	
429	From what is quorn made?	Fungus	Milk	
430	In Madame Bovary, what is the name of Emma's daughter?	Berthe	Sophie	
431	What architectural feature do you walk through to enter a room?	A door	A window	X
432	About how long does fresh milk last in the refrigerator?	About a week	About a month	
433	In the novel 'Sophie's World', what does the first anonymous postcard say?	Who are you?	Who made the world?	
434	How many teams took part in the 2010 Soccer World Cup?	32	24	
435	What is ghee?	Clarified butter	Sweet, spiced tea	
436	The Stanley Cup is contested in which sport?	Ice hockey	Australian Rules Football	
437	What is the character limit for tweets posted on the Twitter social network?	140 characters	160 characters	
438	In computing, what does RAM stand for?	Random Access Memory	Raid Addressed Microprocessor	X
439	Who is widely credited with inventing the telephone?	Alexander Graham Bell	Thomas Edison	X
440	What is the square root of 9?	3	81	X
441	Robert De Niro played which boxer in the 1980 film 'Raging Bull'?	Jake La Motta	Rocky Marciano	
442	What sensory impairment did Beethoven develop in later life?	Deafness	Blindness	
443	What was the name of the IBM supercomputer that played Gary Kasparov at chess in 1996 and 1997?	Deep Blue	Deep Sea	
444	In which year did the September 11th terrorist attacks take place?	2001	2003	X
445	In 'How I Met Your Mother', Season 1 Episode 10, which fruit gave its name to the episode title?	Pineapple	Strawberry	
446	'Highway to Hell' is an album by which band?	AC/DC	Metallica	
447	What is the value of the mathematical constant e, to 2 decimal places?	2.72	3.14	
448	Who wrote the Adventure of Tintin comic book series?	Hergé	Goscinny and Uderzo	
449	What is the puzzle called where the aim is to align a cube so that each side only features one colour?	Rubik's Cube	Magic Box	X
450	The Fifth Amendment to the United States Constitution is often referred to as:	The right to remain silent	The right to bear arms	
451	What is Esperanto?	A language	An alcoholic spirit	
452	What colour top is worn by the points classification leader of the Tour de France?	Green	White	
453	In medicine, what does SSRI stand for?	Selective Serotonin Reuptake Inhibitor	Secondary Somatostatin Receptor Ioniser	
454	Who was Cleopatra's father?	Ptolemy XII Auletes	Caesarion	



Supplementary Materials A: Semantic Recognition Memory Questions

455	Which letter follows 'M' in the English alphabet?	N	L	X
456	Which computer company released the Sonic the Hedgehog video games series?	Sega	Atari	
457	On what day of the Gregorian calendar does the Russian Orthodox Church celebrate Christmas?	7th January	24th December	
458	Who is the author of the children's Adventure novel series featuring Hal and Roger Hunt?	Willard Price	Arthur Ransome	
459	What was the research programme that produced the first atomic bombs during World War II called?	The Manhattan Project	SETI	
460	What is Hopscotch?	A playground game	A drink	
461	What colour are British post boxes?	Red	Blue	
462	Which Coca Cola drink is specifically marketed to men?	Coke Zero	Diet Coke	
463	Who sang 'Barbie Girl'?	Aqua	Fuego	X
464	From which country do the group O-Zone, who sang 'Dragostea din tei', originate?	Moldova	Romania	
465	What is the name of Angelina Jolie's adopted daughter?	Zahara	Shiloh	
466	Which member of the 'Made in Chelsea' cast formerly starred in 'The Hills'?	Stephanie Pratt	Mark-Francis Vandelli	
467	In which dystopian novel was Big Brother said to be watching citizens?	1984	Brave New World	
468	In which city is the stadium known as the 'Bird's Nest Stadium' located?	Beijing	Rio De Janeiro	
469	Mozzarella cheese is traditionally made with the milk of which animal?	Buffalo	Goat	
470	Known as the 'hood' in American English, by what name is this part of a car known in British English?	Bonnet	Boot	
471	What was Charlie Chaplin's middle name?	Spencer	Gladstone	
472	What was One Direction's first studio album called?	Up All Night	Get Lucky	
473	In which ocean are the Seychelles?	Indian	Pacific	
474	Which is the longest river in the world?	Nile	Yangtze	
475	What is the name for a polygon with six sides?	Hexagon	Heptagon	X
476	What is the term for a young hare?	Leveret	Elver	
477	What is the term for a badger's home?	Set	Den	
478	Whirlpool, Sombrero and Andromeda are all:	Galaxies	Bathroom fixture manufacturers	
479	Of what species was the first mammal to be cloned?	Sheep	Mouse	
480	In what country can the peak of Mount Kilimanjaro be found?	Tanzania	Mozambique	
481	Relevant to the brain, what is CSF?	Cerebro-Spinal Fluid	Cortico-Somatic Function	
482	What is the capital of Indonesia?	Jakarta	Bangkok	
483	Pewter is a metal alloy largely made up of which metal?	Tin	Iron	
484	The 'hand of God' describes an infamous soccer incident perpetrated by which player?	Diego Maradona	Eric Cantona	
485	St Patrick is said to have banished which animal from Ireland?	Snake	Bear	
486	In what year was the Treaty of Versailles signed?	1919	1947	
487	What is the capital of New Zealand?	Wellington	Auckland	
488	A man who has lost his hair is said to be:	Bald	Spent	X

Supplementary Materials A: Semantic Recognition Memory Questions

489	Which gland secretes the hormone insulin?	Pancreas	Adrenal	
490	On a violin, what note is the string with the highest pitch?	E	A	
491	Whereas trains runs on tracks, what do cars travel on?	Roads	Rails	X
492	What is the name of the place in which many people gather to watch a game of baseball?	Stadium	Rink	
493	The technology company Philips was founded in which country?	The Netherlands	Great Britain	
494	The Roman numeral D represents what number?	500	50	
495	In what year did the Chernobyl nuclear disaster occur?	1986	1989	
496	Who, in 1841, founded the first modern day travel agency?	Thomas Cook	James Cook	
497	Which Moomin character is mischievous, loves catastrophes and finds untidiness exciting?	Little My	Snork Maiden	
498	Which Jewish festival is also called 'the feast of unleavened bread'?	Passover/Pesach	Yom Kippur	
499	What do vegans not eat?	All animal products and derivatives	Vegetables	X
500	What is the national dish of Portugal?	Dried, Salted Cod	Paella	

Note: Target and Lure indicate true and false response option. Exc. Indicates questions excluded based on the exclusion criteria detailed in the Methods.

## Supplementary Materials B: Comparison of DP-RR and EV Model Fits

The main body of the manuscript compares the fits of the dual process recollect-to-reject (DP-RR) model to the standard dual process (DP) and unequal variance (UEV) signal detection models. Whilst the DP-RR model is newly introduced in the manuscript, the DP and UEV models are widely accepted in the field as providing exceedingly good fits to episodic recognition data (e.g. O'Connor, Guhl, Cox & Dobbins, 2011). However, the way in which these established models generate ROC is arguably disadvantaged by the fact that alterations to their parameters yield asymmetries in ROCs not evident in the semantic recognition data. It may be that a signal detection model constrained to produce symmetrical ROCs provides stronger competition for the DP-RR model.

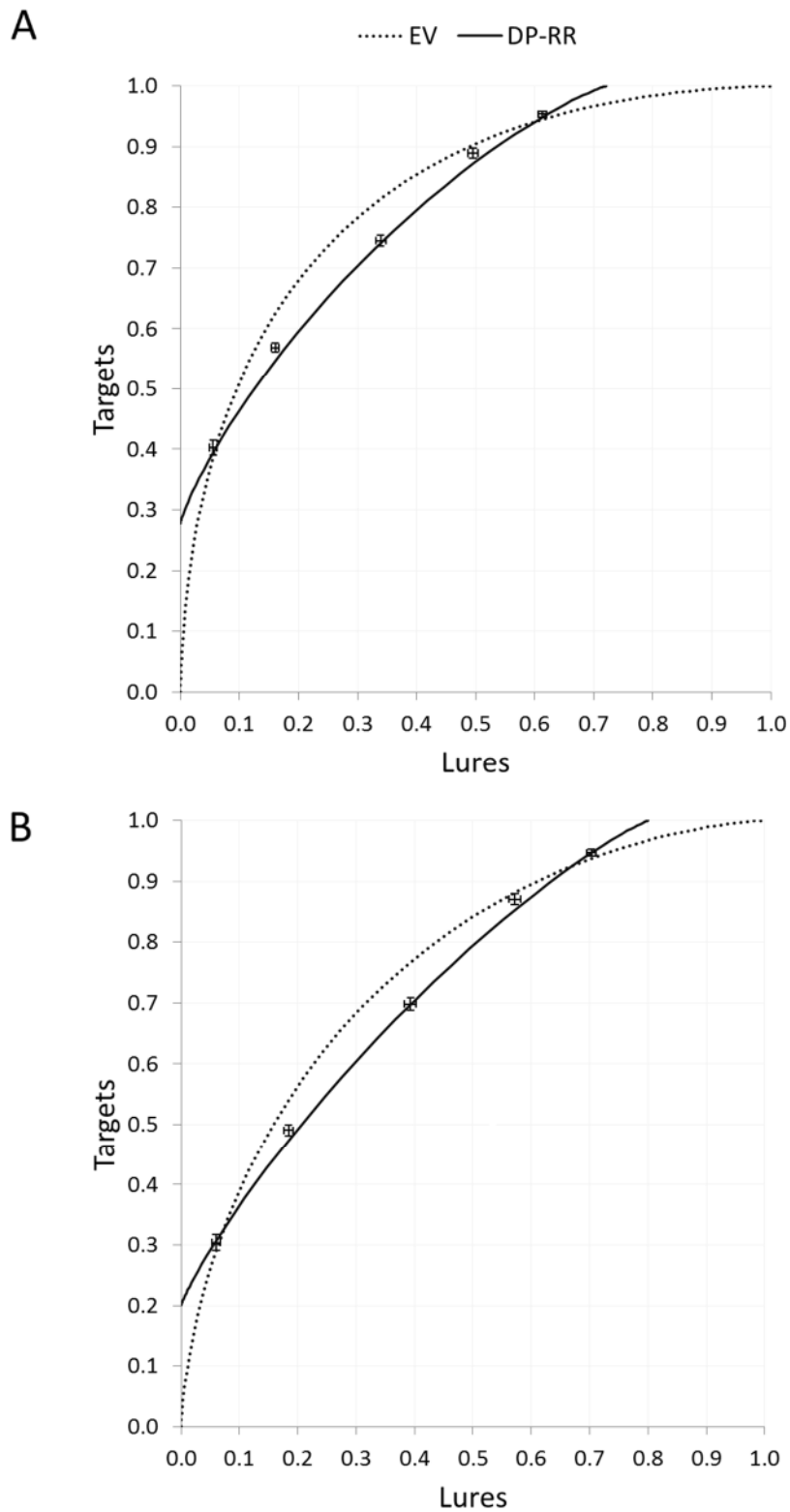
One such model is the equal variance (EV) signal detection model – a model identical to the UEV model, but with the variance of the target distribution held constant at the same value of lure distribution variance. The EV model is a fundamental, minimally complex signal detection model, and parameterises recognition performance with only one distribution parameter,  $d'$  – the distance between the target and lure distribution means. When fitting an EV model ROC to data based on 6 degrees of confidence, 6 parameters can vary: 5 criterion parameters (one for each confidence boundary) and  $d'$ .

When comparing models with differing parameter numbers, the model with the greater number of parameters has the most freedom with which to accommodate any given set of observations. Given the 6 parameters of the EV and 7 of the DP-RR model, log-likelihood ( $LL$ ) fits would therefore tend to favour the DP-RR model. Methods by which the relative quality of models with unequal parameter numbers include the Akaike Information Criterion (AIC; Akaike, 1974). AIC penalises models for having more parameters, meaning that it is theoretically possible to use AIC to compare model fits whilst controlling for the inherent benefit accorded to models with more parameters. Whilst there are no hard and fast rules relating the magnitude of AIC difference to the degree to which one model is superior to another, Burnham and Anderson (2002) have proposed that a minimum difference of 2 AIC units is required before the model returning the lower AIC value can be considered better fitting. In all eventualities, but especially in the absence of such a prominent AIC difference, it would be wise to consider the psychological plausibility of any model being fit to observed data. In the following analyses, we present a comparison of DP-RR and EV model fits, using both  $LL$  and AIC statistics, to the data presented in the manuscript.

### EXPERIMENT 1

Across all experiments, we fit data to the models in question using the procedures described in the manuscript. This holds for the EV model. As such, EV parameters and fit statistics described in this document result from a MLE fitting procedure analogous to that used for the DP-RR and other competing models.

Figure S1 shows the averaged ROCs for the full and question-selective data subsets. Whilst both models are constrained to fit only symmetric ROCs, the EV model's constraint to start at the origin  $[0, 0]$  and end at  $[1, 1]$  causes a prominent ROC overshoot, as was evident in the similarly constrained UEV model's average ROCs (Figures 2 and 5).



**Figure S1: Experiment 1 observed data and mean fits from EV and DP-RR models.**

The five points shown at the intersection of their respective 95% CIs on the ROC plots represent the mean observed ROC points. The two lines are constructed using the mean parameters recovered when participants are fit individually, with the EV model shown as a dotted line, and the DP-RR model shown as a solid line. Panel A shows data and fits from the full dataset. Panel B shows data and fits from the question-selective subset.

Turning to fit statistics, as would be expected from a model with fewer parameters, the DP-RR model's mean *LLs* were significantly higher (less negative) than the EV model's mean *LLs* in both the full and question selective data subsets: Full DP-RR -79.69 [-80.79 -78.59], Full EV -81.00 [-82.11 -79.89],  $t(501) = 21.31, p < .001, d = .951$ ; Question-selective DP-RR -67.91 [-68.92 -66.90], Question-selective EV -78.94 [-69.96 -67.93],  $t(499) = 20.08, p < .001, d = .898$ . The superiority of the DP-RR model's overall capacity to accommodate the data echoes its superiority in visual comparisons of Figure S1.

The more important comparison is of the AIC statistics. As was previously described, a lower AIC is indicative of a better-fitting model once the number of parameters have been accounted for. In both data subsets, there were numeric advantages favouring the DP-RR fits over the EV fits, though there was only a significant difference between the means in the full dataset: Full DP-RR 173.39 [171.19 175.59], Full EV 174.00 [171.79 176.21],  $t(501) = -5.01, p < .001, d = .223$ ; Question-selective DP-RR 149.81 [147.79 151.83], Question-selective EV 149.88 [147.85 151.91],  $t(499) = .681, p = .496, d = .030$ . These AIC differences suggest that, even when controlled for the number of parameters used in each model, there is a very slight advantage for the DP-RR model in fitting the semantic data from Experiment 1.

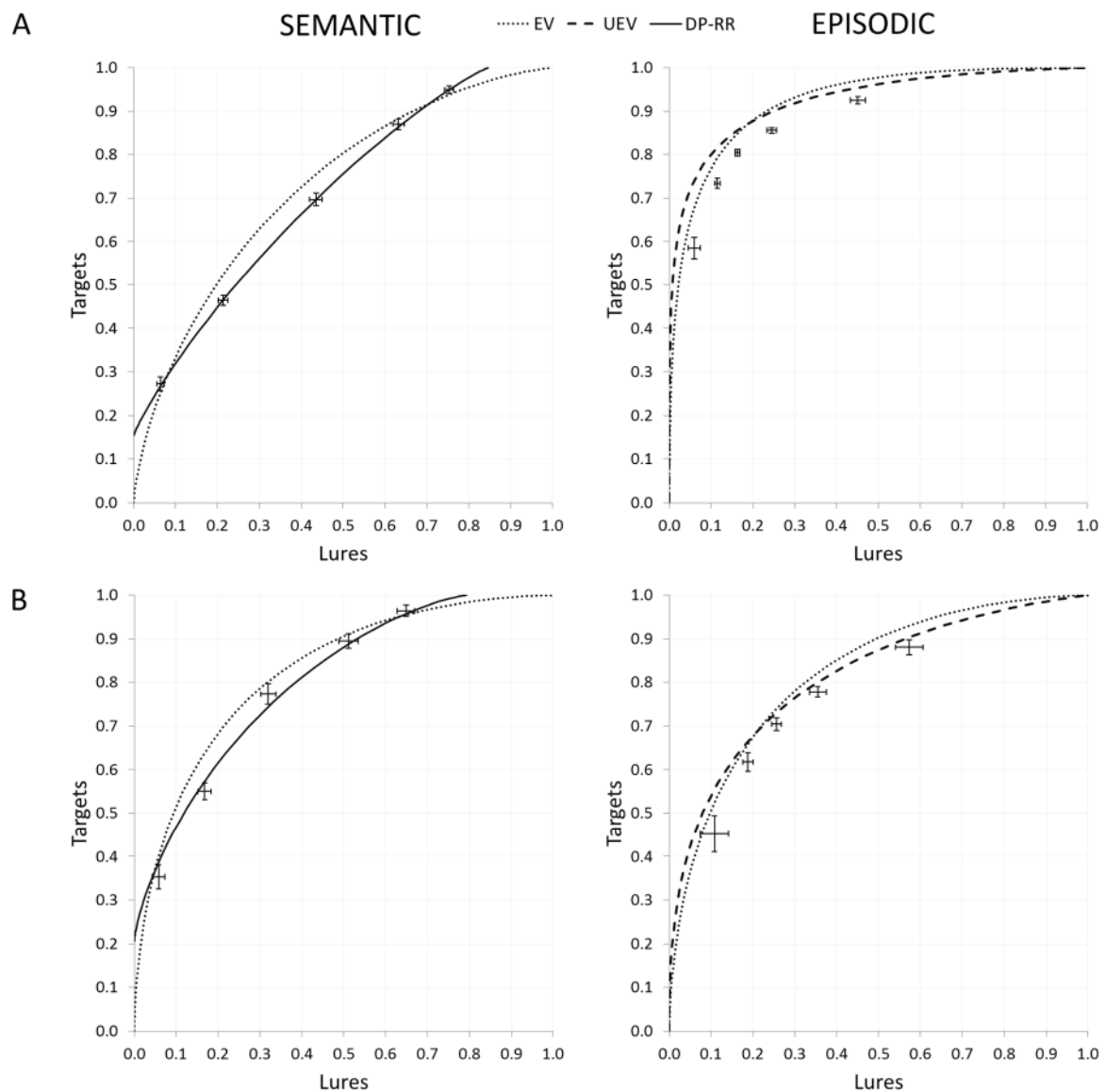
We now turn to the same comparison for the episodic and semantic data obtained in Experiment 2. This is of particular interest given that it provides context for the comparison of the model fits to the novel semantic procedure, in the form of model fits to episodic recognition. In episodic recognition, the unequal variance (UEV) signal detection model is regarded as providing fits far superior to the EV model (e.g. Yonelinas, Dobbins, Szymanski, Dhaliwal, & King, 1996).

## EXPERIMENT 2

Figure S2 shows the averaged ROCs for the semantic and episodic data, in full and sensitivity-matches subsamples.

Across full and sensitivity-matched subsamples, the DP-RR means fits show closer correspondence to the observed semantic data than the EV fits. Interestingly, the fit superiority is less clear-cut for the episodic data, especially for the full sample.

Formal comparison of *LL* statistics in the semantic data found once again that the DP-RR model's mean *LLs* were significantly higher than the EV model's: Full DP-RR -67.13 [-68.45 -65.81], Full EV -68.02 [-69.36 -66.74],  $t(308) = 14.10, p < .001, d = .802$ ; Sensitivity-matched DP-RR -65.93 [-67.81 -64.05], Sensitivity-matched EV -66.63 [-68.54 -64.72],  $t(108) = 8.15, p < .001, d = .780$ . When penalising for the number of parameters using the AIC statistic though, the EV model had the numerical advantage in producing lower means, with a significant difference in the sensitivity-matched subsample: Full DP-RR 148.27 [145.63 150.91], Full EV 148.03 [145.35 150.72],  $t(308) = 1.874, p = .062, d = .107$ ; Sensitivity-matched DP-RR 145.86 [142.10 149.63], Sensitivity-matched EV 145.26 [141.44 149.08],  $t(108) = 3.52, p = .001, d = .337$ .



**Figure S2: Experiment 2 observed data and mean fits to semantic and episodic data from EV, UEV and DP-RR models.**

The five points at the intersections of their 95% CIs represent the means observed ROC points. The lines are constructed using the mean parameters recovered for the DP-RR (semantic data), UEV (episodic data) and EV (semantic and episodic data) signal detection models. Panel A shows fits to the full sample of participants. Panel B shows fits to the sensitivity-matched subsamples. The left-hand-side of panels shows semantic data fits, the right-hand-side shows episodic data fits.

The comparison of the UEV against the EV models in the episodic data yielded significant advantages for the UEV model in *LL* parameters: Full UEV -107.41 [-110.41 -104.13], Full EV -108.16 [-111.45 -104.87],  $t(279) = 12.66$ ,  $p < .001$ ,  $d = .756$ ; Sensitivity-matched UEV -123.96 [-128.66 -119.27], Sensitivity-matched EV -124.53 [-129.24 -119.82],  $t(107) = 8.11$ ,  $p < .001$ ,  $d = .780$ . Notably, the EV model was superior according to AIC statistics, in both full and sensitivity matched samples: Full UEV 228.82 [222.25 235.39], Full EV 228.31 [221.73 234.90],  $t(279) = 4.33$ ,  $p < .001$ ,  $d = .259$ ; Sensitivity-matched UEV 261.93 [252.93 271.33], Sensitivity-matched EV 261.06 [251.65 270.48],  $t(107) = 6.22$ ,  $p < .001$ ,  $d = .599$ .

Taken together these comparisons suggest that DP-RR and UEV, models with a single more parameter than the EV model, fit their respective data better when *LL* statistics are compared. However, the AIC comparisons tend to favour the more sparsely parameterised EV model across both semantic and episodic domains. The lack of an advantage for the UEV model in the episodic domain is noteworthy. Episodic recognition memory researchers have widely discontinued consideration of the EV model as a viable model when compared to UEV (and DP) models. That it continues to show an advantage in the data from Experiment 2 perhaps reflects an over-penalisation of models for additional parameters by the AIC statistic, especially when theoretical considerations that synthesise findings across a range of paradigms are brought to bear on the argument. Beyond this theoretical consideration, quite how much weight should be given to the (sometimes significant) numeric AIC differences of less than one unit across both semantic and episodic domains remains open to interpretation (see Burnham & Anderson, 2002).

In summary, when the semantic data over Experiments 1 and 2 are examined, the DP-RR model does not appear to be compellingly superior to the EV model. However, in the context of the episodic data from Experiment 2 relative to the current position of the recognition memory literature, and the caveat that small differences in AIC should not mask arguments of theoretical superiority, it is not clear that the EV model should be considered alongside the models reviewed in the manuscript as a viable model for recognition memory.

## REFERENCES

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723. <http://dx.doi.org/10.1109/TAC.1974.1100705>
- Burnham, K.P. & Anderson, D. R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed.)*. New York: Springer-Verlag. <http://dx.doi.org/10.1007/b97636>
- O'Connor, A.R., Guhl, E.N., Cox, J.C. & Dobbins, I.G. (2011). Some memories are odder than others: Judgments of episodic oddity violate known decision rules. *Journal of Memory and Language*, 64, 299-315. <http://dx.doi.org/10.1016/j.jml.2011.02.001>
- Yonelinas, A.P., Dobbins, I., Szymanski, M.D., Dhaliwal, H.S., & King, L. (1996). Signal-detection, threshold, and dual-process models of recognition memory: ROCs and conscious recollection. *Consciousness and cognition*, 5(4), 418-441. <http://dx.doi.org/10.1006/ccog.1996.0026>