Linking repetition priming, recognition, and source memory: A single-system signal-detection account☆

Nicholas Lange[a,*], Christopher J. Berry[b], Timothy J. Hollins[b]

[a] Department of Psychology, University of Warwick, United Kingdom
[b] School of Psychology, University of Plymouth, United Kingdom

A B S T R A C T

We present new behavioral data and modeling that links priming, recognition, and source memory. In four experiments, we found that the magnitude of the priming effect, as measured with identification response time in a gradual clarification task, was (1) greater for studied items receiving correct source decisions than incorrect source decisions, and (2) increased as confidence in the source decision increased. Building on the framework for modeling recognition and priming proposed by Berry, Shanks, Speekenbrink, and Henson (2012), we developed a single-system model in which source memory decisions are driven by the same memory strength signal as recognition and priming. We formally compare the model against a multiple-systems model, in which the (implicit) memory signal driving priming is distinct to the (explicit) one driving recognition and source memory. The single-system model reproduces the qualitative patterns of the association between source memory and priming better than the multiple-systems model. Comparison of the quantitative fits was not as clear-cut, however: the single-system model tended to fit better in Experiments 1 and 2, but not in Experiments 3A and 3B, where the observed association between priming and recognition was weaker. Our investigation is an initial attempt at linking priming, recognition, and source memory in the same modelling framework, and provides a basis for further exploration and refinement.

Memory can be expressed in a variety of ways. Three widely studied phenomena are long-term repetition priming, recognition, and source memory. Long-term repetition priming (henceforth priming) refers to a change in identification, detection or production of an item (e.g., a word), which occurs as a result of prior exposure to the same or similar item. For example, identification latencies for words that have recently been studied tend to be faster than those that have not been recently studied. Recognition memory refers to the ability to judge whether an item has been encountered before in a particular context (e.g., a study phase). Source memory refers to the ability to retrieve specific contextual details associated with an item's presentation in the study phase, such as whether an item was previously presented in a study phase towards the bottom or top of a screen. Prominent theories explain these phenomena as being driven by distinct memory systems, signals or processes. For example, under some theoretical accounts, priming is driven by an implicit (unconscious or nondeclarative) memory system, whereas recognition and source memory are driven by a functionally and neurally distinct explicit (conscious or declarative) memory system (e.g., Squire, 1994, 2004, 2009; Squire & Dede, 2015; Tulving & Schacter, 1990).

Key evidence typically cited in support of multiple systems of memory shows that levels of processing manipulations affect recognition memory but not priming (e.g. Jacoby & Dallas, 1981) while changes in modality between study and test affect priming but not recognition memory (e.g. Craik, Moscovitch, & McDowd, 1994). Further, priming is associated with changes in cortical activity (e.g. Schacter, Wig, & Stevens, 2007) while explicit memory is associated with medial temporal lobe activity (e.g. Staresina, Duncan, & Davachi, 2011). Indeed, patients with hippocampal damage typically show intact priming but deficits in explicit memory tasks (e.g. Squire, 2009). Though evidence challenging these dissociations has been reported for functional dissociations (e.g. Buchner & Wippich, 2000; Dunn, 2003; Lukatela, Moreno, Eaton, & Turvey, 2007; Meier, Theiler, Burgi, & Perrig, 2009; Mulligan & Osborn, 2009; Ostergaard, 1992; Poldrack, 1996) and neural dissociations (e.g. Berry, Kessels, Wester, & Shanks, 2014; Addante, 2015; Thakral, Kensinger, & Slotnick, 2016), the multiple systems account of memory is pervasive in psychology textbooks as the default model of memory citing dissociative evidence (e.g. Baddeley,

Eysenck, & Anderson, 2014) and independent memory systems are still used to explain differential memory performance (e.g. Henson et al., 2016).

Particularly fruitful for theoretical advancement has been the development and testing of formal models of priming, recognition, and source memory. Until relatively recently, however, these three phenomena have tended to be modeled in isolation, rather than conjointly (e.g. Rouder, Ratcliff, & McKoon, 2000; Slotnick, Klein, Dodson, & Shimamura, 2000; Wixted, 2007). This is surprising given the potential that conjoint modeling of diverse phenomena has for constraining theory (e.g. Curtis & Jamieson, 2018; Kinder & Shanks, 2003; Zaki, Nosofsky, Jessup, & Unverzagt, 2003). Where efforts of this kind have been made for priming, recognition and source memory, they have only modeled these phenomena in a pairwise fashion (e.g. Berry, Shanks, Speekenbrink, & Henson, 2012; DeCarlo, 2003; Kinder & Shanks, 2003; Shimamura & Wickens, 2009). Here, for the first time, we conjointly model priming, recognition and source memory using a framework based upon signal detection theory. Our modeling incorporates recent advances in the conjoint modeling of recognition and source (e.g. Hautus, Macmillan, & Rotello, 2008) and modeling of recognition and priming (e.g. Berry et al., 2012). First, we describe previous research concerning the conjoint modeling of priming and recognition, and recognition and source. Next, we present a series of experiments that show that priming, recognition and source memory are linked. Finally, we explore the ability of single- and multiple-systems models to explain these findings. We demonstrate that a single-system model with only one underlying memory strength signal explains the qualitative and quantitative pattern of results better than a multiple-systems model assuming stochastically and functionally independent memory signals underlying priming and source memory. More broadly, our aim in this article is to encourage the modeling of different memory phenomena within a common framework and to provide the groundwork for further model development and exploration.

Priming and recognition

Berry et al. (2012) provided support for the view that priming and recognition can be understood as being driven by the same memory system, rather than distinct memory systems, as has been proposed previously (e.g., Squire, 1994, 2004, 2009; Tulving & Schacter, 1990). For instance, in Experiment 2 of Berry et al. (2012), studied and new words were presented in the test phase using a continuous identification with recognition (CID-R) task—a gradual clarification task in which items emerge from a background mask over time (e.g., Feustel, Shiffrin, & Salasoo, 1983; Stark & McClelland, 2000). On each trial, participants pressed a button when they could identify the word, providing an identification response time (henceforth referred to as identification RT), which formed the basis of the measure of the priming effect. Following identification, participants made a recognition judgment for each item using a six-point rating scale (1 = sure new to 6 = sure old). Identification RTs and recognition judgments were found to be linked in several ways: (1) the mean identification RTs for items judged old were faster than that of items judged new, (2) the priming effect, as measured across all studied items, was greater than the priming effect for items not recognized, and (3) identification RTs tended to decrease as recognition confidence increased. Crucially, this pattern of results was predicted by a formal single-system model. The model Berry et al. presented is based on signal detection theory and assumes that, at test, items are associated with a continuous memory strength signal. The greater this signal for an item, the more likely it is to have a relatively fast identification RT and to be judged old with greater confidence.

Other experimental work has confirmed specific predictions of this model, and the model can also explain the differential effect that some variables such as attention, amnesia, and aging have on recognition and priming (e.g. Berry, Henson, & Shanks, 2006; Berry et al., 2014; Berry et al., 2008a, 2008b; Berry, Shanks, Li, Rains, & Henson, 2010; Berry,

Ward, & Shanks, 2017; Ward, Berry, & Shanks, 2013; see Shanks & Berry, 2012, for a review). Formal model comparisons have also shown that the single-system model tends to outperform a variety of multiple-systems models in which distinct memory signals drive performance in priming and recognition tasks (Berry et al., 2012). However, under some dual-process accounts of recognition memory, two distinct processes (recollection and familiarity) contribute to recognition judgments (e.g. Yonelinas, 2002), with repetition priming argued to be the basis of familiarity (e.g. Jacoby & Dallas, 1981; Mandler, 1980). Here, the faster identification of an item (i.e. priming) is attributed to the previous presentation of an item and leads to a judgment of the item as "old". In other words, the association of both memory tasks observed in Berry et al. (2012) may in fact be due to both priming and recognition memory judgments relying on a shared implicit memory component. In this paper, we therefore extend the single-system account to source memory to test if the association of priming and recognition extends to a memory task that does not similarly rely on implicit memory.

Recognition and source memory

Behavioral and modeling work has investigated if performance in tasks that jointly measure recognition and source memory can be explained by a single memory system or by multiple, analogous to the exploration of an association between priming and recognition. In a typical conjoint recognition and source memory task, participants study items in different contexts or presentation formats (i.e., sources), for example words may be studied at the bottom (Source A) or top (Source B) of the screen. In the test phase, previously studied items are presented, intermixed with novel items. Participants are typically asked to give a recognition confidence rating first followed by a source memory confidence rating. The recognition confidence rating scale ranges from sure-new to sure-old ratings, while source confidence ratings range from sure-Source-A to sure-Source-B ratings. Models fitted to the resulting data include those that assume a single memory signal to underlie performance (e.g. Banks, 2000; DeCarlo, 2003; Glanzer, Hilford, & Kim, 2004; Hautus et al., 2008; Slotnick & Dodson, 2005; Starns, Hicks, Brown, & Martin, 2008; Starns, Pazzaglia, Rotello, Hautus, & Macmillan, 2013; Starns, Rotello, & Hautus, 2014) and those that assume that distinct, independent processes drive performance in both tasks (e.g. Yonelinas, 1994, 1999, 2002; Wixted & Mickes, 2010; Onyper, Zhang, & Howard, 2010). The single-system models, just like the model for priming and recognition (Berry et al., 2012), are based upon signal detection theory. They assume that a single underlying memory signal is expressed differently given the task demands of the recognition and source memory task. A stronger underlying signal would therefore tend to give rise to higher recognition confidence and, for studied items, a greater probability of a correct source decision made with greater confidence. Indeed, the patterns of observed data overall can be accounted for better and more parsimoniously by models implementing such a single system idea than those based on multiple memory systems or processes (for a review, see Rotello, 2017).

Linking priming, recognition, and source memory

If, on the one hand, recognition and priming can be modeled as arising from the same strength signal and, on the other hand, so can recognition and source memory, then this begs the question of whether priming and source memory can be modeled as being driven by the same signal. While an association between priming and recognition could potentially be explained on the basis of a shared implicit component, an association of priming and source memory could not. There are hints that this association exists for priming and source memory. For instance, in Experiment 3 of Berry et al. (2012), participants identified each item in the test phase in the CID-R task and then performed a remember-know task for each item (e.g. Gardiner, 1988). They were instructed to respond *remember* if they could retrieve specific

contextual details associated with the item's presentation at study, or to respond *know* if they thought the item was old but in the absence of remembering specific contextual details. In other words, *remember* responses were required to be based on the same type of information typically associated with successful source responding (e.g., Johnson, Hashtroudi, & Lindsay, 1993). Identification RTs for studied items receiving *remember* judgments were found to be faster than for those receiving *know* judgments, suggesting a link between the information in memory that drives the priming effect and the information in memory that drives *remember* responses. The single-system model explained this finding simply by assuming that the strength signal for items receiving a *remember* response is greater than that of *know* responses (following Donaldson, 1996; Dunn, 2004). Greater strength tends to translate into faster identification RTs. Similar experimental findings have been obtained when word-stem completion and lexical decision tasks were used to measure priming (Sheldon & Moscovitch, 2010).

These results suggest that there may be an association between priming and source memory, and that it could be fruitful to extend the single-system model of priming and recognition to source memory. The main assumption of this extended model is that the same underlying memory strength signal that drives priming and recognition also drives source memory judgments. Recognition and source decisions are modeled as in multidimensional signal detection models of recognition and source memory (e.g., Banks, 2000; DeCarlo, 2003; Hautus et al., 2008; Slotnick & Dodson, 2005; Starns et al., 2008). For a given item, the greater its underlying strength value, the higher its performance in a priming task (e.g., faster identification RT in the CID-R task), the greater the confidence with which it will be classified as old and the greater the confidence with which its source will be correctly classified. Thus, a general prediction is that priming, recognition and source decisions will be associated.

One way of conceptualising this model is that it implements the idea that performance in different memory tasks are not driven by distinct memory signals (arising from, for example, distinct implicit and explicit memory systems) but rather that memory tasks access the same information in memory in different ways. If the same memory signal underlies responding in these memory tasks, this could have far-reaching implications. There is an on-going debate in the ERP literature regarding the origin of the FN400. The debate concerns whether the waveform reflects item familiarity or repetition priming (e.g. Strozak, Bird, Corby, Friskoff, & Curran, 2016; Voss & Federmeier, 2011.). If in fact the same memory signal underlies responding in repetition priming and explicit memory tasks, this debate may concern a false dichotomy.

We will contrast that single-system model with a competing multiple systems model that implements the popular assumption that the process underlying priming (familiarity) is stochastically and functionally independent from the one underlying source memory (recollection). While multiple systems models need not implement stochastic independence (e.g., the MS2 model in Berry et al., 2012; DPSD models in Moran & Goshen-Gottstein, 2015; Pratte & Rouder, 2011), it is nevertheless the case that prominent models and widely used experimental methods do make such assumptions (e.g., Yonelinas, 1994, 2002; the process dissociation procedure, Jacoby, 1991). We will describe both the single-system and the multiple-systems models in detail following the description of the behavioural data.

For the behavioral data, we will extend the paradigm in Berry et al. (2012) to include a source manipulation. Thus, we will adapt the study phase such that each item is associated with a source. Source manipulations used in previous research include, for example, distinguishing words spoken by different speakers (e.g. Starns et al., 2014; Slotnick et al., 2000), words and images (e.g. Onyper et al., 2010), visual and auditory information (e.g. Kurilla, 2011) or different locations of the screen (e.g. Yonelinas, 1999). To ensure the same modality for all three memory tests, we will present the words in the study phase in different locations of the screen. Such a spatial manipulation has been shown to activate the hippocampus (Ross & Slotnick, 2008; Slotnick & Thakral,

2013), with hippocampal activation implicated in the retrieval of episodic features as in source memory tasks (Brown & Aggleton, 2001; Diana, Yonelinas, & Ranganath, 2010). Beyond the addition of the source manipulation, the experimental paradigm will be the same as presented in Berry et al. (2012). Following study, participants will complete the test phase. For each item shown at test, participants will first identify it (to provide an identification RT), then give a recognition confidence rating followed by a source confidence rating (CID-RS task). Note, these experiments do not permit us to address the issue of functional independence of priming and source memory but they allow us to address the issue of stochastic independence. In Experiment 1, we attempted to establish whether priming and source memory are linked in this paradigm. In Experiment 2, we looked to replicate and amplify the key findings of Experiment 1. In Experiment 3, we set out to determine whether the associations we observed in Experiments 1 and 2 critically depend upon obtaining recognition and source ratings immediately after identifying an item. To foreshadow our key findings, the magnitude of the priming effect was consistently related to both recognition and source decisions.

Experiment 1

If priming and source decisions are driven by the same memory signal, as outlined in the above single-system account, then we would expect the magnitude of the priming effect to be greater for items receiving correct source decisions than those receiving incorrect decisions. Furthermore, if confidence is a proxy for memory strength, then we would also expect the priming effect to vary with the confidence with which the source decision is made. We investigated this in Experiment 1. In the study phase, words were presented below or above a central fixation point. In the test phase, each word was presented in a CID procedure. Participants identified each word as early in the procedure that they could, providing an identification RT—the measure of priming. They then provided a recognition rating to the word, followed by a source memory rating. This design therefore allowed us to determine whether priming is associated with source memory, whilst simultaneously determining whether the associations between priming and recognition reported by Berry et al. (2012) can be replicated.

Method

Participants

36 individuals (six male; *M* age = 20.06, *SD* = 4.38) took part in the experiment for partial course credit. This sample size provided a power of 0.8 to detect a medium-sized effect in a repeated measures design with two levels (i.e., a Cohen's d_z approximately equal to 0.48). We used the same sample size in each subsequent experiment. Participants in each experiment were recruited using a University of Plymouth participation pool. Ethical approval was gained from a University of Plymouth faculty ethics board.

Materials

The stimulus pool consisted of 384 four-letter low frequency words, selected from the Medical Research Council psycholinguistic database (Coltheart, 1981). The frequency of occurrence ranged from 1 to 13 per million, and there were no concreteness or imageability constraints. Archaic and colloquial terms were excluded. For each participant, 176 words were randomly assigned to be the old stimuli, another 176 words were selected to be the new stimuli, and a further 32 words were selected to be the stimuli appearing on primacy and recency buffer trials in the study phase.

Procedure

At the beginning of the experiment, participants completed six practice trials of the CID procedure (Berry et al., 2012; Feustel et al., 1983; Stark & McClelland, 2000) in order to familiarize themselves with

the task prior to the experimental trials. The CID procedure was the same as that of Berry et al. (2012). On each CID trial a single word was flashed for longer and longer durations, becoming clearer over time. Participants were instructed to press the Enter key as soon as they were sure that they could identify the word correctly. Accuracy and speed were emphasized in the task instructions. At the start of each trial a fixation mask "####" was presented in 24-point Courier font for 1000 ms. Next, the word was presented in 20-point Courier font for 16.7 ms (one screen refresh at 60 Hz). The mask was then presented for 233.3 ms, forming a 250 ms presentation block. There were thirty 250 ms presentation blocks. The stimulus duration increased by 16.7 ms on each alternate block, and the mask was always presented for the remainder of the 250 ms block. Thus, each CID trial was potentially 7500 ms long, but could be terminated prematurely by the participant pressing the Enter key. When the Enter key was pressed, the mask was then re-presented for 16.7 ms. Next, a white outlined box was presented that indicated to the participant that he or she must type the word on the keyboard. Key presses were displayed in the box. Participants were told to press Enter after typing the word to advance to the next trial.

Study phase. Participants were told that they would see words presented below or above the center of the screen for a brief duration and that their task was to remember the location of each word for a later test. Participants completed four study-test blocks, which were identical except that the stimuli in each block were unique. At the start of each study block a "+"-fixation was presented for 500 ms in the center of the screen. The words were presented for 2 s each, with half of them presented 0.9 cm below the central fixation point (i.e., subtending a vertical visual angle of approximately 0.69°, from a viewing distance of approximately 75 cm) and the other half 0.9 cm above the fixation point. The inter-stimulus interval was 100 ms. The assignment of words to the location and the order of presentation was randomized across participants. Participants completed 52 study trials per block, with the first and last four trials in each block designated as primacy and recency buffer trials. The buffer stimuli were not presented in the experiment again.

Test phase. Next, instructions were presented for the first CID-RS test phase. Participants were told that they would again complete identification trials, and that some of the words were from the previous study block and some were novel. They were told that they must decide whether they thought the word was new (i.e., not shown previously) or old (i.e. studied) after each identification, and to indicate whether it was previously shown at the bottom or the top of the screen. They were informed to make that location judgment even for items they indicated were new and to guess if unsure. Participants were told that half of the words would be new and half would be old, and that half of the old words were presented at the bottom of the screen and half were presented at the top. There were 88 trials in each test block, composed of 44 old and 44 new items. On each trial a word was presented in the center of the screen using the same CID procedure as in the practice trials. After participants made their identification, the word was replaced by a recognition probe ("Is the word New or Old?") and a rating scale ("1 = sure new, 2 = probably new, 3 = guess new, 4 = guess old, 5 = probably old, 6 = sure old"). After participants made their recognition judgment a source memory probe was presented ("Was the word presented at the bottom or top?") with a rating scale ("1 = sure bottom, 2 = probably bottom, 3 = guess bottom, 4 = guess top, 5 = probably top, 6 = sure top"). Participants used the number keys 1 through 6 on the main part of a QWERTY keyboard for both judgments. After making their source memory judgment, a prompt was presented instructing participants to press the Enter key to start the next trial. On completion of the test block, participants were presented with the next study block. On completion of the final test block, the experiment terminated.

Initial screening of identification trials

In this experiment and subsequent ones, a trial was not included in the analysis if a word was misidentified during the identification phase of a trial. Identification responses were corrected for minor typographical errors (e.g., where a number or a symbol was typed after the correctly typed word). We excluded trials on which the word was misidentified after correction for typographical errors ($M = 5.57\%$, $SD = 3.76$), no response was given ($M = 0.09\%$, $SD = 0.25$) and on which the identification RT was less than 200 ms or greater than three standard deviations above the mean identification RT (within participant) ($M = 1.15\%$ of trials, $SD = 0.50$). These trials were not analyzed further, following Berry et al. (2012). After exclusions the number of valid trials was $M = 93.19\%$ ($SD = 3.86$, Min $= 82.10\%$).

Measures

All analyses were conducted in R (R Core Team, 2017). For all relevant statistical comparisons, we excluded participants listwise if they had missing data in any cell of that analysis. ANOVAs were calculated using aov_car in the *afex* package (Singmann, Bolker, & Westfall, 2015), with posthoc contrasts calculated with *emmeans* (Lenth, 2016). Degrees of freedom were corrected for violation of sphericity where necessary using the Greenhouse-Geisser correction. An alpha level of 0.05 was used for all statistical analyses and all *t*-tests were two-tailed. We also conducted equivalent Bayesian analyses for all reported frequentist tests using the *BayesFactor* package (Morey & Rouder, 2018), using the package's default default priors for all tests. We report the following effect sizes: η_p^2 for ANOVAs, Cohen's d_z (d_z; mean difference of two dependent measures, divided by the average standard deviation of the difference of the two measures) for *t* tests. Trials were collapsed across study-test blocks for all analyses.

The priming effect was calculated as the mean identification RT for new items minus the mean identification RT for old items. Recognition discrimination was measured with d' (henceforth referred to as *recognition d'*), which is calculated as $z[p(\text{"old"}|\text{old})] - z[p(\text{"old"}|\text{new})]$, where $p(\text{"old"}|\text{old}) = (\text{number of hits} + 0.5)/(\text{number of old items} + 1)$ and $p(\text{"old"}|\text{new}) = (\text{number of false alarms} + 0.5)/(\text{number of new items} + 1)$, following Snodgrass and Corwin (1988). The pattern of results for P_r, which is the measure of discriminability in the two-high threshold model and is calculated as $p(\text{"old"}|\text{old}) - p(\text{"old"}|\text{new})$, was the same, so we only report recognition d' throughout. Recognition response bias was measured with c (henceforth referred to as *recognition c*), which is calculated as $-0.5 * (z[p(\text{"old"}|\text{old})] + z[p(\text{"old"}|\text{new})])$. Source discrimination was also measured with d' (henceforth referred to as *source d'*). For this measure, source-top items were arbitrarily designated as targets and source-bottom items as non-targets; thus, source $d' = z[p(\text{"top"}|\text{top})] - z[p(\text{"top"}|\text{bottom})]$, where $p(\text{"top"}|\text{top}) = (\text{number of correct top responses} + 0.5)/(\text{number of source-top items} + 1)$ and $p(\text{"top"}|\text{bottom}) = (\text{number of incorrect top responses} + 0.5)/(\text{number of source-bottom items} + 1))$. The pattern of results for source accuracy—calculated as (number of "top"|top items + number of "bottom"|bottom items)/number of old items—was the same, so only the former is reported. Source bias was measured with c (henceforth referred to as *source c*) and calculated as $-0.5 * (z[p(\text{"top"}|\text{top})] + z[p(\text{"top"}|\text{bottom})])$.

For the analysis of identification RTs classified according to source confidence ratings, responses were collapsed across source-top and source-bottom items. Source ratings 3, 2 and 1 for source-bottom items and 4, 5, and 6 for source-top items constituted correct source decisions with increasing certainty of response, while source ratings 4, 5 and 6 for source-bottom items and 3, 2 and 1 for source-top items constituted incorrect source decisions.

Reliability of measures. Prior research has shown that it is important to consider the relative reliabilities of direct and indirect memory tasks when comparing task performance (Buchner & Wippich, 2000). Accordingly, split-half correlations were used to determine the

Table 1

Mean identification RTs for new and old items across experiments and mean hit and false alarm rates for the recognition and source memory tasks.

	Identification RT (ms)			Recognition						Source		
	Old items	New items	Priming effect	Hit	False alarm (completely new)	False alarm (partially new)	d' (completely new)	d' (partially new)	Hit	False alarm	d'
Experiment 1											
M	2065	2217	152	0.69	0.40	–	0.80	–	0.59	0.41	0.48
SE	85	83	19	0.03	0.02	–	0.09	–	0.02	0.02	0.08
Experiment 2											
M	1983	2242	259	0.73	0.29	–	1.27	–	0.65	0.35	0.84
SE	73	73	23	0.02	0.02	–	0.11	–	0.02	0.02	0.13
Experiment 3A											
M	1869	2142	273	0.79	0.34	0.60	1.35	0.57	0.64	0.36	0.75
SE	60	66	22	0.02	0.03	0.04	0.11	0.10	0.02	0.02	0.11
Experiment 3B											
M	1980	2230	250	0.70	0.23	0.49	1.38	0.61	0.57	0.43	0.38
SE	73	90	30	0.02	0.03	0.03	0.10	0.09	0.01	0.01	0.07

reliability of the priming, recognition and source measures in all experiments. To calculate these, we first split the data from each participant into odd and even numbered trials and then calculated the priming effect, recognition d' and source d' in each half. The split-half correlations were then given as the Pearson correlation between performance in each half across participants. In Experiment 1, these were large and significant (priming, $r(34) = 0.51$, $p = .001$, $BF = 32$; recognition d', $r(34) = 0.90$, $p < .001$, $BF = 8.61 \times 10^9$; source d', $r(34) = 0.77$, $p < .001$, $BF = 1.85 \times 10^5$), indicating that the measures were highly reliable. The lower reliability of the priming measure relative to the recognition and source measures is consistent with previous findings (e.g., Buchner & Wippich, 2000).

Results

Considering first overall levels of memory performance, the priming effect, recognition d' and source d' all exceeded chance (0) (M priming = 152 ms, $SE = 19$, $t(35) = 8.05$, $p < .001$, $d = 1.34$, $BF = 6.52 \times 10^6$; M recognition $d' = 0.80$, $SE = 0.09$, $t(35) = 8.80$, $p < .001$, $d = 1.47$, $BF = 4.90 \times 10^7$; M source $d' = 0.48$, $SE = 0.08$, $t(35) = 5.96$, $p < .001$, $d = 0.99$, $BF = 1.97 \times 10^4$). Table 1 shows the mean identification RT for new and old items, and also the mean hit rate and false alarm rate for recognition and source decisions. Recognition responding was significantly biased (recognition $c = -0.12$, $SE = 0.06$, $t(35) = 2.15$, $p = .039$, $d = 0.36$, $BF = 1.37$, suggesting overall liberal responding, but source responding was not biased (source $c = 0.03$, $SE = 0.04$), $t(35) = 0.80$, $p = .43$, $d = 0.13$, $BF = 0.24$. Across participants, the priming effect (in ms) was significantly correlated with recognition d', $r(34) = 0.54$, $p < .001$, $BF = 57.96$, and also source d', $r(34) = 0.57$, $p < .001$, $BF = 103.13$. Recognition d' and source d' were also highly positively correlated, $r(34) = 0.84$, $p < .001$, $BF = 1.09 \times 10^7$.

Turning to the relation between priming and source decisions, we analyzed the priming effect associated with correct and incorrect source decisions, regardless of whether those items had been recognized (deemed "old") in the recognition task to test for the association of priming and source memory irrespective of recognition judgments. Two associations were evident. First, the priming effect for items with correct source decisions was significantly greater than for items with incorrect source decisions (M difference = 45 ms, $SE = 14$), $t(35) = 3.13$, $p = .003$, $d = 0.52$, $BF = 10.51$ (see the left-hand side of Fig. 1A). This difference was consistent across individuals, being present in 75% of participants. Second, identification RTs tended to decrease (i.e., the priming effect was greater) as confidence in the source decision increased, as is shown in the right-hand side of Fig. 1A. This trend was confirmed in a 3 (source confidence: guess, probably, sure) × 2 (source

decision: correct, incorrect) repeated measures ANOVA, which yielded a significant main effect of source confidence, $F(1.49, 44.80) = 13.58$, $MSE = 51670$, $p < .001$, $\eta_P^2 = 0.31$, $BF = 4.70 \times 10^4$. Five participants could not be included in this ANOVA because they had zero responses for particular cells of the analysis (hence $N = 31$ for this analysis). The main effect of source decision $F(1, 30) = 0.71$, $MSE = 23682$, $p = .41$, $\eta_P^2 = 0.023$, $BF = 0.19$, and the interaction were not significant, $F(1.40, 42.05) = 2.89$, $MSE = 38313$, $p = .083$, $\eta_P^2 = 0.088$, $BF = 0.65$, though visual inspection of Fig. 1A reveals a trend for RTs to only be faster for correct versus incorrect source decisions when source decisions were made with the highest level of confidence (i.e., "sure" responses).

Next, regarding the relation between the priming effect and recognition decisions, the results reported by Berry et al. (2012, Exp. 2) were replicated. First, identification RTs for items judged old were faster than those judged new. This was the case for old items (i.e., M RT(hit) < M RT(miss)) ($M = 167$ ms, $SE = 28$), $t(35) = 6.07$, $p < .001$, $d = 1.01$, $BF = 2.65 \times 10^4$, and also new items (i.e., M RT(false alarm) < M RT(correct rejection)) ($M = 131$ ms, $SE = 24$), $t(35) = 5.57$, $p < .001$, $d = 0.93$, $BF = 6.54 \times 10^3$. Second, the priming effect for items judged new (i.e., M RT(correct rejection) – M RT(miss)) was significantly greater than chance (0 ms) ($M = 83$ ms, $SE = 23$), $t(35) = 3.62$, $p < .001$, $d = 0.60$, $BF = 33.74$, and, crucially, this priming effect was significantly smaller than the priming effect calculated across all items (M difference = 69 ms, $SE = 20$), $t(35) = 3.46$, $p = .001$, $d = 0.58$, $BF = 22.77$. Third, identification RTs tended to decrease as confidence that an item is old increased (Fig. 2A), as shown by significant linear trends for old stimuli, $t(150) = 6.65$, $p < .001$, and new stimuli, $t(150) = 5.18$, $p < .001$, with higher order relationships not reaching significance for either type of item (all $ps > .09$). Again, five participants could not be included in this ANOVA because they had zero responses for particular cells of the analysis (hence $N = 31$ for this analysis).

Finally, regarding the relation of recognition and source memory, source d' significantly exceeded chance for old items that were correctly recognized (source $d' = 0.60$, $SE = 0.10$, $t(35) = 5.96$, $p < .001$, $d = 0.99$, $BF = 1.97 \times 10^4$, but did not significantly differ from chance for old items that were not correctly recognized (source $d' = 0.09$, $SE = 0.07$), $t(35) = 1.26$, $p = .21$, $d = 0.21$, $BF = 0.37$. Source d' was significantly higher for recognized than unrecognized old stimuli (M difference = 0.51, $SE = 0.13$), $t(35) = 4.08$, $p < .001$, $d = 0.68$, $BF = 109.68$. Analysis of source accuracy led to the same pattern of results.
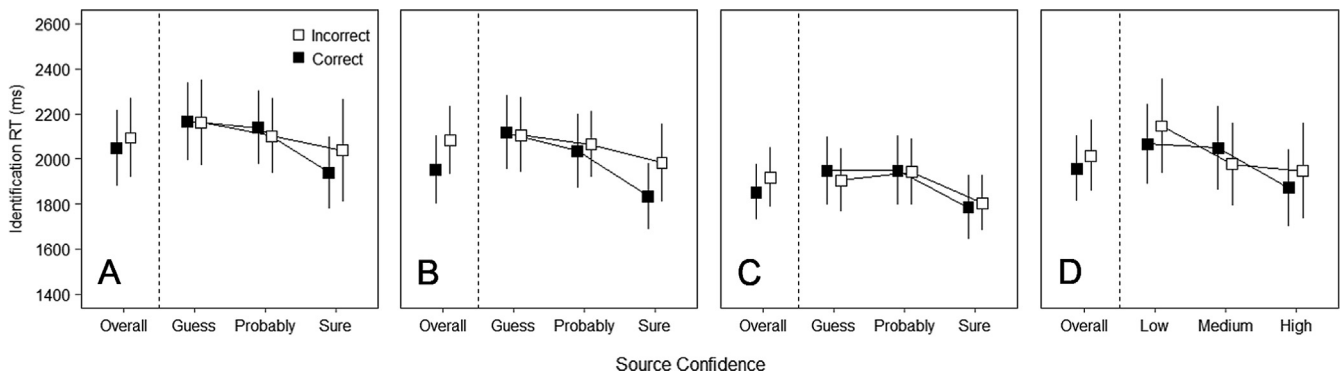
Fig. 1. Identification RTs across source confidence for Experiment 1 (A), Experiment 2 (B), Experiment 3A (C), and Experiment 3B (D) for correct and incorrect source decisions. Error bars are 95% between-subjects confidence intervals of the mean.

Discussion

The main finding from Experiment 1 was that the priming effect varied according to the correctness and confidence of the source decision. Specifically, the priming effect was greater for items with correct versus incorrect source decisions, and the priming effect tended to increase as confidence in the source decision increased. There was a suggestion of an interaction, such that the priming effect for correct source decisions was greatest for high confidence ratings (i.e., "sure" decisions; Fig. 1A), but the Source Decision × Source Confidence interaction suggested this was not reliable. It is possible that the failure to detect this interaction (and the main effect of source decision) is due to high variance in the identification RTs for some of the source responses made infrequently (e.g., incorrect source responses made with high confidence; see Supplemental Materials for details), and we test for this interaction again in the subsequent experiments.

The priming effect also varied according to the recognition decision, replicating key results of Berry et al. (2012), see Fig. 2A. Thus, the magnitude of the priming effect was linked to both recognition and source decisions. An important consideration with this set of results is that, compared with many published studies, source memory in Experiment 1 was relatively low (i.e., M source $d' = 0.48$), and many correct source decisions were guesses or were made with low confidence. Accordingly, in Experiment 2 we examined whether the associations between the priming effect and source decisions would persist with superior source memory.

Experiment 2

In Experiment 2, we aimed to increase source memory by using a shorter study list length than was used in Experiment 1. Previous research has shown that both source discrimination and recognition are superior for items presented in a shorter study list than a longer one (Glanzer et al., 2004; Slotnick et al., 2000). Similarly, priming effects also tend to be greater for items in shorter study lists (e.g., Berry et al., 2006). Accordingly, in Experiment 2, we used half the number of items in the study phase, compared to Experiment 1. In order to keep the total number of old and new stimuli the same as in Experiment 1, participants completed eight, rather than four, study-test cycles. Experiment 2 was otherwise identical to Experiment 1.

Method

Participants

36 individuals (11 male; M age = 21.61, SD = 3.12) took part in the experiment for partial course credit.

Procedure

The materials, experimental set up and measures were identical to Experiment 1. Participants completed eight study-test blocks (compared to four in Experiment 1). Each study block consisted of 26 items, with the first and last two designated as primacy and recency buffer items. In the test phase, participants were shown 22 old items and 22 new items in a random order. On each trial, they first identified the word in the CID phase, and then gave a recognition and a source memory judgement using separate six-point scales.

Initial screening of identification trials

The proportion of misidentified trials, after correction for typographical errors, was M = 4.08% (SD = 2.35). The proportion of trials on which participants did not make a response was M = 0.43% (SD = 0.24). The proportion of trials on which the identification RT
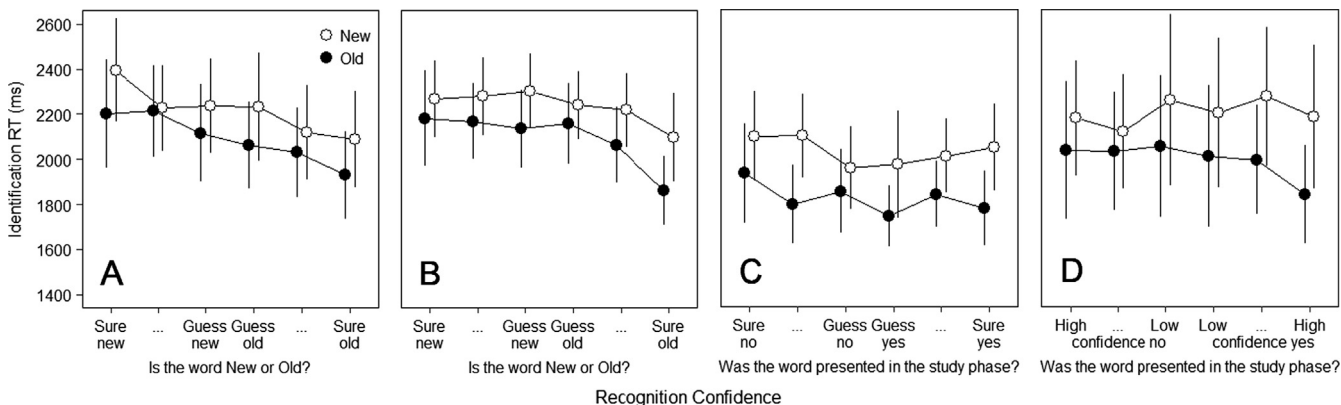

Fig. 2. Identification RTs across recognition confidence for Experiment 1 (A), Experiment 2 (B), Experiment 3A (C) and Experiment 3B (D) for old and new items. Error bars indicate 95% between-subjects confidence intervals of the mean.

was less than 200 ms or greater than three standard deviations above mean identification RT (within-participants) was $M = 1.15\%$ ($SD = 0.45$). After exclusions the number of valid trials was $M = 94.70\%$ ($SD = 2.49$, Min = 86.65%).

Reliability of measures

Once again, the measures of priming, recognition and source memory were reliable (priming effect, $r(34) = 0.59$, $p < .001$, $BF = 181.23$; recognition d', $r(34) = 0.92$, $p < .001$, $BF = 1.91 \times 10^{11}$; source d', $r(34) = 0.90$, $p < .001$, $BF = 6.16 \times 10^{9}$).

Results

As in Experiment 1, the priming effect, recognition d' and source d' all exceeded chance (M priming = 259 ms, $SE = 23$, $t(35) = 11.42$, $p < .001$, $d = 1.90$, $BF = 3.23 \times 10^{10}$; M recognition $d' = 1.27$, $SE = 0.11$, $t(35) = 11.10$, $p < .001$, $d = 1.85$, $BF = 1.54 \times 10^{10}$; M source $d' = 0.84$, $SE = 0.13$, $t(35) = 6.38$, $p < .001$, $d = 1.06$, $BF = 6.39 \times 10^{4}$). Table 1 shows the mean identification RT for new and old items, and also the mean hit rate and false alarm rate for recognition and source decisions. Recognition responding was not significantly biased (recognition $c = -0.03$, $SE = 0.05$), $t(35) = 0.51$, $p = .61$, $d = 0.08$, $BF = 0.20$, nor was source responding (source $c = 0.04$, $SE = 0.04$), $t(35) = 1.23$, $p = .23$, $d = 0.20$, $BF = 0.36$. The magnitude of the priming effect, recognition d' and source d' were significantly greater than in Experiment 1 ($ts > 2.32$, $ps < .024$, $d > 0.61$, $BF > 2.36$), as was intended by the shorter study list length. Across participants, the priming effect was significantly correlated with recognition d', $r(34) = 0.46$, $p = .005$ $BF = 11.14$, and source d', $r(34) = 0.52$, $p = .001$, $BF = 37.01$. Recognition d' and source d' were also significantly correlated, $r(34) = 0.85$, $p < .001$, $BF = 4.80 \times 10^{7}$.

Regarding the association between the priming effect and source decision, as was found in Experiment 1, the priming effect was greater for items with correct source decisions than those with incorrect decisions ($M = 129$ ms, $SE = 28$), $t(35) = 4.62$, $p < .001$, $d = 0.77$, $BF = 470.96$, see the left-hand side of Fig. 1B. Once again, this effect was consistent, being present in 80% of participants. A 3 (Source Confidence) × 2 (Source Decision) repeated measures ANOVA, revealed that, as in Experiment 1, identification RTs decreased (i.e., the priming effect increased) with increasing source confidence, $F(2, 66) = 18.74$, $MSE = 39489$, $p < .001$, $\eta_p^2 = 0.36$, $BF = 2.59 \times 10^{7}$, as is shown in the right-hand side of Fig. 1B. The main effect of source decision was significant though the evidence for this main effect was not substantive, $F(1, 33) = 6.03$, $MSE = 25944$, $p = .020$, $\eta_p^2 = 0.15$, $BF = 1.01$. The main effects were moderated by the Source Confidence × Source Decision interaction, $F(2, 66) = 6.81$, $MSE = 17403$, $p = .002$, $\eta_p^2 = 0.17$, $BF = 3.90$. Note that two participants could not be included in this ANOVA because they had zero responses for particular cells of the analysis (hence $N = 34$ for this analysis). Closer inspection of the data showed that priming was greater for correct source decisions than incorrect source decisions, only when the source decision was made with high confidence, $F(1, 33) = 11.93$, $MSE = 31611$, $p = .002$, $\eta_p^2 = 0.27$, $BF = 19.30$. Identification RTs did not differ for items with correct and incorrect source decisions at lower levels of source confidence, both $ps > .30$, $BFs < 0.40$.

With regards to the relation between the priming effect and recognition decisions, as in Experiment 1, the results replicated Berry et al. (2012, Exp. 2). First, identification RTs for items judged old were faster than those judged new. This was the case for old items ($M = 222$ ms, $SE = 29$), $t(35) = 7.57$, $p < .001$, $d = 1.26$, $BF = 1.77 \times 10^{6}$, and also new items ($M = 84$ ms, $SE = 25$), $t(35) = 3.36$, $p < .001$, $d = 0.56$, $BF = 17.94$. Second, the priming effect for items judged new ($M = 112$ ms, $SE = 22$) exceeded chance, $t(35) = 5.02$, $p < .001$, $d = 0.84$, $BF = 1.41 \times 10^{3}$, and was smaller than the priming effect across all items (M difference = 147 ms,

$SE = 25$), $t(35) = 5.87$, $p < .001$, $d = 0.98$, $BF = 1.51 \times 10^{4}$. Third, identification RTs tended to decrease as confidence that an item is old increased (Fig. 2B). For old items, there were significant linear, quadratic, and cubic trends ($ps < .03$), although, pairwise comparisons (Bonferroni-adjusted) indicated that these trends were driven by the identification RTs for items receiving sure-old ratings being faster than those of items receiving any other rating ($ps < .001$; for all other pairwise comparisons, $ps > .17$). For new items, there were significant linear and quadratic trends ($ps < .01$), with pairwise comparisons (Bonferroni-adjusted) indicating that these trends were again driven by the identification RTs for new items receiving sure-old ratings being significantly faster than those receiving sure-new, probably-new and guess-new ratings, $ps < .01$ (all other comparisons, $ps > .07$). Thus, the variations in the priming effect with recognition confidence were primarily driven by identification RTs being shortest for "sure old" ratings. Note that $N = 31$ for this trend analysis due to five participants having zero responses in some of the response options.

In regards to the relationship of recognition and source memory, source d' was significantly higher for recognized than unrecognized items (M difference = 1.05, $SE = 0.13$), $t(35) = 7.81$, $p < .001$, $d = 1.30$, $BF = 3.47 \times 10^{6}$. Source d' significantly exceeded chance for recognized ($M = 1.11$, $SE = 0.15$), $t(35) = 7.38$, $p < .001$, $d = 1.23$, $BF = 1.05 \times 10^{6}$, but not unrecognized items ($M = 0.06$, $SE = 0.08$), $t(35) = 0.74$, $p = .46$, $d = 0.12$, $BF = 0.23$.

Discussion

As expected, a shorter list length led to superior source memory, recognition, and priming, relative to Experiment 1. Importantly, the association between priming and source decisions was, if anything, stronger than in Experiment 1. That is, the priming effect was greater for items with correct versus incorrect source decisions, and also tended to increase as confidence in the source decision increased. Furthermore, although only weakly present in Experiment 1, the Source Confidence × Source Decision interaction was significant in Experiment 2. This arose because the priming effect was only greater for items with correct versus incorrect source decisions at the highest level of source confidence. Once again, the associations between the priming effect and the recognition decision reported by Berry et al. (2012, Exp. 2) were found in this experiment.

Experiment 2 thus confirmed that the magnitude of the priming effect is related to both the correctness and confidence of the source decision, even when overall memory strength is higher. These findings are consistent with an account in which priming, recognition and source memory are driven by a single continuous strength signal. Before considering these findings in more detail and modeling the data, we consider an alternative explanation for the associations between the three tasks. It is possible that the interleaved nature of the identification and memory rating trials encourages speed of identification to influence the memory rating. Such a link has been proposed between priming and recognition. For example, if an item is identified relatively quickly, participants may attribute this relative ease of identification to the item's prior exposure at study (Jacoby & Dallas, 1981). This would lead to items with fast identification RTs being given higher recognition ratings. Given that high confidence source ratings tend to follow high confidence "old" ratings (Hautus et al., 2008; Onyper et al., 2010; see also Rotello, 2017), it follows that identification RTs would also decrease as confidence in the source decision increases, and this is the pattern we observed in Experiments 1 and 2. We investigated this alternative explanation of our findings in Experiments 3A and 3B.

Experiments 3A and 3B

In Experiments 3A and 3B, we tested whether the association between priming and source memory would still be found under conditions where the fluency of an item's identification would not affect the

memory decisions it received. This was achieved by presenting the CID trials in a separate phase to the memory ratings. Following the study phase, participants were first asked to identify old and new items in CID trials. Participants then completed a memory rating phase containing all old items, the new items from the CID phase (henceforth referred to as *partially-new items*), and previously unseen (henceforth *completely-new*) items.[1] Both types of new items were included in the rating phase in order to encourage recognition decisions to be based on whether an item was presented in the study phase or not. That is, including partially new items helped to prevent participants from making their recognition ratings solely on the basis of whether an item seemed familiar from its presentation in the CID phase, since partially new items were also presented in this stage. Experiment 3A was based upon Experiment 1, and had four study-test blocks. Experiment 3B had only one study-test block and further differed minimally in wording and presentation of the recognition and source memory probes from previous experiments to eliminate possible design artefacts as an explanation for the observed experimental effects.

Method

Participants

36 individuals (5 male; *M* age 19.75, *SD* = 1.76) took part in Experiment 3A and 37 individuals (6 male, *M* age = 19.95, *SD* = 3.09) took part in Experiment 3B for partial course credit. Two participants in Experiment 3B failed to follow instructions and were excluded from analyses.

Procedure

The materials, experimental set up, and study phase of Experiment 3A were identical to Experiment 1 except that, in the study phase, each item was shown twice in the same spatial location, with at least one intervening item before being repeated. An identification phase followed the study phase in which participants identified all 44 old items and 22 new items using the CID procedure. In the following memory rating phase, participants were shown all 44 old items, 22 partially-new items and 22 completely-new items. For each trial, the item was presented in the center of the screen with the recognition probe ("Was the word presented in the previous study phase?" 1 = sure no, 2 = probably no, 3 = guess no, 4 = guess yes, 5 = probably yes, 6 = sure yes") presented beneath the word. When the recognition rating was made, the probe was replaced with the source memory probe ("Was the word presented at the bottom or top? 1 = sure bottom, 2 = probably bottom, 3 = guess bottom, 4 = guess top, 5 = probably top, 6 = sure top").

Experiment 3B had the following differences to Experiment 3A: First, there was one study-test cycle, rather than four, and there were 64 old items in the study phase rather than 44 old items. In the CID phase, the 64 old items were presented with 48 new words. Then, in the rating phase, participants were shown all 64 old and 48 partially-new items, together with 16 completely-new items. There were also minor changes in the wording of the recognition and source probes. For the recognition rating, the probe was "Was the word presented in the previous study phase? 1 = high confidence no, 2 = medium confidence no, 3 = low confidence no, 4 = low confidence yes, 5 = medium confidence yes, 6 = high confidence yes". For the source rating, the probe was "Was the word previously presented towards the top or the bottom of the screen? 1 = high confidence top, 2 = medium confidence top, 3 = low confidence top, 4 = low confidence bottom, 5 = medium confidence bottom, 6 = high confidence bottom".

In both Experiments 3A and 3B, the rating phase instructions explicitly informed participants that the study phase was the phase in

---

[1] See the Supplemental Material for a pilot study for this experiment in which the new items were unique in the identification and rating phases (i.e., there was no repetition of new items across phases).

which words were presented towards the bottom or top of the screen. This was done to avoid confusion with the CID phase in which the items appeared in the center of the screen. Also, in each experiment, there were four additional trials at the start and end of the study phase (to control for primacy and recency effects), and the words from these trials were not presented again at test.

Initial screening of identification trials

The proportion of misidentified trials, after correction for typographical errors, was *M* = 3.58% (*SD* = 2.68) in Experiment 3A and *M* = 3.35% (*SD* = 2.86) in Experiment 3B. The proportion of trials on which no response was given in the identification phase was *M* = 0.03% (*SD* = 0.09) in Experiment 3A and *M* = 0.02% (*SD* = 0.13) in Experiment 3B. The proportion of trials on which the identification RT was less than 200 ms or greater than three standard deviations above mean identification RT (within participant) was *M* = 1.01% (*SD* = 0.43) in Experiment 3A and *M* = 1.32% (*SD* = 0.82) in Experiment 3B. After excluding these trials, the number of valid trials was *M* = 95.38% (*SD* = 2.68, Min = 86.93) in Experiment 3A and *M* = 95.31% (*SD* = 3.27, Min = 85.94) in Experiment 3B.

Measures

To take into account the two types of new items presented during the rating phase, we report recognition measures separately for the discrimination of old and partially-new and old and completely-new items. Beyond that, calculation of measures was the same as in previous experiments. In Experiment 3B, the source memory rating scale was reversed compared to previous experiments, so here source-bottom items were designated target items, and source confidence for source-top items was recoded when collapsing across source-top and source-bottom items.

In Experiment 3A, priming, recognition and source memory measures were reliable (priming, $r(34) = 0.40$, $p = .015$, $BF = 4.73$; recognition d' for old and partially-new items, $r(34) = 0.81$, $p < .001$, $BF = 1.67 \times 10^6$; recognition d' for old and completely-new items, $r(34) = 0.87$, $p < .001$, $BF = 2.11 \times 10^8$; source d', $r(34) = 0.82$, $p < .001$, $BF = 5.46 \times 10^6$). In Experiment 3B, priming, $r(33) = 0.39$, $p = .022$, $BF = 3.58$, recognition d' for old and partially-new items, $r(33) = 0.61$, $p < .001$, $BF = 280.33$, and for old and completely-new items, $r(33) = 0.39$, $p = .020$, $BF = 4.01$, were reliable, but for source memory was not, $r(33) = 0.13$, $p = .44$, $BF = 0.50$.

Results

Experiment 3A

The priming effect, recognition d' for both types of new items and source d' all exceeded chance (*M* priming = 273 ms, *SE* = 22, $t(35) = 12.53$, $p < .001$, $d = 2.09$, $BF = 4.02 \times 10^{11}$; *M* recognition partially-new $d' = 0.57$, *SE* = 0.10, $t(35) = 5.60$, $p < .001$, $d = 0.93$, $BF = 7.03 \times 10^3$; *M* recognition completely-new $d' = 1.35$, *SE* = 0.11, $t(35) = 11.85$, $p < .001$, $d = 1.97$, $BF = 8.81 \times 10^{10}$; *M* source $d' = 0.75$, *SE* = 0.11, $t(35) = 6.84$, $p < .001$, $d = 1.14$, $BF = 2.35 \times 10^5$). Table 1 shows the mean identification RT for new and old items, and also the mean hit rate and false alarm rate for recognition and source decisions. As in Experiment 1, recognition responding was significantly liberally biased for both discrimination of old and partially-new items (recognition $c = -0.58$, *SE* = 0.07), $t(35) = 8.55$, $p < .001$, $d = 1.42$, $BF = 2.52 \times 10^7$, and also old and completely-new items (recognition $c = -0.19$, *SE* = 0.07), $t(35) = 2.83$, $p = .008$, $d = 0.47$, $BF = 5.28$. Source responding was not significantly biased (source $c = 0.03$, *SE* = 0.04), $t(35) = 0.88$, $p = .39$, $d = 0.15$, $BF = 0.26$. Across participants, the priming effect was correlated with recognition d' for old and completely-new items, $r(34) = 0.51$, $p = .002$, $BF = 28.84$, but only marginally for old and partially-new items, $r(34) = 0.32$, $p = .055$, $BF = 1.84$. The priming effect was significantly correlated with source d', $r(34) = 0.56$,

$p < .001$, $BF = 84.40$. Recognition d' and source d' were significantly correlated: for old and partially-new items, $r(34) = 0.67$, $p < .001$, $BF = 2.26 \times 10^3$; for old and completely-new items, $r(34) = 0.73$, $p < .001$, $BF = 2.22 \times 10^4$.

As in Experiments 1 and 2, the priming effect was greater for items with correct source decisions than those with incorrect source decisions (M difference = 66 ms, $SE = 23$), $t(35) = 2.90$, $p = .006$, $d = 0.48$, $BF = 6.25$, see the left-hand side of Fig. 1C. Two-thirds of participants showed this difference. Like in Experiment 1, there was no significant corresponding main effect of source decision in a 3 (Source Confidence) $\times$ 2 (Source Decision) ANOVA, $F(1, 29) = 0.11$, $p = .74$, $\eta_P^2 = 0.004$, $BF = 0.16$. Once again, identification RTs decreased with increasing confidence in the source decision (see the right-hand side of Fig. 1C), $F(2, 58) = 9.68$, $MSE = 42400$, $p < .001$, $\eta_P^2 = 0.25$, $BF = 365$, but, as was found in Experiment 1, the Source Confidence $\times$ Source Decision interaction was not significant, $F(2, 58) = 0.39$, $p = .68$, $\eta_P^2 = 0.01$, $BF = 0.18$. Six participants could not be included in this ANOVA because they had zero responses for particular cells of the analysis (hence $N = 30$ for this analysis).

With regards to the relation between the priming effect and recognition decisions, the association was generally weaker than in Experiments 1 and 2. First, identification RTs tended to be faster for items judged old than those judged new, but this was only significant for old items (M difference = 79 ms, $SE = 32$, $t(35) = 2.51$, $p = .017$, $d = 0.42$, $BF = 2.70$, and not for new items ($M = 37$ ms, $SE = 34$), $t(35) = 1.09$, $p = .28$, $d = 0.18$, $BF = 0.31$. Second, the priming effect for items judged new ($M = 236$ ms, $SE = 26$) exceeded chance, $t(35) = 9.02$, $p < .001$, $d = 1.50$, $BF = 8.64 \times 10^7$, but there was no substantial support for this effect being smaller than the priming effect across all items (M difference = 37 ms, $SE = 21$), $t(35) = 1.75$, $p = .088$, $d = 0.29$, $BF = 0.71$. Third, as shown in Fig. 2C, the change in identification RTs with recognition confidence was clearly less pronounced than in Experiments 1 and 2. Only 18 out of 36 participants provided the full range of confidence ratings to permit this trend analysis, and it must be kept in mind that this analysis therefore has low power relative to Experiments 1 and 2. In these individuals, a significant linear trend was evident in old items, $t(85) = 2.06$, $p = .042$, although, as can be seen in Fig. 2C, identification RTs did not follow a clear monotonically decreasing function as old-new confidence increased (for both old and new items). For example, for old items, the mean identification RT for guess-old responses was actually numerically faster than sure-old responses. Furthermore, the linear trend for new items was not significant, $t(85) = 1.07$, $p = .29$.

Finally, regarding the relation of recognition and source memory, as in the previous experiments, source d' was significantly higher for recognized than unrecognized stimuli (M difference = 0.78, $SE = 0.16$), $t(35) = 4.80$, $p < .001$, $d = 0.80$, $BF = 763$. Source d' significantly exceeded chance for recognized ($M = 0.91$, $SE = 0.13$), $t(35) = 6.83$, $p < .001$, $d = 1.14$, $BF = 2.31 \times 10^5$, but not unrecognized stimuli ($M = 0.13$, $SE = 0.10$), $t(35) = 1.31$, $p = .19$, $d = 0.22$, $BF = 0.39$.

Experiment 3B

The priming effect, recognition d' for both types of new items and source d' all exceeded chance (M priming = 250 ms, $SE = 30$, $t(34) = 8.26$, $p < .001$, $d = 1.40$, $BF = 9.30 \times 10^6$; M recognition partially-new $d' = 0.61$, $SE = 0.09$, $t(34) = 7.15$, $p < .001$, $d = 1.21$, $BF = 4.78 \times 10^5$; M recognition completely-new $d' = 1.38$, $SE = 0.10$, $t(34) = 13.60$, $p < .001$, $d = 2.30$, $BF = 2.42 \times 10^{12}$; M source $d' = 0.38$, $SE = 0.07$, $t(34) = 5.27$, $p < .001$, $d = 0.89$, $BF = 2.63 \times 10^3$). Table 1 shows the mean identification RT for new and old items, and also the mean hit rate and false alarm rate for recognition and source decisions. In line with Experiment 3A, recognition responding was significantly liberally biased for the discrimination of old and completely-new items (recognition $c = -0.28$, $SE = 0.08$), $t(34) = 3.43$, $p = .002$, $d = 0.58$, $BF = 21.03$. In contrast to Experiment 3A, recognition responding for old and completely-new items was

conservatively biased (recognition $c = 0.15$, $SE = 0.07$), $t(34) = 2.23$, $p = .033$, $d = 0.38$, $BF = 1.60$. Source responding was conservatively biased (source $c = 0.11$, $SE = 0.04$), $t(34) = 2.92$, $p = .006$, $d = 0.49$, $BF = 6.52$, indicating a tendency to respond "top" over responding "bottom".

Across participants, the priming effect was significantly correlated with recognition d' for partially-new items, $r(33) = 0.48$, $p = .003$, $BF = 15.54$, and completely-new items, $r(33) = 0.49$, $p = .003$, $BF = 18.24$, but not significantly correlated with source d', $r(33) = 0.10$, $p = .57$, $BF = 0.43$. Recognition d' for partially-new items was significantly correlated with source d', $r(33) = 0.48$, $p = .004$, $BF = 14.13$, but not for completely-new items, $r(33) = 0.28$, $p = .097$, $BF = 1.23$.

As in Experiment 3A, the priming effect was greater for items with correct source decisions than those with incorrect source decisions (M difference = 57 ms, $SE = 21$), $t(34) = 2.76$, $p = .009$, $d = 0.47$, $BF = 4.54$, see the left-hand side of Fig. 1D. Three-quarters of participants showed this difference. As in Experiment 3A, there was no significant corresponding main effect of source decision in a 3 (Source Confidence) $\times$ 2 (Source Decision) ANOVA, $F(1, 27) = 0.66$, $MSE = 48265$, $p = .42$, $\eta_P^2 = 0.02$, $BF = 0.21$. Once again, identification RTs decreased with increasing confidence in the source decision (see the right-hand side of Fig. 1D), $F(1.51, 40.83) = 8.54$, $MSE = 83875$, $p = .002$, $\eta_P^2 = 0.24$, $BF = 187.32$, and, as was found in Experiment 3A, the Source Confidence $\times$ Source Decision interaction was not significant, $F(2, 54) = 2.08$, $MSE = 50991$, $p = .14$, $\eta_P^2 = 0.07$, $BF = 0.46$. Eight participants could not be included in this ANOVA because they had zero responses for particular cells of the analysis (hence $N = 27$ for this analysis).

With regards to the relation between the priming effect and recognition decisions, the pattern of the association was similar to the one observed in Experiment 3A. First, identification RTs tended to be faster for items judged old than those judged new, but this was only significant for old items (M difference = 128 ms, $SE = 39$), $t(34) = 3.26$, $p = .003$, $d = 0.55$, $BF = 14.08$, and not for new items ($M = -9$ ms, $SE = 38$), $t(34) = -0.24$, $p = .82$, $d = 0.04$, $BF = 0.19$. Second, the priming effect for items judged new ($M = 142$ ms, $SE = 37$) exceeded chance, $t(34) = 3.84$, $p < .001$, $d = 0.65$, $BF = 58.10$. In line with Experiments 1 and 2, and more clearly than in Experiment 3A, this was smaller than the overall priming effect (M difference = 108 ms, $SE = 38$), $t(34) = 3.06$, $p = .004$, $d = 0.52$, $BF = 8.73$. Third, as shown in Fig. 2D, the change in identification RTs with recognition confidence was similar to Experiment 3A. Only 16 out of 35 participants provided the full range of confidence ratings to permit this trend analysis, so this analysis, as was the case in Experiment 3A, has relatively low power. In these individuals, a significant linear trend was evident in old items, $t(75) = 2.72$, $p = .008$, though the quadratic trend was also significant, $p = .049$. As Fig. 2D shows, identification RTs more clearly follow a monotonically decreasing function as old-new confidence increased than was the case in Experiment 3A. However, none of the pairwise comparisons were significant, all $ps = 1$ (Bonferroni-adjusted). Furthermore, the linear trend for new items was not significant, $t(75) = 0.88$, $p = .38$.

Finally, regarding the relation of recognition and source memory, once again, source d' was significantly higher for recognized than unrecognized stimuli (M difference = 0.42, $SE = 0.12$), $t(34) = 3.42$, $p = .002$, $d = 0.58$, $BF = 20.27$. Source d' also significantly exceeded chance for recognized ($M = 0.50$, $SE = 0.09$), $t(34) = 5.79$, $p < .001$, $d = 0.98$, $BF = 1.13 \times 10^4$, but not unrecognized stimuli ($M = 0.09$, $SE = 0.10$), $t(34) = 0.92$, $p = .36$, $d = 0.16$, $BF = 0.27$.

Discussion

The results of Experiments 3A and 3B show that an association between the priming effect and the source decision persists, even when identification RTs and memory ratings are measured in separate test

phases. Under these conditions, fluency of identification of an item would not be expected to affect the recognition and source ratings. As was found in Experiments 1 and 2, the priming effect was greater for items with correct versus incorrect source decisions. The size of these effects was comparable across experiments: Cohen's d_z was 0.52 in Experiment 1 versus 0.48 and 0.47 in Experiments 3A and 3B, respectively. Also like previous experiments, the priming effect tended to increase as confidence in the source decision increased: η_p^2 for the main effect of source confidence was 0.31 in Experiment 1 versus 0.25 and 0.24 in Experiments 3A and 3B, respectively. Thus, the association between the priming effect and the source decision is not attributable to participants using the speed with which they identify an item in a CID procedure to inform their memory decisions. The association also persists despite minor procedural variations across experiments (i.e., the wording and polarity of source and recognition probes, and whether performance is measured across multiple study-test phase cycles or not).

Although evidence for an association between the magnitude of the priming effect and recognition decisions was found in these two experiments, the association was clearly weaker than in Experiments 1 and 2. At first glance, this might be taken to support an account in which fluency of identification drove the observed association between priming and recognition in Experiments 1 and 2 (and Berry et al., 2012, Exp.2), but not in Experiments 3A and 3B. However, if the association between priming and recognition is driven by fluency, then this does not explain why an association persisted in Experiment 3A and 3B when the influence of fluency was precluded. That is, the priming effect was still greater for studied items that were recognized than those that were not recognized (see also Ostergaard, 1998; Sheldon & Moscovitch, 2010, for a similar findings). It is also not clear how a fluency attribution account would explain why the priming-source association persisted in Experiments 3A and 3B.

A more plausible explanation for the weaker association between priming and recognition in Experiments 3A and 3B is that the ability to decide whether a studied item was presented in the study phase was affected by its additional presentation in the CID phase. Recognition decisions may then have sometimes been based on memory strength that originated from an item's presentation in the CID phase, rather than from the study phase. The effect of this would be to weaken the association between recognition and priming because both would no longer be driven by the same memory strength signal (i.e., the signal from the study phase). Consistent with this explanation, the false alarm rate to partially-new items was greater than that of completely-new items in Experiment 3A (M difference = 0.26, SE = 0.03, $t(35)$ = 8.96, $p < .001$, $d = 1.49$, $BF = 7.54 \times 10^8$, and Experiment 3B (M difference = 0.27, SE = 0.03), $t(34)$ = 9.39, $p < .001$, $d = 1.59$, $BF = 1.72 \times 10^8$. This suggests that recognition ratings were indeed sometimes based on the familiarity from an item's presentation in the CID phase. Regardless of the cause of the weaker association between priming and recognition in Experiments 3A and 3B, most important for our purposes is that an association between priming and source memory was still found.

The main findings from our behavioral work are thus: (1) priming is associated with source memory, such that higher priming is associated with correct source decisions and source decisions made with greater confidence, (2) priming is associated with recognition memory, such that higher priming is associated with recognition of studied items as old. Additionally, (3) recognition memory is associated with source memory, such that studied items recognized as old are associated with correct source decisions, whereas source d' does not differ from chance for studied items that are not recognized. Fluency, as a result of interleaved identification and memory rating trials, cannot account completely for these associations. In the next section we explore the potential for an extended version of the single-system model proposed by Berry et al. (2012) to explain these associations simultaneously. We will additionally formulate a simple multiple-systems version of the model

that assumes a distinct, stochastically independent memory system for the priming task and memory rating tasks to compare both models' quantitative fit and qualitative predictions for the associations.

Modeling

In this section we will formally test whether the assumption of a single memory signal underlying performance in priming, recognition, and source memory is sufficient to explain the observed data, or if we need to assume that an additional distinct, stochastically independent memory signal drives priming. To formalize the single-system model, we extended the single-system (SS) model of priming and recognition (Berry et al., 2012) to source memory. We extended this model in a straightforward manner by assuming that, for a given item, the greater its underlying strength value, the faster it will tend to be identified, the greater the confidence with which it will tend to be judged as old, and the greater the confidence with which its source will tend to be correctly classified. As a baseline for comparison, we also derived a multiple systems (MS) version of model. The MS model tests the idea that the memory system driving priming is stochastically independent from the one driving the explicit memory measures of recognition and source. This implements a multiple systems view of implicit and explicit memory where "memory systems of the mammalian brain operate independently and in parallel to support behavior" (Squire & Dede, 2015, p. 9). This view suggests that "the priming effects […] are independent of recognition memory" (Tulving, Schacter, & Stark, 1982, p. 1) and results in predictions such as that an "individual might have a fear of large dogs [resulting from implicit memory], quite independently of whether the event itself is remembered" (Squire & Dede, 2015, p. 3). Further, such a multiple systems model also implements the assumption of independence central to dual-process theory. While dual-process theory can be formulated to not assume stochastic independence (e.g., similar to the MS2 model in Berry et al., 2012; DPSD models in Moran & Goshen-Gottstein, 2015; Pratte & Rouder, 2011), it is nevertheless the case that prominent models and widely used experimental methods do make such assumptions (e.g., Yonelinas, 1994, 2002; the process dissociation procedure, Jacoby, 1991). This MS model thus formally implements multiple, independent systems within our signal detection framework. We first present formal specification of these models, then describe how their parameters are estimated, and how well the models account for the data from Experiments 1-3A/3B. To preview, both the SS and MS model show grave misfits of model predictions to data at the outset. Accordingly, we then modify the SS model (along with the MS version of it) in order to incorporate recent developments in modeling conjoint recognition and source rating data. These modifications turn out to be particularly important for the SS model to better explain the pattern of results of Experiments 1-3A/3B. While the SS model accounts for the pattern of data in Experiments 1 and 2 particularly well, results are less conclusive for Experiment 3A and 3B.

Model specification

SS model

The main assumption of the SS model is that the same memory strength signal drives priming, recognition and source memory. Each item in the test phase is associated with a memory strength-of-evidence variable f, which is a random variable drawn from a normal (Gaussian) distribution with mean μ_k and variance σ_f^2 (i.e. $f_k \sim N(\mu_k, \sigma_f^2)$). The k subscript stands for the item type, $k = \{N, A, B\}$, where N = new items, A = source-A items and B = source-B items. As in Berry et al. (2012), the mean f of studied items is higher than that of new items due to being presented in the study phase, hence if μ_N is fixed to zero, then $\mu_A = \mu_B = \mu \geq 0$.

Identification RT and recognition judgments are modeled as in Berry et al. (2012). One value of f is sampled for each individual item from the relevant distribution, and this value is combined with a randomly

sampled, normally distributed noise value. Identification RT is modeled as $RT = b - sf_k + e_p$, where $e_p \sim N(0, \sigma_p^2)$ and b and s are scaling parameters whose value is greater than 0. The variable giving rise to recognition judgments, J_r, is modeled as $J_r = f_k + e_r$, where $e_r \sim N(0, \sigma_r^2)$. The noise variables e_p and e_r are uncorrelated, and both are uncorrelated with f. In line with signal detection theory, each resulting value of J_r is compared with decision criteria in order to determine the recognition confidence rating (e.g., $C_{r1} - C_{r5}$ when there are six recognition rating options).

When making the source decision, a participant must decide whether an item was previously presented in one of a two sources (e.g., towards the bottom or top of the computer screen). We assume that participants make this decision by comparing the source-A and source-B strengths of an item, that is, the source decision is a relative judgment, rather than an absolute one, as is the case with recognition. Responses are therefore based upon the difference in the source-A and source-B strengths of an item, $f_A - f_B$. As with the generation of RT and J_r, the assessment of $f_A - f_B$ is also subjected to a noise variable, e_s, to give J_s, where $J_s = f_A - f_B + e_s$, where $e_s \sim N(0, \sigma_s^2)$. As with e_p and e_r, e_s is not correlated with the other noise variables or with f.

For source-A items, $f_A \sim N(\mu, \sigma_f^2)$, and there will be no memory strength associated with source-B for these items on average, and so $f_B \sim N(0, \sigma_f^2)$. The converse is true for source-B items: that is, $f_B \sim N(\mu, \sigma_f^2)$ and $f_A \sim N(0, \sigma_f^2)$. Thus, for source-A items, $J_s \sim N(\mu, 2\sigma_f^2 + \sigma_s^2)$, and for source B items, $J_s \sim N(-\mu, 2\sigma_f^2 + \sigma_s^2)$. For new items, the value of f_N is not associated with either source, as is assumed in multivariate signal detection models of item recognition and source memory (e.g., DeCarlo, 2003). Accordingly, we make the simplifying assumption that only noise drives ratings for new items. Hence, for new items, $J_s = e_s$, and $J_s \sim N(0, \sigma_s^2)$. Thus, we can write J_s for each type of item k as $J_{s|k} \sim N(\beta_k \mu_k, 2\beta_k^2 \sigma_f^2 + \sigma_s^2)$, where $\beta_N = 0$, $\beta_A = 1$, and $\beta_B = -1$. As in signal detection models of source decisions (DeCarlo, 2003), an item's value of J_s is compared with decision criteria in order to determine the source rating (e.g., C_{s1}-C_{s5} when there are six source rating options), with the source decision criteria the same across all recognition ratings, i.e. assuming linear decision bounds. This closely matches how source memory is typically modeled in two-dimensional signal detection models of recognition and source memory (e.g. DeCarlo, 2003; Hautus et al., 2008; Starns et al., 2008; Slotnick & Dodson, 2005). In contrast to those implementations, we merely make it explicit that source judgments are based on a relative judgment of strength values associated with both sources. Thus, we assume that participants' source judgments are relative strength judgments, following from participants weighing an item's source-A strength against an item's source-B strength. As a consequence of this implementation, correct source decisions will tend to arise if an item's source strength exceeds the noise from the competing source, and incorrect source decisions will tend to arise if there is little to no signal or memory strength from the correct source, and thus noise exceeds the signal. This means that correct source decisions will tend to be associated with high source-specific memory strength (e.g., a relatively high value of f_A for a source-A item) and incorrect source decisions will tend to be associated with low source-specific memory strength (e.g., a relatively low value of f_A for a source-A item).

MS model

The MS model can be derived from the SS model by extending it to multiple signals such that we define f_p, f_r and f_s as distinct memory signals underlying performance in the priming, recognition, and source tasks respectively. Whereas one value of f is sampled in the SS model to derive an item's priming, recognition and source memory performance, in the MS model, each item is associated with three values of f at test. Now the distribution underlying priming is defined as $f_{p|k} \sim N(\mu_{p|k}, \sigma_f^2)$. The distribution underlying recognition performance is defined as $f_{r|k} \sim N(\mu_{r|k}, \sigma_f^2)$ respectively. As in the SS model, $\mu_{p|N} = 0$ and $\mu_{p|A} = \mu_{p|B} = \mu_p \geq 0$ and $\mu_{r|N} = 0$ and $\mu_{r|A} = \mu_{r|B} = \mu_r \geq 0$. Analogously the source distribution is defined as $f_{s|k} \sim N(\mu_{s|k}, \sigma_f^2)$.

RT, J_r and J_s are derived as in the SS model by combining each value of f with their respective task-specific noise terms. J_s is again derived by comparing $f_{s|B}$ and $f_{s|A}$ for both source-A and source-B items. In the multiple systems model μ_p, μ_r and μ_s are free to vary.

Since values of f_p, f_r, and f_s are sampled from three distinct distributions, an additional consideration has to be given to the correlations (w) between these variables (i.e., w_{pr}, w_{ps}, w_{rs}). In the SS model, the strength distributions underlying performance in priming, recognition and source memory tasks are identical, hence $w_{pr} = w_{ps} = w_{rs} = 1$. A fully unconstrained version of the MS model (as in the MS2 model in Berry et al., 2012) that does not assume stochastic independence of the memory signal underlying priming and that underlying recognition or source memory would implement μ_p, μ_r and μ_s as parameters that are free to vary but would allow f_p, f_r and f_s to be positively correlated across items. This means that an item's priming, recognition and source performance may be related, even though the mean memory strengths in the distinct systems are unrelated. A strict version of the multiple-systems model (as in the MS1 model in Berry et al., 2012) would not allow f_p, f_r and f_s to correlate across items. Under such a model, not only would μ_p, μ_r and μ_s be free to vary, identification RTs, recognition and source memory judgments would not be correlated across trials.

A fully constrained version of this MS model (as in the MS1 model in Berry et al., 2012) would assume stochastic independence of all memory signals. This would be achieved by allowing μ_p, μ_r and μ_s to vary freely and constrain f_p, f_r and f_s to not correlate across items. Critically, such a model would assume not only priming and recognition or priming and source memory, but also recognition and source memory are driven by distinct, stochastically independent signals.

Here we implement a version of this MS1 model that assumes stochastic independence only for the priming and recognition as well as the priming and source memory relationship, while assuming that the same memory signal underlies recognition and source memory responding. Within the framework we have presented thus far, this is implemented such that μ_p and μ_r are allowed to vary for old items, while $\mu_s = \mu_r$. Additionally, as in the SS model, f_r and f_s correlate perfectly ($w_{rs} = 1$), but f_p and f_r ($w_{pr} = 0$) and f_p and f_s ($w_{ps} = 0$) do not. This means this MS model implements the idea that the signal driving priming is independent of the one that drives recognition and source memory (e.g., Squire & Dede, 2015; Tulving et al., 1982). Given that recognition and source memory are modeled in this MS model in the same way as in the SS model, this means that any differences in goodness of fit of the models can be largely attributed to the nature of the assumed relationship of priming with recognition and source memory, as the key relationships we are investigating in this paper.

Formal specification

Based on the above distributional and decisional assumptions, the mean vector for RT, J_r and J_s can then be defined for both the SS and MS models as

$$\mu_k = \begin{bmatrix} b - s\mu_{p|k} \\ \mu_{r|k} \\ \beta_k \mu_{s|k} \end{bmatrix},$$

and the covariance matrix as

$$\Sigma_k = \begin{bmatrix} s^2\sigma_f^2 + \sigma_p^2 & -sw_{pr}\sigma_f^2 & -sw_{ps}\beta_k\sigma_f^2 \\ -sw_{pr}\sigma_f^2 & \sigma_f^2 + \sigma_r^2 & w_{rs}\beta_k\sigma_f^2 \\ -sw_{ps}\beta_k\sigma_f^2 & w_{rs}\beta_k\sigma_f^2 & 2\beta_k^2\sigma_f^2 + \sigma_s^2 \end{bmatrix}$$

In the SS model, $\mu_{p|k} = \mu_{r|k} = \mu_{s|k} = \mu_k$ with $\mu_N = 0$ and $\mu_A = \mu_B = \mu \geq 0$, as well as $w_{pr} = w_{ps} = w_{rs} = 1$. For the MS model, $\mu_{p|k}$ and $\mu_{r|k} = \mu_{s|k}$ are free to vary, but $\mu_{p|N} = \mu_{r|N} = \mu_{s|N} = 0$ and $\mu_{p|A} = \mu_{p|B} \geq 0$, $\mu_{r|A} = \mu_{r|B} \geq 0$, $\mu_{s|A} = \mu_{s|B} \geq 0$. Further, in the MS model, $w_{rs} = 1$, and $w_{pr} = w_{ps} = 0$. Additionally, as in Berry et al.

(2012), the slope s in the MS model is set to be equal to the slope estimated in the SS model.[2] To implement equal variance for old and new items, in line with the implementation in Berry et al. (2012), $\sigma_f = \sigma_r = 1/\sqrt{2}$ for both old and new items. σ_p and σ_s are free to vary but are set equal for old and new items. The SS model for priming, recognition and source memory is thus described by 15 parameters in total (μ_r, b, s, σ_p, σ_s, C_{r1}, C_{r2}, C_{r3}, C_{r4}, C_{r5}, C_{s1}, C_{s2}, C_{s3}, C_{s4}, C_{s5}) as is the MS model (μ_r, μ_p, b, σ_p, σ_s, C_{r1}, C_{r2}, C_{r3}, C_{r4}, C_{r5}, C_{s1}, C_{s2}, C_{s3}, C_{s4}, C_{s5}). In the first instance, we are therefore describing a SS and MS model with linear decision bounds and equal variance assumptions (henceforth referred to as SS-Lin-EV, MS-Lin-EV models). We will later modify these models by replacing the linear source decision bounds with criteria converging with increasing recognition confidence (SS-Con-EV, MS-Con-EV models), and finally adding an unequal variance assumption (SS-Con-UV, MS-Con-UV). As suggested, these modifications will iteratively improve model fits of the SS model in particular.

Parameter estimation

The parameters of all models were determined with maximum likelihood estimation procedures.[3] The likelihood for a tuple of observations (RT, R, S), where RT denotes the identification RT, R the recognition judgement and S the source memory judgement on a given test trial is given as

$$P(RT, R, S) = P(R, S|RT)P(RT) = \left(\iint P(J_r J_s |RT) dJ_r dJ_s \right) P(RT)$$

where the conditional distribution $P(R, S|RT)$ is multivariate Normal, with mean

$$\boldsymbol{\mu}|RT = \boldsymbol{\mu}_1 + \Sigma_{12}\Sigma_{22}^{-1}(RT - \mu_{RT})$$

and covariance

$$\Sigma|RT = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}^{T}$$

where

$$\boldsymbol{\mu}_1 = \begin{bmatrix} \mu_{J_r} \\ \mu_{J_s} \end{bmatrix} = \begin{bmatrix} \mu_{r|k} \\ \beta_k \mu_{s|k} \end{bmatrix}$$

$$\Sigma_{11} = \begin{bmatrix} \sigma_f^2 + \sigma_r^2 & w_{rs}\beta_k\sigma_f^2 \\ w_{rs}\beta_k\sigma_f^2 & 2\beta_k^2\sigma_f^2 + \sigma_s^2 \end{bmatrix}$$

$$\Sigma_{12} = \begin{bmatrix} -sw_{pr}\sigma_f^2 \\ -sw_{ps}\beta_k\sigma_f^2 \end{bmatrix}$$

and

$$\Sigma_{22} = \begin{bmatrix} s^2\sigma_f^2 + \sigma_p^2 \end{bmatrix}$$

This becomes

$$\boldsymbol{\mu}|RT = \begin{bmatrix} \mu_{J_r|RT} \\ \mu_{J_s|RT} \end{bmatrix} = \begin{bmatrix} \mu_{r|k} - \dfrac{sw_{pr}\sigma_f^2(RT - b + s\mu_{p|k})}{s^2\sigma_f^2 + \sigma_p^2} \\ \beta_k\mu_{s|k} - \dfrac{sw_{ps}\beta_k\sigma_f^2(RT - b + s\mu_{p|k})}{s^2\sigma_f^2 + \sigma_p^2} \end{bmatrix}$$

and

$$\Sigma|RT = \begin{bmatrix} \sigma_f^2 + \sigma_r^2 - \dfrac{s^2w_{pr}^2\sigma_f^4}{s^2\sigma_f^2 + \sigma_p^2} & w_{rs}\beta_k\sigma_f^2 - \dfrac{s^2w_{ps}w_{pr}\beta_k\sigma_f^4}{s^2\sigma_f^2 + \sigma_p^2} \\ w_{rs}\beta_k\sigma_f^2 - \dfrac{s^2w_{ps}w_{pr}\beta_k\sigma_f^4}{s^2\sigma_f^2 + \sigma_p^2} & 2\beta_k^2\sigma_f^2 + \sigma_s^2 - \dfrac{s^2w_{ps}^2\beta_k^2\sigma_f^4}{s^2\sigma_f^2 + \sigma_p^2} \end{bmatrix}$$

The full likelihood of a trial can therefore be expressed as

$$L(RT, R, S|k) = F_m(C_{upper}, C_{lower}|\boldsymbol{\mu}|RT, \Sigma|RT) \times \phi(RT| \mu_{RT}, \sigma_{RT}^2)$$

where k = {New, source-A = Top, source-B = Bottom}; RT is identification RT; R is recognition confidence rating 1…6; S is the source confidence rating 1…6; F_m denotes the cumulative normal distribution function for the multivariate normal distribution; $\boldsymbol{\mu}|RT$ denotes the mean of the conditional distribution of R and S given RT and $\Sigma|RT$ denotes the covariance matrix; C_{upper} denotes the upper truncation points corresponding to the regions in a bivariate signal detection space for judgments R and S, and C_{lower} denotes the lower truncation points of these regions, such that for an item classified with $R = i$, $S = j$, $C_{lower} = \{C_{ri\text{-}1}, C_{rj\text{-}1}\}$ and $C_{upper} = \{C_{ri}, C_{sj}\}$, where $C_{r0} = -\infty$, $C_{r1\text{-}5}$ are the recognition criteria, $C_{r6} = \infty$ and $C_{s0} = -\infty$, $C_{s1\text{-}5}$ are the recognition criteria, $C_{s6} = \infty$; ϕ denotes the normal density function; $\mu_{RT} = b - s\mu_{p|k}$ and $\sigma_{RT}^2 = s^2\sigma_f^2 + \sigma_p^2$.

Fitting procedure and expected values

The maximum likelihood parameters of each model were determined for each participant using the following procedure. For each participant's data, the likelihood function was used to determine the likelihood of every valid trial (i.e., on which the item was correctly identified, the identification RT > 200 ms and < mean identification RT ± 3 *SD*), given a set of parameter values. The log-likelihood was summed across all trials and converted to a negative value to be used by a function minimization algorithm (Nelder-Mead), as implemented by *optim* in R (R Core Team, 2017). The minimization routine was run ten times for each participant's data set in the first instance. Different starting values of the parameters to be estimated were used for the minimization routine in order to maximize the chance of finding the global minimum for the negative log-likelihood of each model, with the starting parameters for the first half of runs drawn from normal distributions centered on the respective mean parameter values estimated from the data, and half drawn from uniform distributions with appropriate constraints. We next fitted the model again a further five times, using the parameter estimates from the best-fitting run of the initial ten runs as the starting parameters for the first of the five runs. For the following four runs, we used the estimated parameter values from the preceding run as starting parameters to avoid premature termination of the simplex fitting procedure to identify the parameter estimates associated with the lowest negative log-likelihood.[4]

The maximum likelihood parameter estimates from the best-fitting run for each participant were used to simulate an identification RT, recognition decision and source decision for 25,000 simulated old items (12,500 source-top, 12,500 source-bottom items) and 25,000 simulated new items. Priming, recognition and source memory performance measures were then calculated from the simulated trials in the same way as the experimental data. This gave the expected priming, recognition and source results for each model (see Figs. 3, 4 and 7 in this paper, and Figs. S2–S6 in the Supplemental Material).

---

[2] Alternatively the slope could be estimated as a free parameter or eliminated from the MS models altogether (by setting s = 1). We explored these possibilities, see the Supplemental Material for details. In brief, estimating the slope as a free parameter prevents the MS models from being identifiable. Eliminating it from the model, and leaving other parameters to account for the elimination, leads to the same fits as fixing it to the slope estimated by the SS model for the equal variance models, but a worse fit for the unequal variance, multiple systems model.

[3] We thank Maarten Speekenbrink for assistance with deriving an earlier version of the likelihood function.

[4] In the Supplemental Material, we report results of parameter recovery simulations for all models. With increasing complexity of the model, parameter recovery becomes less successful, in particular when simulating a relatively low number of trials.
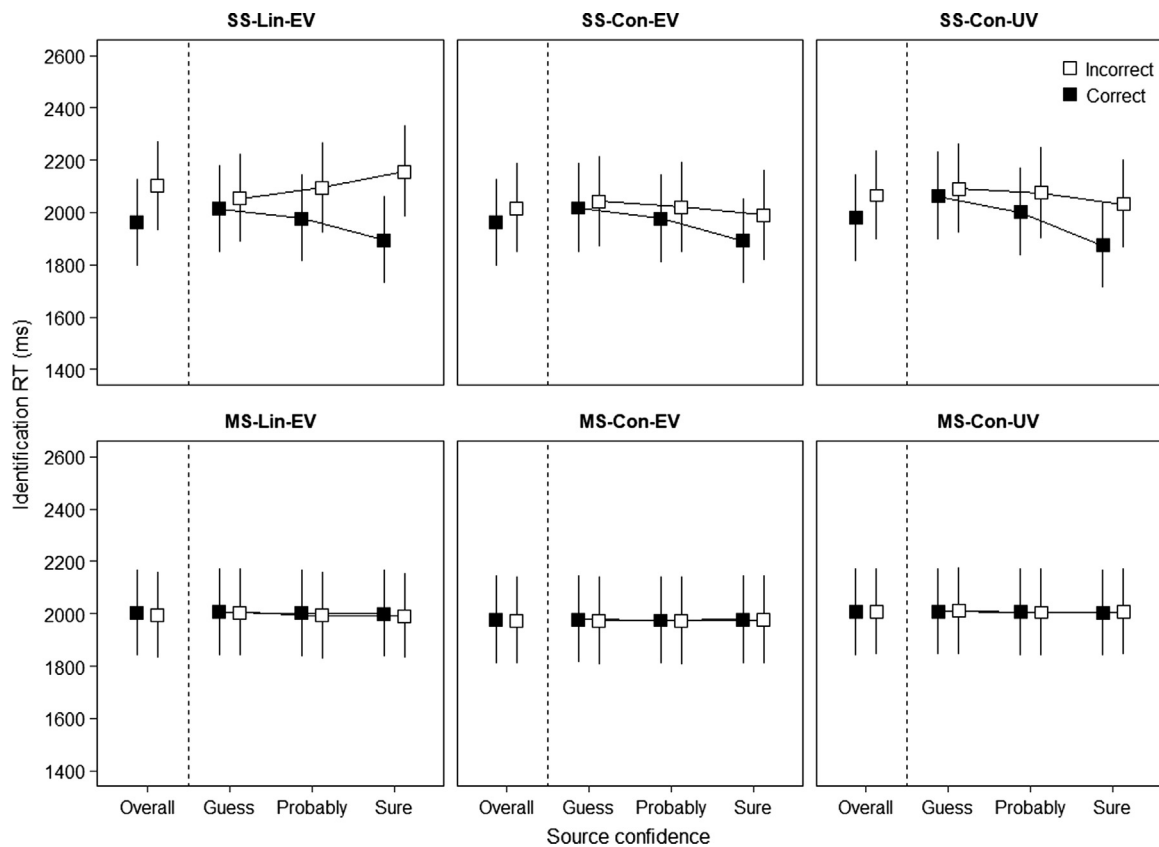
Fig. 3. Illustrative model predictions for the relationship of identification RTs and source confidence for correct and incorrect source decisions to studied items. The predictions were generated by simulating 25,000 old (and 25,000 new) items according to the maximum likelihood parameter estimates for each participant in Experiment 2. Error bars indicate 95% between-subjects confidence intervals of the mean. SS = single-system model, MS = multiple-systems model, -Lin = linear decision bounds, -Con = converging criteria, -EV = equal variance, -UV = unequal variance.

Model fitting results

SS and MS models with linear decision bounds and equal variance assumption (SS-Lin-EV, MS-Lin-EV)

Parameters estimates of the models are shown in Table 2; their qualitative predictions are illustratively shown in Figs. 3 and 4 (panels SS-Lin-EV, MS-Lin-EV)[5]. The SS-Lin-EV model predicts that identification RTs will be faster for items with correct versus incorrect source decisions, as found in Experiments 1-3A/3B, see Fig. 3, left-hand side. For items with correct source decisions, it also reproduces the trend for identification RTs to decrease as confidence in the source decision increases, see Fig. 3 (SS-Lin-EV, right-hand side). However, in contrast to the observed data (Fig. 1, A-D, right-hand side), for items with incorrect source decisions it predicts that identification RTs increase as confidence in the source decision increases (Fig. 3, SS-Lin-EV, right-hand side). Thus, the SS-Lin-EV model does not completely explain the observed relation between priming and source confidence. With regards to the association between priming and recognition (for the observed data, see Fig. 2, A–D), the SS-Lin-EV model correctly predicts that identification RTs will decrease as recognition confidence increases (Fig. 4, SS-Lin-EV). In contrast, the MS-Lin-EV model incorrectly predicts that identification RTs do not vary according to the source decision (Fig. 3, MS-Lin-EV) or recognition confidence (Fig. 4, MS-Lin-EV).

Despite the obvious inability of either the SS-Lin-EV or the MS-Lin-EV model to account for qualitative pattern of data, we also compared the quantitative model fit across experiments. In Experiments 1 and 2, the SS-Lin-EV model outperforms the MS-Lin-EV model when summed across participants (Table 5, top rows)[6] and a higher percentage of participants is best fit by the SS model (Fig. 5, -Lin-EV panel). For Experiments 3A and 3B, the MS-Lin-EV model is preferred over the SS-Lin-EV model. While the difference in the likelihood summed across participants is small in Experiment 3B (Table 5, top rows), in both Experiments 3A and 3B the majority of participants are better accounted for by the MS-Lin-EV than SS-Lin-EV model (Fig. 5, -Lin-EV panel). However, contrasting the observed and predicted relationship of priming and source memory (observed: Fig. 1, panel C and D; predicted: Fig. 3, left column) shows that in both experiments correct source decisions are associated with faster identification than incorrect source decisions, as predicted by the SS model. The gains of the MS-Lin-EV model relative to the SS-Lin-EV model in fitting the data of Experiments 3A and 3B are likely driven by the weaker relationship of priming and recognition in these experiments, compared to Experiments 1 and 2 (compare Fig. 2, panel C and D, with Fig. 2, panel A and B, and see

---

[5] Figures showing direct comparisons of observed and predicted data for all experiments for overall identification RT, proportions of items assigned recognition and source ratings, identification RT across recognition ratings and identification RT across source ratings are in the Supplemental Material. We focus on general trends predicted by the models for ease of illustrating individual models' predictions in the main of the paper.

[6] We report AIC and BIC in the table alongside the negative log-likelihood for later comparison of fit between SS models. Given SS-Lin-EV and MS-Lin-EV model have the same number of parameters (as do the later pairs of models discussed), the quantitative comparison within the SS/MS model pair here is naturally equivalent for negative log-likelihood, AIC and BIC. In the Supplemental Material, we report five-fold cross-validation analyses for all models for Experiment 2 and 3A. The results of these analyses lead to equivalent conclusions as the use of AIC and BIC for model selection reported in the paper.
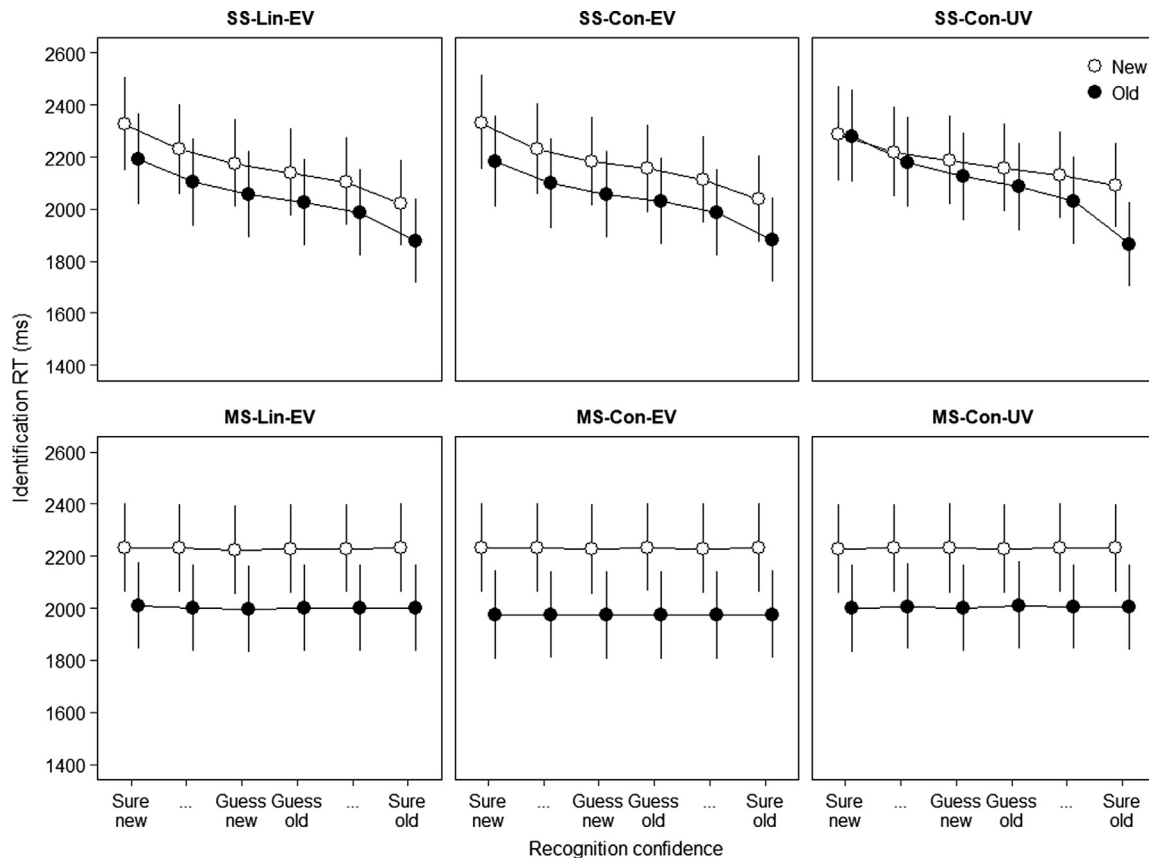
Fig. 4. Illustrative model predictions for the relationship of identification RTs and recognition decisions to new and old items. The predictions were generated by simulating 25,000 old and 25,000 new items according to the maximum likelihood parameter estimates for each participant in Experiment 2. Error bars indicate 95% between-subjects confidence intervals. SS = single-system model, MS = multiple-systems model, -Lin = linear decision bounds, -Con = converging criteria, -EV = equal variance, -UV = unequal variance.

Table 2
Means and standard deviations (in parentheses) of the parameter estimates of the standard SS and MS models across participants in Experiments 1, 2, 3A, and 3B.

Parameter	SS-Lin-EV				MS-Lin-EV			
	Exp 1 $N = 25$	Exp 2 $N = 30$	Exp 3A $N = 25$	Exp 3B $N = 24$	Exp 1 $N = 25$	Exp 2 $N = 30$	Exp 3A $N = 25$	Exp 3B $N = 24$
μ_r	0.75 (0.32)	1.07 (0.50)	0.76 (0.36)	0.61 (0.24)	0.76 (0.31)	1.07 (0.52)	0.74 (0.37)	0.58 (0.24)
μ_p	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	0.92 (0.52)	1.34 (0.47)	1.85 (1.11)	1.46 (1.04)
μ_s	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$
w_{pr}	$=1$	$=1$	$=1$	$=1$	$=0$	$=0$	$=0$	$=0$
w_{ps}	$=1$	$=1$	$=1$	$=1$	$=0$	$=0$	$=0$	$=0$
w_{rs}	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$
b	2159 (477)	2200 (463)	2054 (422)	2173 (481)	2179 (482)	2229 (453)	2162 (443)	2247 (517)
S	213 (114)	215 (83)	207 (101)	202 (117)	$=SS$	$=SS$	$=SS$	$=SS$
$\sigma_{f|new}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$
$\sigma_{f|old}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$
σ_r	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$	$=1/\sqrt{2}$
σ_p	553 (125)	560 (127)	545 (129)	497 (163)	551 (124)	557 (126)	530 (127)	484 (157)
σ_s	1.06 (0.62)	1.09 (0.60)	0.90 (0.85)	1.70 (1.45)	1.18 (0.83)	1.08 (0.55)	0.89 (0.81)	1.74 (1.71)
C_{r1}	−1.10 (0.56)	−0.51 (0.58)	−1.59 (1.36)	−1.73 (1.29)	−1.09 (0.56)	−0.52 (0.59)	−1.50 (1.15)	−1.59 (1.04)
C_{r2}	−0.30 (0.48)	0.09 (0.47)	−0.57 (0.78)	−0.76 (0.67)	−0.30 (0.47)	0.09 (0.47)	−0.59 (0.81)	−0.77 (0.67)
C_{r3}	0.11 (0.31)	0.42 (0.38)	−0.17 (0.45)	−0.08 (0.55)	0.12 (0.30)	0.43 (0.39)	−0.18 (0.45)	−0.10 (0.55)
C_{r4}	0.49 (0.33)	0.73 (0.37)	0.13 (0.38)	0.46 (0.47)	0.50 (0.33)	0.74 (0.37)	0.12 (0.39)	0.45 (0.47)
C_{r5}	1.10 (0.40)	1.18 (0.37)	0.70 (0.58)	1.10 (0.58)	1.10 (0.40)	1.19 (0.37)	0.70 (0.58)	1.08 (0.57)
C_{s1}	−2.02 (1.20)	−1.77 (0.91)	−1.82 (2.09)	−3.01 (2.72)	−2.26 (1.74)	−1.76 (0.88)	−1.83 (2.12)	−3.10 (3.16)
C_{s2}	−0.89 (0.78)	−0.80 (0.63)	−0.73 (0.91)	−1.62 (1.93)	−1.03 (1.08)	−0.80 (0.60)	−0.70 (0.86)	−1.64 (2.14)
C_{s3}	0.24 (0.30)	0.13 (0.34)	0.06 (0.19)	0.19 (0.57)	0.26 (0.33)	0.13 (0.32)	0.06 (0.18)	0.13 (0.56)
C_{s4}	1.28 (0.91)	1.03 (0.53)	0.83 (0.81)	1.63 (1.38)	1.44 (1.28)	1.03 (0.49)	0.80 (0.77)	1.54 (1.41)
C_{s5}	2.24 (1.53)	1.91 (0.89)	1.80 (1.55)	2.90 (1.95)	2.52 (2.19)	1.89 (0.84)	1.78 (1.49)	2.88 (2.04)

Note. A value preceded by an equal sign indicates that the value was fixed.
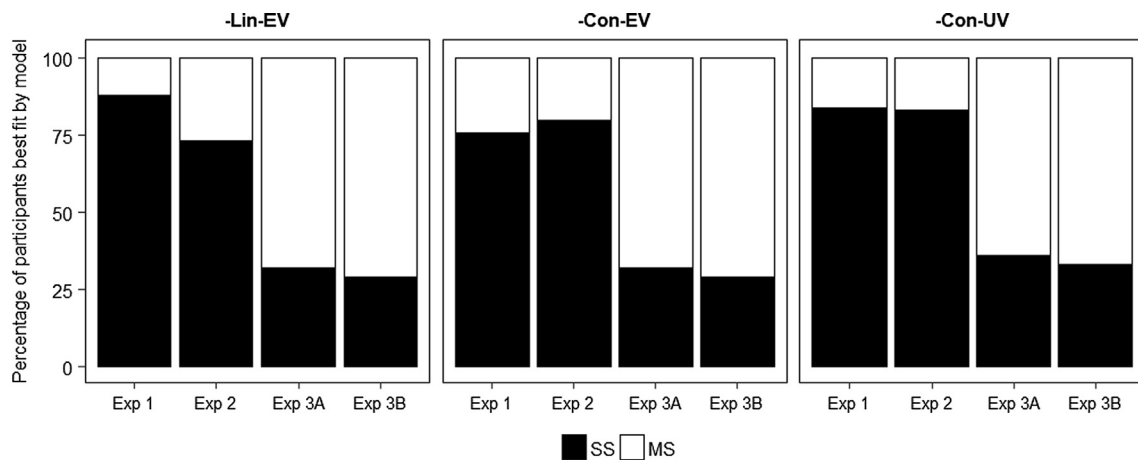
Fig. 5. Percentage of participants best fit by SS compared to MS models for the –Lin-EV models, -Con-EV models, and –Con-UV models according to AIC. SS = single-system model, MS = multiple-systems model, -Lin = linear decision bounds, -Con = converging criteria, -EV = equal variance, -UV = unequal variance.

Discussion of Experiments 3A and 3B). Since both the SS-Lin-EV model and MS-Lin-EV model attempt to model all three bivariate relationships at the same time, the weaker recognition and priming relationship in Experiments 3A and 3B will lead to the MS-Lin-EV model making gains on the SS-Lin-EV model. This is what we observe. A second aspect of the data that both SS-Lin-EV and MS-Lin-EV models struggle to explain is that of the association of recognition and source memory. This relationship has not been central to our investigation here, but the misfits here (see Fig. 7) naturally influence each models' overall ability to predict the observed data. We will discuss this aspect further in the General Discussion.

In sum, neither the SS-Lin-EV or MS-Lin-EV model was able to reproduce the full qualitative pattern of results. Measures of fit tended to favor the SS-Lin-EV model across Experiments 1 and 2, while the MS-Lin-EV model was favored in Experiments 3A and 3B. The SS-Lin-EV model was able to reproduce many of the data patterns, but, crucially for the key investigation here, for items with incorrect source decisions, it incorrectly predicted that identification RTs would decrease as source confidence increases, and this is the opposite pattern to the one observed in Experiments 1-3A/3B.

Recent attempts to model recognition and source memory with signal detection theory indicate two ways in which the SS-Lin-EV model can be modified in order to provide a better account of the data. The first is to allow convergence in the source criteria as recognition confidence increases. The second is to allow old and new item memory strength variance to be unequal. We consider each of these modifications in turn. For comparison, we also applied these modifications to the MS-Lin-EV model.

SS and MS models with converging source criteria and equal variance assumption (SS-Con-EV, MS-Con-EV)

A variety of published data shows that in joint recognition and source memory paradigms, source ratings following high confidence "old" ratings tend to be more confident than those that follow "old" ratings made with lower confidence or "new" ratings (e.g. Slotnick & Dodson, 2005; Onyper et al., 2010; Starns et al., 2013), that is, source ratings change with recognition ratings, rather than vary independently of them. To account for this pattern of data, Hautus et al. (2008) defined a signal detection model of recognition and source memory that implemented non-linear likelihood-ratio decision bounds rather than linear decision bounds that resulted in more closely predicting the observed association of recognition and source ratings. In more recent models, linear approximations of these likelihood-ratio bounds simply defined separate source criteria for some or all recognition ratings, which improves model predictions but necessitates a high number of additional parameters (Onyper et al., 2010; Starns et al., 2014). To

account for the same pattern of data, Klauer and Kellen (2010) adapted guessing parameters in their discrete-state model of recognition and source memory. Here, the idea of converging criteria was implemented by a function that compressed probabilities of responses to the middle of the source scale with lower recognition confidence, such that mid-scale source responses (i.e. guessing the source) were more likely when they were preceded by lower recognition ratings. In contrast to the separate estimation of criteria for all levels of recognition confidence, and thereby increasing the number of estimated parameters, the advantage of this compression function in the discrete-state model is that convergence of responses is implemented with only one additional parameter.

We adapted the compression function (Klauer & Kellen, 2010) to our signal detection based model to implement the idea of converging criteria here with a minimal number of additional parameters. As in the SS-Lin-EV and MS-Lin-EV models, we estimated one set of recognition criteria for all levels of source response. For the baseline source criteria to compress (C_S), we first estimated source criteria for the items given a high-confidence old (6 = "sure-old") rating. We then derived criteria for lower recognition ratings by fanning or de-compressing the criteria, such that with decreasing recognition confidence, mid-scale source ratings (i.e. guess-source responses) become more likely. In addition to estimating source criteria directly for 6 = sure-old, we therefore derived source criteria for 5 = probably-old and 4 = guess-old ratings as well as 3–1 = guess to sure-new ratings. Thus, while we do not assume guessing following "new" ratings (e.g. as in Hautus et al., 2008; Klauer & Kellen, 2010), we assume the same criteria for all "new" responses. This means that we assume that "new" responses still follow from evaluation of a memory strength signal.

Extremity of a confidence rating is quantified as the distance of rating i from the high-confidence sure-old category ($h = 6$), that is $|h - i|$, for $i = \{6, 5, 4, 3\}$ (with 3 as the representative for the 3–1 bin of recognition ratings). The decompressed version of the criteria C_S^* is thus generated by $C_S^* = \exp(|h - i|\lambda)C_S$ (adapted from Klauer & Kellen, 2010). Thus, in the SS and MS models with converging source criteria (henceforth referred to as the SS-Con and MS-Con models, respectively), one additional compression parameter λ for the source criteria is estimated. The maximum likelihood parameter estimates of these models are shown in Table 3.

The effect of converging criteria for the relationship of priming and source memory is striking. Compared to the models with linear decision bounds, both the SS-Con-EV and MS-Con-EV models made gains of on average 70 points by AIC and BIC for each individual's fit. The SS-Con-EV model is now able to predict the association of priming and source memory, such that identification RTs decrease as source confidence increases for *both* correct and incorrect source decisions (Fig. 3, SS-Con-

Table 3

Means and Standard Deviations (in parentheses) of the Parameter Values of the Converging Criteria SS and MS models Across Participants in Experiments 1, 2 and 3A, and 3B.

Parameter	SS-Con-EV				MS-Con-EV			
	Exp 1 $N = 25$	Exp 2 $N = 30$	Exp 3A $N = 25$	Exp 3B $N = 24$	Exp 1 $N = 25$	Exp 2 $N = 30$	Exp 3A $N = 25$	Exp 3B $N = 24$
μ_r	0.99 (0.58)	1.30 (0.66)	0.87 (0.49)	0.73 (0.36)	0.99 (0.54)	1.31 (0.68)	0.85 (0.51)	0.70 (0.36)
μ_p	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	0.99 (0.65)	1.46 (0.57)	1.80 (1.09)	1.42 (1.00)
μ_s	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$
w_{pr}	$=1$	$=1$	$=1$	$=1$	$=0$	$=0$	$=0$	$=0$
w_{ps}	$=1$	$=1$	$=1$	$=1$	$=0$	$=0$	$=0$	$=0$
w_{rs}	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$
b	2178 (480)	2215 (462)	2070 (422)	2193 (488)	2178 (483)	2229 (455)	2163 (442)	2246 (519)
s	205 (115)	200 (80)	209 (100)	194 (112)	$=SS$	$=SS$	$=SS$	$=SS$
$\sigma_{f|new}$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$
$\sigma_{f|old}$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$
σ_r	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$
σ_p	554 (124)	562 (126)	545 (129)	496 (162)	553 (123)	561 (125)	530 (127)	486 (156)
σ_s	2.75 (1.60)	2.14 (0.74)	1.37 (0.71)	3.02 (1.67)	2.71 (1.06)	2.34 (1.20)	1.40 (0.95)	3.99 (3.56)
C_{r1}	−1.02 (0.59)	−0.43 (0.61)	−1.54 (1.42)	−1.60 (1.14)	−1.01 (0.58)	−0.43 (0.62)	−1.59 (1.53)	−1.62 (1.14)
C_{r2}	−0.20 (0.50)	0.20 (0.51)	−0.51 (0.88)	−0.71 (0.70)	−0.19 (0.50)	0.21 (0.52)	−0.52 (0.90)	−0.72 (0.69)
C_{r3}	0.24 (0.35)	0.56 (0.44)	−0.08 (0.51)	−0.01 (0.59)	0.24 (0.34)	0.57 (0.44)	−0.10 (0.53)	−0.03 (0.58)
C_{r4}	0.64 (0.39)	0.89 (0.43)	0.22 (0.45)	0.55 (0.51)	0.65 (0.38)	0.90 (0.43)	0.21 (0.47)	0.54 (0.51)
C_{r5}	1.27 (0.42)	1.36 (0.42)	0.81 (0.63)	1.20 (0.62)	1.28 (0.42)	1.37 (0.42)	0.80 (0.64)	1.17 (0.62)
C_{s1}	−2.18 (1.73)	−1.64 (1.00)	−1.49 (1.94)	−2.77 (1.80)	−2.38 (1.95)	−1.76 (1.22)	−1.55 (1.97)	−3.72 (3.43)
C_{s2}	−0.76 (0.98)	−0.58 (0.49)	−0.44 (0.61)	−1.00 (0.86)	−0.75 (0.94)	−0.65 (0.61)	−0.49 (0.81)	−1.29 (1.43)
C_{s3}	0.19 (0.22)	0.15 (0.32)	0.05 (0.14)	0.15 (0.68)	0.24 (0.29)	0.13 (0.35)	0.05 (0.13)	0.17 (0.72)
C_{s4}	1.27 (1.02)	0.94 (0.62)	0.59 (0.58)	1.11 (1.36)	1.37 (1.19)	1.00 (0.64)	0.61 (0.65)	1.38 (1.80)
C_{s5}	2.96 (1.99)	2.00 (1.09)	1.65 (1.33)	2.69 (1.98)	3.16 (2.35)	2.12 (1.28)	1.71 (1.55)	3.43 (3.24)
λ	0.55 (0.25)	0.43 (0.23)	0.53 (0.36)	0.60 (0.38)	0.55 (0.26)	0.43 (0.25)	0.52 (0.35)	0.61 (0.40)

Note. A value preceded by an equal sign indicates that the value was fixed.

EV, left-hand-side). Additionally, the SS-Con-EV model still captures the decrease in identification RTs with increasing recognition confidence (Fig. 4, SS-Con-EV). The MS-Con-EV model still incorrectly predicts that identifications RTs do not vary with source confidence (Fig. 3, MS-Con-EV) or recognition confidence (Fig. 4, MS-Con-EV). This may make the quantitative gains in fit surprising for the MS-Con-EV relative to the MS-Lin-EV model. These gains are primarily due to the better fit of the association of recognition and source memory compared to the MS-Lin-EV model (Fig. 7). We will briefly discuss that aspect of the data later.

Comparing SS-Con-EV and MS-Con-EV models replicates the overall pattern shown in the comparison of the SS-Lin-EV and MS-Lin-EV models. The SS-Con-EV model is clearly preferred over the MS-Con-EV model in Experiments 1 and 2, while the reverse is the case in Experiment 3A, both when considering summed log-likelihoods (Table 5, middle rows) and percentage of participants best fit by each of the models (Fig. 5, -Con-EV panel). For Experiment 3B, both models perform equally well when examining the summed log-likelihoods but a greater percentage of participants is best fit by the MS-Con-EV model, with differences between models marginal even on an individual basis. As with the previous models, examination of predicted patterns of data compared to observed ones (see the Supplemental Material for a direct contrast) suggests that the relationship of priming and source is reasonably well described in Experiments 3A and 3B while the relationship with recognition is not, leading to the misfits of the SS-Con-EV model overall for these experiments.

SS and MS models with converging source criteria and unequal variance assumption (SS-Con-UV, MS-Con-UV)

While implementing converging criteria significantly improved the SS model's fit to the data, we modified a final aspect of the model to bring it further in line with current signal detection models of recognition and source memory. The models implemented above and in Berry et al. (2012) assume that the variance of f is the same for old and new items. Models of recognition typically allow for unequal variance in the old and new item evidence distributions (Hautus et al., 2008; Ratcliff, Sheu, & Gronlund, 1992; Starns et al., 2014; Yonelinas, 1994). In the unequal variance signal detection model, for example, the new item distribution is set as the reference distribution with mean strength 0 and variance 1; the mean strength and variance of the old item distribution is free to vary. Thus, in this final modification of the SS (and MS model), we assumed converging criteria and separate σ_f for old and new item distributions, such that for new items $\sigma_{f|new} = \sigma_r = 1/\sqrt2$, while for old items $\sigma_r = 1/\sqrt2$, with $\sigma_{f|old}$ allowed to vary. Thus, in the SS and MS models with converging source criteria and unequal old-new item strength variance (henceforth referred to as the SS-Con-UV and MS-con-uv models, respectively), one additional parameter ($\sigma_f|_{old}$) is estimated for the standard deviation of the old item strength distribution. The parameter estimates are shown in Table 4. Fig. 8 shows a graphical representation of the SS-Con-UV model for all three pairwise relationships (priming – source memory, priming – recognition memory, recognition memory – source memory) based on parameter estimates from the model fitted to Experiment 2.

The unequal variance modification affected the fit of the SS model to the priming and recognition relationship as well as the one between priming and source memory, with only minimal changes in the fit of the relationship of recognition and source memory (Fig. 7, -Con-UV panels). As Fig. 3 (SS-Con-UV, left-hand-side) shows, the decrease of identification RTs with source confidence is steeper for correct than incorrect source decisions, while Fig. 4 (SS-Con-UV) shows that identification RTs decrease more steeply with increasing recognition confidence for old items than new items. As with the previous instantiations of the MS model, the MS-Con-UV model predicts that identification RTs do not vary with source or recognition confidence (Figs. 3 and 4, MS-Con-UV), but as with the SS-Con-UV model, the fit of the association of recognition and source memory improves. Quantitatively, the SS-Con-UV model outperforms the MS-Con-UV model in Experiments 1 and 2 (Table 5, bottom rows). In Experiments 3A, the MS-Con-UV model outperforms the SS-Con-UV model overall and by

Table 4
Means and Standard Deviations (in parentheses) of the Parameter Values of the Unequal Variance Converging Criteria SS and MS Models Across Participants in Experiments 1, 2, 3A, and 3B.

Parameter	SS-Con-UV				MS-Con-UV			
	Exp 1 $N = 25$	Exp 2 $N = 30$	Exp 3A $N = 25$	Exp 3B $N = 24$	Exp 1 $N = 25$	Exp 2 $N = 30$	Exp 3A $N = 25$	Exp 3B $N = 24$
μ_r	1.30 (1.03)	1.97 (1.45)	1.13 (0.98)	0.81 (0.41)	1.32 (0.96)	1.93 (1.37)	1.16 (1.03)	0.75 (0.39)
μ_p	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	1.47 (1.23)	2.43 (1.43)	2.28 (1.51)	1.53 (0.85)
μ_s	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$	$=\mu_r$
w_{pr}	$=1$	$=1$	$=1$	$=1$	$=0$	$=0$	$=0$	$=0$
w_{ps}	$=1$	$=1$	$=1$	$=1$	$=0$	$=0$	$=0$	$=0$
w_{rs}	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$	$=1$
b	2168 (481)	2205 (467)	2062 (424)	2188 (484)	2178 (481)	2228 (454)	2169 (441)	2245 (518)
s	169 (124)	138 (72)	234 (196)	174 (121)	$=$SS	$=$SS	$=$SS	$=$SS
$\sigma_{f\mid new}$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$
$\sigma_{f\mid old}$	1.17 (0.43)	1.47 (0.59)	0.97 (0.63)	0.87 (0.34)	1.18 (0.37)	1.45 (0.50)	1.01 (0.61)	0.81 (0.32)
σ_r	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$	$=1/\sqrt2$
σ_p	557 (122)	565 (125)	540 (134)	496 (164)	553 (121)	563 (124)	518 (145)	489 (159)
σ_s	2.99 (1.52)	2.71 (1.13)	1.50 (0.80)	3.17 (1.69)	3.27 (1.42)	3.26 (1.85)	1.54 (0.89)	3.46 (2.61)
C_{r1}	−1.11 (0.61)	−0.48 (0.65)	−1.74 (1.79)	−1.78 (1.45)	−1.10 (0.61)	−0.48 (0.65)	−1.85 (1.93)	−1.68 (1.17)
C_{r2}	−0.22 (0.51)	0.23 (0.62)	−0.62 (1.29)	−0.75 (0.72)	−0.22 (0.52)	0.23 (0.59)	−0.62 (1.29)	−0.75 (0.71)
C_{r3}	0.25 (0.51)	0.65 (0.57)	−0.08 (0.60)	−0.02 (0.62)	0.25 (0.37)	0.64 (0.52)	−0.08 (0.58)	−0.04 (0.62)
C_{r4}	0.70 (0.44)	1.08 (0.62)	0.26 (0.55)	0.58 (0.54)	0.72 (0.43)	1.06 (0.57)	0.27 (0.55)	0.56 (0.54)
C_{r5}	1.47 (0.57)	1.75 (0.74)	0.94 (0.83)	1.27 (0.67)	1.50 (0.55)	1.72 (0.63)	0.99 (0.90)	1.23 (0.67)
C_{s1}	−2.42 (1.61)	−2.31 (1.25)	−1.48 (1.44)	−2.90 (1.96)	−2.90 (2.47)	−2.77 (2.04)	−1.75 (2.15)	−3.27 (2.77)
C_{s2}	−0.83 (0.94)	−0.83 (0.74)	−0.40 (0.46)	−1.05 (1.00)	−0.96 (1.25)	−1.01 (1.03)	−0.50 (0.70)	−1.19 (1.21)
C_{s3}	0.25 (0.28)	0.20 (0.45)	0.06 (0.14)	0.17 (0.56)	0.30 (0.32)	0.18 (0.52)	0.06 (0.13)	0.13 (0.56)
C_{s4}	1.50 (1.15)	1.33 (0.81)	0.58 (0.46)	1.24 (1.36)	1.72 (1.57)	1.49 (1.06)	0.66 (0.57)	1.27 (1.41)
C_{s5}	3.47 (2.26)	2.83 (1.37)	1.80 (1.59)	2.90 (2.21)	4.01 (3.19)	3.21 (1.97)	2.05 (2.34)	3.09 (2.44)
λ	0.53 (0.25)	0.38 (0.19)	0.52 (0.35)	0.59 (0.36)	0.53 (0.24)	0.39 (0.24)	0.50 (0.33)	0.60 (0.39)

*Note.* A value preceded by an equal sign indicates that the value was fixed.

percentage of participants best fit by each model, while in Experiment 3B both models perform equally well overall, while the SS-Con-UV model edges past the MS-Con-UV model in the percentage of participants' best fit by the model. As mentioned previously, this better performance of the MS-Con-UV model in Experiments 3A and 3B likely reflects the weaker association between recognition and priming in this experiment. Thus, while the SS-Con-UV model better accounts for the relationship of priming and source memory observed throughout all experiments, the MS-Con-UV model makes gains for those experiments where the association of priming and recognition memory was attenuated (i.e., Experiments 3A and 3B).

Comparing the AIC and BIC of all the models in each experiment, for Experiments 1 and 2, the SS-Con-UV model provides the best fit overall, and by a substantial margin. For Experiment 3A, the MS-Con-EV model provides the best fit overall. For Experiment 3B, the MS-Con-EV provides a marginally better fit than the SS-Con-EV model, by a fraction of a point, with the unequal variance assumption not improving the fit. Comparing the SS models by AIC for percentage of participants best fit (Fig. 6) shows that the SS-Con-UV model fits the data in Experiments 1, 2 and 3A better than the simpler models, while in Experiment 3B, the majority of participants are better described by the SS-Con-EV model. The pattern of results is similar for comparison by BIC though weakened, with SS-Con-EV and SS-Con-UV equally preferred in Experiments 1 and 3A.
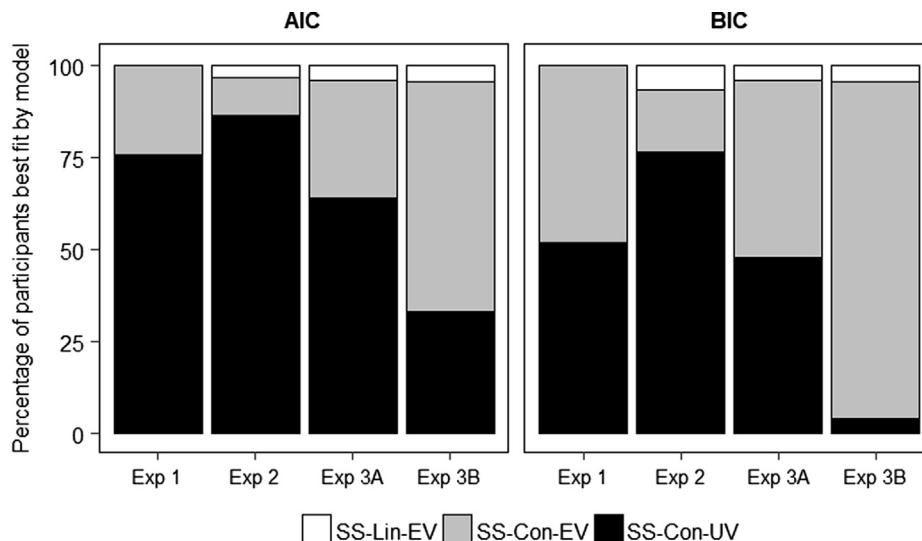
Fig. 6. Percentage of participants best fit by the single-system models in Experiments 1-3A/B according to the AIC and BIC. SS = single-system model, MS = multiple-systems model, -Lin = linear decision bounds, -Con = converging criteria, -EV = equal variance, -UV = unequal variance.
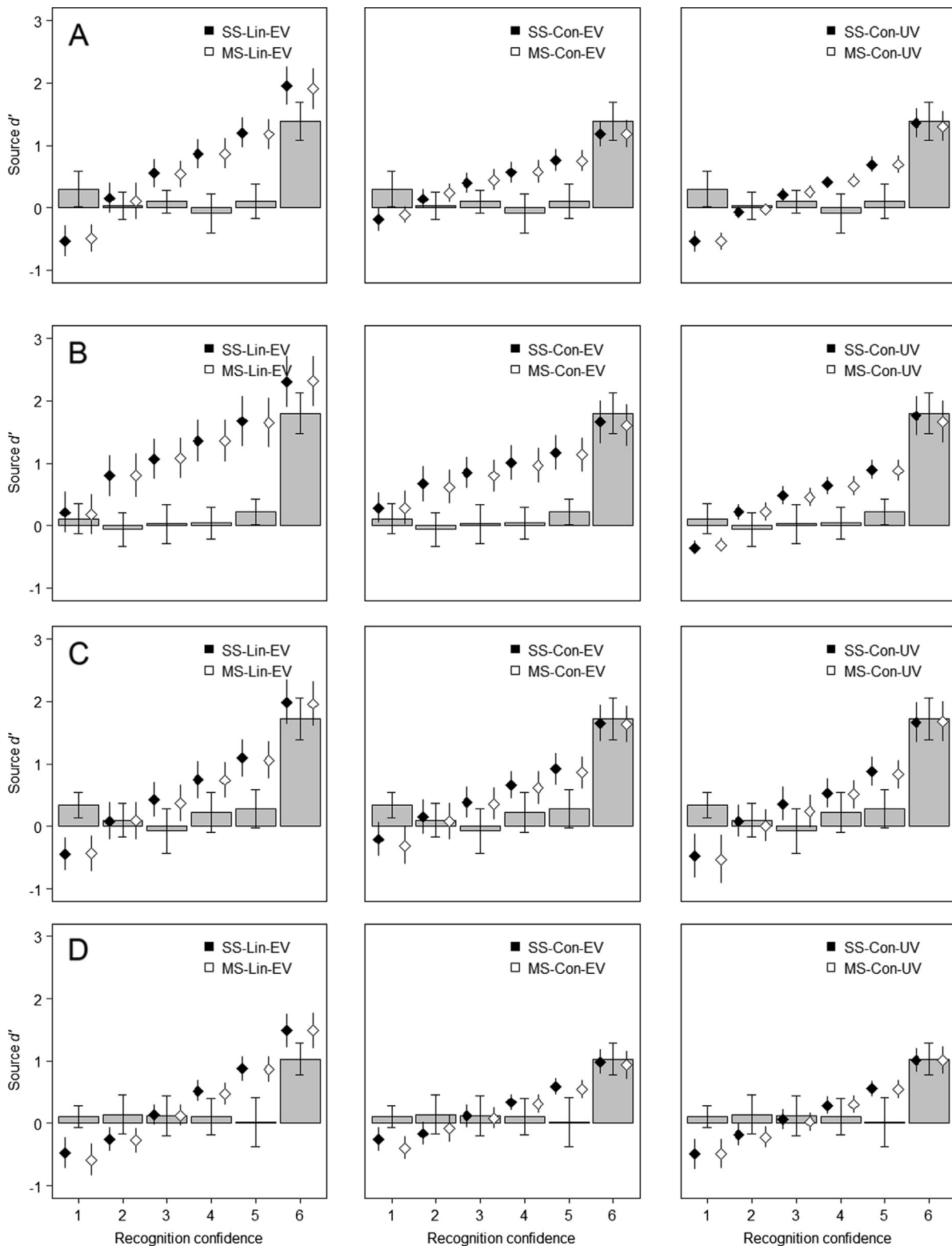
Fig. 7. Source *d′* across recognition confidence for the observed data (bars) and data predicted by the models (points) for the -Lin-EV, -Con-EV and -Con-UV models for Experiment 1 (A), Experiment 2 (B), Experiment 3A (C), and Experiment 3B (D). The model predictions were generated by using each participant's maximum likelihood parameter estimates to simulate 25,000 old and 25,000 new items. SS = single-system model, MS = multiple-systems model, -Lin = linear decision bounds, -Con = converging criteria, -EV = equal variance, -UV = unequal variance.

General discussion

In this article we provided an initial attempt at linking priming, recognition and source memory within a single experimental and modeling framework, by extending the work of Berry et al. (2012). In four experiments, we replicated the associations between priming and recognition reported by Berry et al. (2012, Experiment 2) and, moreover, found that the magnitude of the priming effect is also linked to
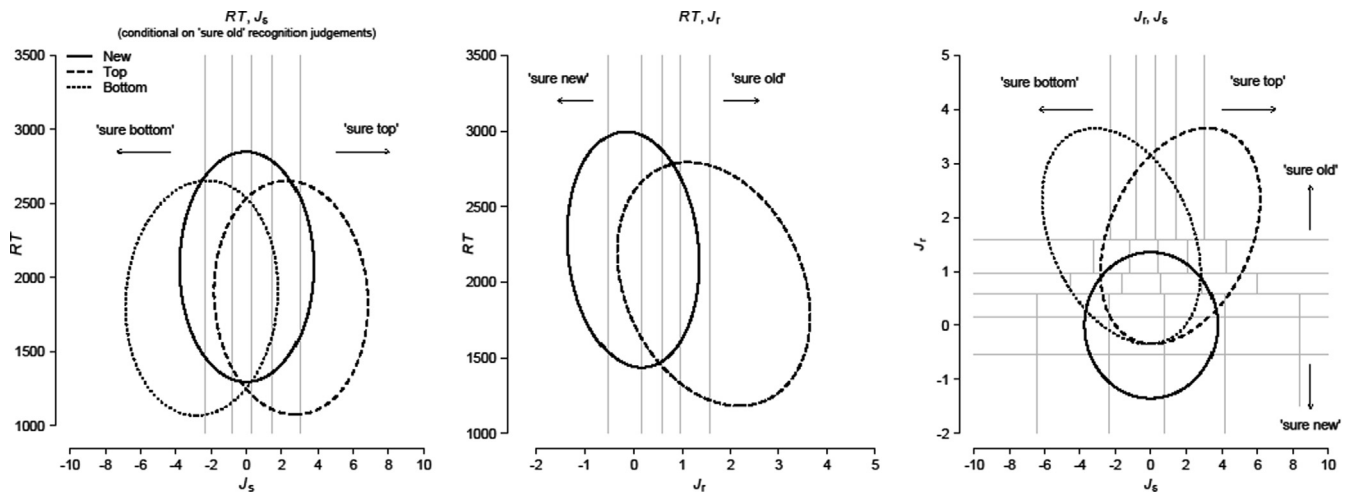
Fig. 8. Graphical representation of the SS-Con-UV model, based on mean parameter estimates from Experiment 2 (Table 4) for the relationship of *RT* and source memory (J_s, *RT*), *RT* and recognition memory (J_r, *RT*) and recognition memory and source memory (J_r, J_s). Gray lines indicate recognition and source criteria. The *RT*, J_s panel shows the joint distribution of *RT* and J_s, conditional on 'sure old' recognition judgments. The ellipses represent contours of equal probability in a bivariate normal distribution.

source memory decisions. Specifically, in all experiments, the magnitude of the priming effect was reliably greater for items with correct source decisions than those with incorrect source decisions; priming was also linked to confidence in the source decision in that identification RTs tended to be faster as confidence in the source decision increased, regardless of whether the source decision was correct or incorrect. The association between priming and source decisions was present even when priming was measured in a distinct phase to recognition and source ratings (Experiments 3A and 3B), and so is not an artefact of measuring priming, recognition and source decisions concurrently. The association also persisted despite variations in (1) overall source discriminability (i.e., Experiment 1 vs. Experiment 2), and (2) the strength of the association between priming and recognition across experiments (i.e., Experiments 1 and 2 vs. Experiments 3A and 3B), and so does not appear to be driven by this association. This is consistent with priming and source memory being driven by a single signal rather than multiple distinct memory signals.

However, it is possible that other processes contributed to the association of priming and source memory we observed. For one, it may be the case that some words were simply more memorable than others with specific orthographic or lexical characteristics. We did not model item effects directly, so cannot rule out that such item effects

contributed to the association we observed here. However, Sheldon and Moscovitch (2010) showed that the association they observed between recollection and priming (using remember-know and lexical decision tasks respectively) could not be attributed to item effects. This suggests that the association here between priming and source memory is also unlikely to arise from item effects alone.

Second, participants were informed prior to the test phase that half the words they would be asked to identify had been studied while half would be new. This means that they could rely on their memory to reduce the number of candidate answers to complete the identification task. We believe this is unlikely. Ward, Berry, and Shanks (2013) showed that even under circumstances that would be optimal for relying on memory (participants were informed before each trial if the upcoming item had been studied), identification performance was unaffected. This suggests that the observed associations are not due to strategic uses of memory during identification.

Analysis of the behavioral data thus suggests that a single memory signal could be driving responding in the priming and source memory task, as well as the recognition task and formal modeling supports this conclusion to a degree. An extended version of the SS model in Berry et al. (2012) correctly predicted that the priming effect would be greater for items with correct versus incorrect source decisions, and also

Table 5
Goodness-of-Fit for the Models in Experiments 1, 2, 3A, and 3B, Summed Across Participants.

Model	p	Experiment 1 (N = 25)*			Experiment 2 (N = 30)*			Experiment 3A (N = 25)*			Experiment 3B (N = 24)*		
		ln(L)	AIC	BIC	ln(L)	AIC	BIC	ln(L)	AIC	BIC	ln(L)	AIC	BIC
SS-Lin-EV	15	87,510	175,771	178,400	105,642	212,185	215,429	64,538	129,826	132,352	26,837	54,393	56,498
MS-Lin-EV	15	87,647	176,043	178,672	105,779	212,458	215,703	64,498	129,747	132,274	26,833	54,385	56,490
SS-Con-EV	16	86,135	173,071	175,875	104,601	210,161	213,622	63,668	128,136	130,832	26,420	53,608	55,853
MS-Con-EV	16	86,249	173,297	176,102	104,720	210,400	213,861	63,631	128,061	130,757	26,420	53,607	55,853
SS-Con-UV	17	86,007	172,863	175,843	104,343	209,706	213,383	63,565	127,980	130,843	26,406	53,628	56,013
MS-Con-UV	17	86,137	173,124	176,103	104,497	210,014	213,691	63,527	127,903	130,767	26,406	53,627	56,013

Note. The Akaike information criterion (AIC) is calculated as AIC $= -2\ln(L) + 2P$, where $P = p \times N$ is the total number of free parameters for each fit, where p is the number of free parameters for each model, and N is the number of participants modeled in each experiment. The Bayesian information criterion (BIC) is calculated as BIC $= -2\ln(L) + P$, where $P = p \times \ln(d)$ is the number of free parameters for each model, where p is the number of free parameters for each model and d is the number of trials modeled across participants. A smaller ln(L), AIC or BIC value indicates a relatively better model fit. The average number of fitted trials per participant was in Experiment 1: $n = 328$ (SD $= 13$) for n total $= 8196$; Experiment 2: $n = 333$ (SD $= 8$) for n total $= 9996$; Experiment 3A: $n = 249$ (SD $= 9$) for n total $= 6237$; Experiment 3B: $n = 107$ (SD $= 4$) for n total $= 2557$. SS = single system, MS = multiple system, -Lin = linear criteria bounds, -Con = converging criteria, -EV = equal variance, -UV = unequal variance.

* Participants were excluded for whom model fits resulted in extreme parameter estimates for at least one of the models. The exclusions did not affect the pattern of results; see the Supplemental Material for the table including all participants and tables excluding participants for pairs of models and SS models only.

that identification RTs would decrease with increasing source confidence. However, this model incorrectly predicted that the identification RTs of items with incorrect source decisions would be slower as confidence in the source decision increased. In this model, the decision criteria used for source decisions were assumed to be invariant across recognition ratings (e.g. as in DeCarlo, 2003), which gave rise to this incorrect prediction. In order to bring the model more into line with current signal detection modeling of recognition and source ratings (Hautus et al., 2008; Onyper et al., 2010; Starns et al., 2013, 2014), we modified it to allow the source decision criteria to converge as recognition confidence increases (i.e., in the SS-Con-EV model). This drastically improved the fit of the model, which then correctly predicted that identification RTs decrease as source confidence increases for both correct and incorrect source decisions.

Allowing source criteria to converge with increasing recognition confidence is one way of representing a dependency between recognition and source ratings. Specifically, items receiving high confidence old recognition ratings will also tend to receive high confidence source decisions. With this modification, a source rating made with high confidence, even if incorrect, is relatively likely to be preceded by a high confidence recognition rating, which in turn is likely to have a relatively fast identification RT. This gives rise to the trend for identification RTs to decrease as source confidence increases for incorrect source decisions, and suggests that this trend critically depends upon the association between recognition and source ratings. The single-system model predictions and fit were improved even further by allowing the variance of the strength signal for old items to differ from that of new items (i.e., in the SS-Con-UV model), which again brings the single-system model more in line with existing signal detection accounts of recognition (e.g. Ratcliff et al., 1992; Rotello, 2017; Wixted, 2007).

As a baseline to compare the single-system models against, we implemented "multiple-systems" versions of these models in which the signal driving priming is independent of the one driving recognition and source decisions, implementing a theory in the literature that implicit and explicit memory operate with stochastic and functional independence (Squire & Dede, 2015; Tulving et al., 1982). These models could not reproduce the aforementioned associations between priming and source decisions; the quantitative fit of these models to the data also tended to be inferior to that of the SS-Con-EV and SS-Con-UV models. Of course, this does not rule out a multiple-systems account altogether. Distinct memory systems or signals may drive recognition, priming and source, but our results suggest that if this is the case then these signals are highly interrelated, not independent, in contrast to assumptions of stochastically and functionally independent implicit and explicit memory systems (e.g., Squire & Dede, 2015). This would be in line with dual-process models that do not implement stochastic independence (e.g., the MS2 model in Berry et al., 2012; DPSD models in Moran & Goshen-Gottstein, 2015; Pratte & Rouder, 2011). In order to implement a multiple-systems model with highly interrelated systems, we could implement a model in which the signals driving priming, recognition, and source (i.e., f_p, f_r, and f_s) can all be correlated to various degrees (i.e., by allowing w_{pr}, w_{ps}, w_{rs} to all be free to vary). This model would be a more general case of the single-system model (in which $f_p = f_r = f_s$ and $w_{pr} = w_{ps} = w_{rs} = 1$) and the multiple-systems models (in which $w_{pr} = w_{ps} = 0$). A problem with this model is that it is not at all clear that it would be falsifiable—it would not make firm predictions in advance because it can mimic the single-system and multiple-systems models, producing any result that these models can. The MS2 model in Berry et al. (2012) fit the data well by closely mimicking the single-system model, and its greater flexibility meant that, although it could fit the data more closely, it appears to achieve this by overfitting (Shanks & Berry, 2012). Thus, although our findings could also be explained by a multiple-systems model with highly interrelated systems, for the sake of parsimony, we think a single-system interpretation of our findings should be preferred compared to such an

interrelated, multiple-systems model. This has implications for, for example, the debate concerning the origin of the FN400 (e.g. Strozak et al., 2016; Voss & Federmeier, 2011). This debate concerns whether that waveform reflect repetition priming or familiarity. The results here and in Berry et al. (2012) suggest that this debate is based on a false dichotomy.

While we clearly favour the single-system model to account for the data, such a model would need to fit not only the priming and source memory relationship but also account for the other bivariate relationships. While the single-system model we presented here did well in accommodating the association of priming and source memory for all experiments, it fared less well relative to the multiple-systems model in terms of its quantitative fit to the data when applied to Experiments 3A and 3B. Those experiments show an attenuated association of priming and recognition memory, due to the changes in experimental design. While the association persists overall, the multiple systems model we implemented here was relatively better able to accommodate the overall pattern of data in those experiments than in Experiment 1 and 2. A model of performance in these three memory tasks needs to be able to account for these variations in data but neither the SS or the MS model was able to do that for all patterns of data.

Association of recognition and source memory

Although our focus has been on the link between priming and source memory, and to a lesser degree priming and recognition memory, our empirical and modeling results also have relevance for the ongoing discourse concerning how recognition and source memory should be modeled (for a review, see Rotello, 2017). The single-system and multiple-systems models we tested do not make distinct predictions for this relationship. In both models, we assumed the same memory signal to underlie recognition and source memory. Thus, both types of models predict that source memory should be higher for recognized than unrecognized items, and source memory should monotonically increase with increasing recognition confidence.

Nevertheless, we want to briefly relate the pattern of the observed association of recognition and source memory to the model predictions. As already reported, in all four experiments, source memory was higher for recognized than unrecognized items. However, the shape of the association is not closely predicted by the models (see Fig. 7 and a more detailed analysis in the Supplemental Materials). First, source memory only clearly exceeded chance for items recognized with high confidence. Second, the models frequently predicted sub-chance source performance for items with lower recognition ratings (see Fig. 7), but this trend was not found in any of the experiments.

The lack of graded source memory with increasing recognition is well known (e.g. Slotnick & Dodson, 2005). Although this could be argued to indicate threshold rather than continuous processes, continuous models can account for this pattern of data (e.g. Slotnick & Dodson, 2005; Rotello, 2017).

The curious prediction of sub-chance performance, that is, participants reliably respond source-A for source-B items for low recognition confidence ratings (evident in Fig. 7), has similarly been discussed previously (e.g. Hautus et al., 2008; Starns et al., 2008). This prediction follows from the distributional assumptions of the signal detection models. Since participants typically need to provide source ratings even after making "sure-new" responses, that pattern is theoretically observable. However, realistically these data points are likely noisy due to participants guessing or responding with random, not memory-driven, responses. Some of this noise can be reduced experimentally by providing participants with an option for 'guessing' as the mid-point on a scale with an uneven number of rating categories (e.g. Onyper et al., 2010; Slotnick & Dodson, 2005). Even so, subchance source performance for unrecognized items is not observed reliably (Malejka & Bröder, 2015; Starns et al., 2008). Formal models of recognition and source memory performance are therefore frequently modified to avoid

this prediction by implementing source guessing for unrecognized items (Hautus et al., 2008, Model 2) or bounded distributions that avoid crossover of correlated distributions and the following predictions of sub-chance performance (Starns et al., 2014). In our experimental set-up and models, we made no such assumptions, given that the relationship of recognition and source memory was not the focus of our theoretical interest. Under a modified model, with formal implementation of source guessing for unrecognized items, we would expect chance source performance for items given 'new' responses, with graded, increasing source memory with increasing confidence in 'old' responses. This modification would likely provide a closer fit of the model to the association of recognition and source memory than the current model implementation does.

Limitations and future research

One potential limitation of our investigation concerns the way in which identification RTs were modeled. For computational simplicity, we assumed all variables in the models, including identification RTs, are normally distributed. Although identification RTs were generally normally distributed in the experiments in Berry et al. (2012), this was only the case for a minority (< 20%) of participants in the experiments in the present paper. This leaves the possibility that the observed associations and resulting modeling results are an artefact of this distributional analysis. We addressed this in two ways. We repeated the statistical analyses twice: First, we replaced per-participant mean identification RTs with mean log-transformed identification RTs and also per-participant mean identification RTs with median identification RTs (see the Supplemental Materials). The pattern of results and substance of the effects in those analyses was comparable to the ones we reported in the paper. Second, we also fitted the SS-Con-UV and MS-Con-UV model to all experiments using log-transformed identification RTs to replace identification RTs in ms. The results broadly followed the same pattern as the model fits for the data with untransformed identification RTs (see the Supplemental Materials). This demonstrates that our main conclusions hold, even though the identification RTs in our CID-RS task are not normally distributed.

The modeling framework we presented here is very much based in multi-dimensional signal-detection theory models of recognition and source memory (e.g. DeCarlo, 2003; Hautus et al., 2008; Slotnick & Dodson, 2005; Starns et al., 2013). While some of our choices in the implementation (e.g. relative judgments for source decisions, converging criteria) have not been explicitly implemented in the same way in those models, the modeling framework here is in its essence as much of an extension of those multi-dimensional recognition and source models to priming as it is of the priming and recognition model in Berry et al. (2012) to source memory. This is not to say that this is the only way to model these data. Recently, signal-detection and discrete-state models have been extended to model reaction times of recognition and source decisions using diffusion processes (e.g. Dube, Starns, Rotello, & Ratcliff, 2012; Kellen, Singmann, Vogt, & Klauer, 2015; Starns, 2014). While, to our knowledge, RT distributions have not been used as the basis for the simultaneous modeling of recognition and source decisions, one can easily imagine a model that incorporates response times for all three tasks (priming, recognition memory, source memory).

One outstanding question is whether a major dual-process model of recognition and source (Yonelinas, 1994, 1999, 2002) could be extended to priming, and how it would explain our experimental findings; a related issue is whether recollection and familiarity make independent contributions to the recognition of an item. Berry et al. (2012) extended the dual-process model of recognition (Yonelinas, 1994) to priming in the form of the DPSD1 model. The DPSD1 model assumes that the same memory strength signal (i.e., familiarity, or f in the single-system model) drives priming and recognition performance, but that an independent threshold recollection process also contributes to recognition. This model provided a good account of the relation

between priming and recognition confidence ratings (Berry et al., 2012, Experiment 2), but not the relation between priming and remember-know judgments (see Berry et al., 2012, Experiment 3). Rather than predicting that the magnitude of the priming effect would be greater for *remember* responses than *know* responses, as was observed empirically, the DPSD1 model incorrectly predicted the opposite pattern—namely that the priming effect for *remember* responses would be smaller than that of *know* responses. It made this incorrect prediction because, according to the model, *remember* judgments are driven by recollection, and whether an item is recollected or not is independent of an item's familiarity (i.e., because stochastic independence between recollection and familiarity was assumed; Yonelinas, 2002). A recollected old item will therefore not necessarily have a relatively fast identification RT in the DPSD1 model. If source decisions are being driven solely by recollection, as in early applications of the dual-process theory model (Yonelinas, 1999), such a model would similarly fail to predict the faster identification RTs for correct source decisions we observed in the experiments here and predicted with a single system model.

More recent applications, however, allow for a greater contribution of familiarity to source memory decisions (e.g. Diana, Yonelinas, & Ranganath, 2008). If priming and familiarity-based responding are assumed to depend upon the same memory strength variable, and if source decisions in our Experiments 1–3A/3B were driven by familiarity, rather than recollection, then a dual-process account of our findings would be equivalent to that of the single-system model and would be able to predict the association of priming and source memory we observed. Yet, source memory for spatial location is typically associated with hippocampal activation (Ekstrom and Bookheimer, 2007; Slotnick & Thakral, 2013), which is in turn associated with recollection (e.g., Brown & Aggleton, 2001; Diana, Yonelinas, & Ranganath, 2007; Diana et al., 2010; Eichenbaum, Yonelinas, & Ranganath, 2007), leaving little room to explain the association of priming and source memory as familiarity-based in a dual-process context (though see Taylor & Henson, 2012).

A direct test of the predictions of the single-system model compared to an extended, multivariate dual-process instantiation would be to use source memory tasks that are thought to differ in the extent to which they rely upon the familiarity and recollection processes. Recent work within the dual-process framework suggests that familiarity contributes more to source performance if item and source information are encoded as a single unit (i.e. are unitized), compared to when they are not unitized, where conversely source memory is argued to be based on recollection to a greater degree (e.g., Bastin et al., 2013; Diana et al., 2008; Diana, Van den Boom, Yonelinas, & Ranganath, 2011). The single-system model we propose here, where we assume a single memory signal drives performance, would predict the same pattern of association between priming and source memory for both unitized and non-unitized stimuli. A dual-process, multivariate extension of the single system model (following Starns et al., 2014) would instead predict an association for priming and source memory only for unitized stimuli, where source memory would arguably rely on the same memory signal as priming. We aim to test these predictions in future investigations.

Conclusion

We found that priming is linked to source memory for spatial location. This extends previous work that established a link between priming and recognition, which we also replicated here (Berry et al., 2012). A single-system signal detection model that allows for convergence in source criteria with increasing recognition confidence explains these associations well, and tended to outperform a "multiple-systems" version of the model in which the signals driving priming and source memory are independent. While neither the single-system or the multiple-systems models we tested are able to account for all aspects of the data we presented here, our work provides a new basis to explore

the relationship between priming, recognition memory and source memory.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jml.2019.104039.

References

Addante, R. J. (2015). A critical role of the human hippocampus in an electrophysiological measure of implicit memory. *NeuroImage, 109*, 515–528. https://doi.org/10.1016/j.neuroimage.2014.12.069.

Baddeley, A., Eysenck, M. W., & Anderson, M. C. (2014). *Memory.* Hove: Psychology Press.

Banks, W. P. (2000). Recognition and source memory as multivariate decision processes. *Psychological Science, 11*, 267–273. https://doi.org/10.1111/1467-9280.00254.

Bastin, C., Diana, R. A., Simon, J., Collette, F., Yonelinas, A. P., & Salmon, E. (2013). Associative memory in aging: The effect of unitization on source memory. *Psychology and Aging, 28*, 275–283. https://doi.org/10.1037/a0031566.

Berry, C. J., Henson, R. N., & Shanks, D. R. (2006). On the relationship between repetition priming and recognition memory: Insights from a computational model. *Journal of Memory and Language, 55*, 515–533. https://doi.org/10.1016/j.jml.2006.08.008.

Berry, C. J., Kessels, R. P., Wester, A. J., & Shanks, D. R. (2014). A single-system model predicts recognition memory and repetition priming in amnesia. *The Journal of Neuroscience, 34*, 10963–10974. https://doi.org/10.1523/jneurosci.0764-14.2014.

Berry, C. J., Shanks, D. R., & Henson, R. N. (2008a). A single-system account of the relationship between priming, recognition, and fluency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 97–111.

Berry, C. J., Shanks, D. R., & Henson, R. N. A. (2008b). A unitary signal-detection model of implicit and explicit memory. *Trends in Cognitive Sciences, 12*, 367–373. https://doi.org/10.1016/j.tics.2008.06.005.

Berry, C. J., Shanks, D. R., Li, S., Rains, L. S., & Henson, R. N. (2010). Can "pure" implicit memory be isolated? A test of a single-system model of recognition and repetition priming. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale, 64*, 241–255.

Berry, C. J., Shanks, D. R., Speekenbrink, M., & Henson, R. N. (2012). Models of recognition, repetition priming, and fluency: Exploring a new framework. *Psychological Review, 119*, 40–79. https://doi.org/10.1037/a0025464.

Berry, C. J., Ward, E. V., & Shanks, D. R. (2017). Does study duration have opposite effects on recognition and repetition priming? *Journal of Memory and Language, 97*, 154–174. https://doi.org/10.1016/j.jml.2017.07.004.

Brown, M. W., & Aggleton, J. P. (2001). Recognition memory: What are the roles of the perirhinal cortex and hippocampus? *Nature Reviews Neuroscience, 2*, 51–61. https://doi.org/10.1038/35049064.

Buchner, A., & Wippich, W. (2000). On the reliability of implicit and explicit memory measures. *Cognitive Psychology, 40*, 227–259. https://doi.org/10.1006/cogp.1999.0731.

Coltheart, M. (1981). The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology, 33*, 497–505. https://doi.org/10.1080/14640748108400805.

Craik, F. I. M., Moscovitch, M., & McDowd, J. M. (1994). Contributions of surface and conceptual information to performance on implicit and explicit memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(4), 864–875. https://doi.org/10.1037/0278-7393.20.4.864.

Curtis, E. T., & Jamieson, R. K. (2018). Computational and empirical simulations of selective memory impairments: Converging evidence for a single-system account of memory dissociations. *Quarterly Journal of Experimental Psychology.* https://doi.org/10.1177/1747021818768502 (advance online publication).

DeCarlo, L. T. (2003). Source monitoring and multivariate signal detection theory, with a model for selection. *Journal of Mathematical Psychology, 47*, 292–303. https://doi.org/10.1016/s0022-2496(03)00005-1.

Diana, R. A., Van den Boom, W., Yonelinas, A. P., & Ranganath, C. (2011). ERP correlates of source memory: Unitized source information increases familiarity-based retrieval. *Brain Research, 1367*, 278–286. https://doi.org/10.1016/j.brainres.2010.10.030.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2007). Imaging recollection and familiarity in the medial temporal lobe: A three-component model. *Trends in Cognitive Sciences, 11*, 379–386. https://doi.org/10.1016/j.tics.2007.08.001.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2008). The effects of unitization on familiarity-based source memory: Testing a behavioral prediction derived from neuroimaging data. *Journal of Experimental Psychology Learning, Memory, and Cognition, 34*, 730–740. https://doi.org/10.1037/0278-7393.34.4.730.

Diana, R. A., Yonelinas, A. P., & Ranganath, C. (2010). Medial temporal lobe activity during source retrieval reflects information type, not memory strength. *Journal of Cognitive Neuroscience, 22*, 1808–1818. https://doi.org/10.1162/jocn.2009.21335.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24*, 523–533. https://doi.org/10.3758/bf03200940.

Dube, C., Starns, J. J., Rotello, C. M., & Ratcliff, R. (2012). Beyond ROC curvature: Strength effects and response time data support continuous-evidence models of recognition memory. *Journal of Memory and Language, 67*, 389–406. https://doi.org/10.1016/j.jml.2012.06.002.

Dunn, J. C. (2003). The elusive dissociation. *Cortex, 39*(1), 177–179. https://doi.org/10.1016/s0010-9452(08)70096-0.

Dunn, J. C. (2004). Remember-know: A matter of confidence. *Psychological Review, 111*, 524–542. https://doi.org/10.1037/0033-295x.111.2.524.

Eichenbaum, H., Yonelinas, A. P., & Ranganath, C. (2007). The medial temporal lobe and recognition memory. *Annual Review of Neuroscience, 30*, 123–152. https://doi.org/10.1146/annurev.neuro.30.051606.094328.

Ekstrom, A. D., & Bookheimer, S. Y. (2007). Spatial and temporal episodic memory retrieval recruit dissociable functional networks in the human brain. *Learning & Memory, 14*, 645–654. https://doi.org/10.1101/lm.575107.

Feustel, T. C., Shiffrin, R. M., & Salasoo, A. (1983). Episodic and lexical contributions to the repetition effect in word identification. *Journal of Experimental Psychology: General, 112*, 309–346. https://doi.org/10.1037/0096-3445.112.3.309.

Gardiner, J. M. (1988). Functional aspects of recollective experience. *Memory & Cognition, 16*, 309–313. https://doi.org/10.3758/bf03197041.

Glanzer, M., Hilford, A., & Kim, K. (2004). Six regularities of source recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1176–1195. https://doi.org/10.1037/0278-7393.30.6.1176.

Hautus, M. J., Macmillan, N. A., & Rotello, C. B. (2008). Toward a complete decision model of item and source recognition. *Psychonomic Bulletin & Review, 15*, 889–905. https://doi.org/10.3758/PBR.15.5.889.

Henson, R. N., Campbell, K. L., Davis, S. W., Taylor, J. R., Emery, T., Erzinclioglu, S., ... Cusack, R. (2016). Multiple determinants of lifespan memory differences. *Scientific Reports, 6*, 32527.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30*(5), 513–541. https://doi.org/10.1016/0749-596X(91)90025-F.

Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General, 110*, 306. https://doi.org/10.1037/0096-3445.110.3.306.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114*, 3–28. https://doi.org/10.1037/0033-2909.114.1.3.

Kellen, D., Singmann, H., Vogt, J., & Klauer, K. C. (2015). Further evidence for discrete-state mediation in recognition memory. *Experimental Psychology, 76*, 216–236. https://doi.org/10.1027/1618-3169/a000272.

Kinder, A., & Shanks, D. R. (2003). Neuropsychological dissociations between priming and recognition: A single-system connectionist account. *Psychological Review, 110*, 728–744. https://doi.org/10.1037/0033-295X.110.4.728.

Klauer, K. C., & Kellen, D. (2010). Toward a complete decision model of item and source recognition: A discrete-state approach. *Psychonomic Bulletin & Review, 17*, 465–478. https://doi.org/10.3758/PBR.17.4.465.

Kurilla, B. P. (2011). Enhanced processing fluency leads to biases in source memory. *The Quarterly Journal of Experimental Psychology, 64*, 1609–1631. https://doi.org/10.1080/17470218.2011.561866.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software, 69*, 1–33. https://doi.org/10.18637/jss.v069.i01.

Lukatela, G., Eaton, T., Moreno, M. A., & Turvey, M. T. (2007). Equivalent inter-and intramodality long-term priming: Evidence for a common lexicon for words seen and words heard. *Memory & Cognition, 35*, 781–800.

Malejka, S., & Bröder, A. (2015). No source memory for unrecognized items when implicit feedback is avoided. *Memory & Cognition, 44*, 63–72. https://doi.org/10.3758/s13421-015-0549-8.

Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252–271. https://doi.org/10.1037/0033-295X.87.3.252.

Meier, B., Theiler-Bürgi, M., & Perrig, W. (2009). Levels of processing and amnesia affect perceptual priming in fragmented picture naming. *International Journal of Neuroscience, 119*(8), 1061–1075. https://doi.org/10.1080/00207450802336691.

Moran, R., & Goshen-Gottstein, Y. (2015). Old processes, new perspectives: Familiarity is correlated with (not independent of) recollection and is more (not equally) variable for targets than for lures. *Cognitive Psychology, 79*, 40–67. https://doi.org/10.1016/j.cogpsych.2015.01.005.

Morey, Richard D., & Rouder, Jeffrey N. (2018). BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-4.2. https://CRAN.R-project.org/package=BayesFactor.

Mulligan, N. W., & Osborn, K. (2009). The modality-match effect in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*(2), 564–571. https://doi.org/10.1037/a0014524.

Onyper, S. V., Zhang, Y. X., & Howard, M. W. (2010). Some-or-none recollection: Evidence from item and source memory. *Journal of Experimental Psychology: General, 139*, 341–364. https://doi.org/10.1037/a0018926.

Ostergaard, A. L. (1992). A method for judging measures of stochastic dependence: Further comments on the current controversy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(2), 413–420. https://doi.org/10.1037/0278-7393.18.2.413.

Ostergaard, A. L. (1998). The effects on priming of word frequency, number of repetitions, and delay depend on the magnitude of priming. *Memory & Cognition, 26*, 40–60. https://doi.org/10.3758/bf03211369.

Poldrack, R. A. (1996). On testing for stochastic dissociations. *Psychonomic Bulletin & Review, 3*(4), 434–448. https://doi.org/10.3758/bf03214547.

Pratte, M. S., & Rouder, J. N. (2011). Hierarchical single- and dual-process models of recognition memory. *Journal of Mathematical Psychology, 55*, 36–46. https://doi.org/10.1016/j.jmp.2010.08.007.

R Core, & Team (2017). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Ratcliff, R., Sheu, C.-F., & Gronlund, S. D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*, 518–535. https://doi.org/10.1037/0033-295x.99.3.518.

Ross, R. S., & Slotnick, S. D. (2008). The hippocampus is preferentially associated with

memory for spatial context. *Journal of Cognitive Neuroscience, 20*, 432–446. https://doi.org/10.1162/jocn.2008.20035.

Rotello, C. M. (2017). Signal detection theories of recognition memory. *Learning and Memory: A Comprehensive Reference,* 201–225. https://doi.org/10.1016/b978-0-12-809324-5.21044-4.

Rouder, J. N., Ratcliff, R., & McKoon, G. (2000). A neural network model of implicit memory for object recognition. *Psychological Science, 11*(1), 13–19. https://doi.org/10.1111/1467-9280.00208.

Schacter, D. L., Wig, G. S., & Stevens, W. D. (2007). Reductions in cortical activity during priming. *Current Opinion in Neurobiology, 17*(2), 171–176. https://doi.org/10.1016/j.conb.2007.02.001.

Shanks, D. R., & Berry, C. J. (2012). Are there multiple memory systems? Tests of models of implicit and explicit memory. *The Quarterly Journal of Experimental Psychology, 65*, 1449–1474. https://doi.org/10.1080/17470218.2012.691887.

Sheldon, S. A. M., & Moscovitch, M. (2010). Recollective performance advantages for implicit memory tasks. *Memory, 18*, 681–697. https://doi.org/10.1080/09658211.2010.499876.

Shimamura, A. P., & Wickens, T. D. (2009). Superadditive memory strength for item and source recognition: The role of hierarchical relational binding in the medial temporal lobe. *Psychological Review, 116*, 1–19. https://doi.org/10.1037/a0014500.

Singmann, H., Bolker, B., & Westfall, J. (2015). afex: Analysis of factorial experiments. Retrieved from http://CRAN.R-project.org/package=afex.

Slotnick, S. D., & Dodson, C. S. (2005). Support for a continuous (single-process) model of recognition memory and source memory. *Memory & Cognition, 33*, 151–170. https://doi.org/10.3758/BF03195305.

Slotnick, S. D., Klein, S. A., Dodson, C. S., & Shimamura, A. P. (2000). An analysis of signal detection and threshold models of source memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 1499–1517. https://doi.org/10.1037/0278-7393.26.6.1499.

Slotnick, S. D., & Thakral, P. P. (2013). The hippocampus operates in a threshold manner during spatial source memory. *NeuroReport, 24*, 265–269. https://doi.org/10.1097/wnr.0b013e32835f282d.

Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General, 117*, 34–50. https://doi.org/10.1037/0096-3445.117.1.34.

Squire, L. R. (1994). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. In D. L. Schacter, & E. Tulving (Eds.). *Memory systems* (pp. 203–231). Cambridge: MIT Press.

Squire, L. R. (2004). Memory systems of the brain: A brief history and current perspective. *Neurobiology of Learning and Memory, 82*, 171–177.

Squire, L. R. (2009). Memory and brain systems: 1969–2009. *Journal of Neuroscience, 29*, 12711–12716. https://doi.org/10.1523/JNEUROSCI.3575-09.2009.

Squire, L. R., & Dede, A. J. O. (2015). Conscious and unconscious memory systems. *Cold Spring Harbor Perspectives in Biology, 7*, a021667. https://doi.org/10.1101/cshperspect.a021667.

Staresina, B. P., Duncan, K. D., & Davachi, L. (2011). Perirhinal and parahippocampal cortices differentially contribute to later recollection of object- and scene-related event details. *Journal of Neuroscience, 31*(24), 8739–8747. https://doi.org/10.1523/jneurosci.4978-10.2011.

Stark, C. E., & McClelland, J. L. (2000). Repetition priming of words, pseudowords, and nonwords. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 26*, 945–972. https://doi.org/10.1037/0278-7393.26.4.945.

Starns, J. J. (2014). Using response time modeling to distinguish memory and decision processes in recognition and source tasks. *Memory & Cognition, 42*, 1357–1372. https://doi.org/10.3758/s13421-014-0432-z.

Starns, J. J., Hicks, J. L., Brown, N. L., & Martin, B. A. (2008). Source memory for unrecognized items: Predictions from multivariate signal detection theory. *Memory & Cognition, 36*, 1–8. https://doi.org/10.3758/MC.36.1.1.

Starns, J. J., Pazzaglia, A. M., Rotello, C. M., Hautus, M. J., & Macmillan, N. A. (2013). Unequal-strength source zROC slopes reflect criteria placement and not (necessarily) memory processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 39*, 1377–1392. https://doi.org/10.1037/a0032328.

Starns, J. J., Rotello, C. M., & Hautus, M. J. (2014). Recognition memory zROC slopes for items with correct versus incorrect source decisions discriminate the dual process and unequal variance signal detection models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 1205–1225. https://doi.org/10.1037/a0036846.

Stróżak, P., Bird, C. W., Corby, K., Frishkoff, G., & Curran, T. (2016). FN400 and LPC memory effects for concrete and abstract words. *Psychophysiology, 53*(11), 1669–1678. https://doi.org/10.1111/psyp.12730.

Taylor, J. R., & Henson, R. N. (2012). You can feel it all over: Many signals potentially contribute to feelings of familiarity. *Cognitive Neuroscience, 3*, 209–210. https://doi.org/10.1080/17588928.2012.689966.

Thakral, P. P., Kensinger, E. A., & Slotnick, S. D. (2016). Familiarity and priming are mediated by overlapping neural substrates. *Brain Research, 1632*, 107–118. https://doi.org/10.1016/j.brainres.2015.12.008.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science, 247*, 301–306. https://doi.org/10.1126/science.2296719.

Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 336–342. https://doi.org/10.1037/0278-7393.8.4.336.

Voss, J. L., & Federmeier, K. D. (2011). FN400 potentials are functionally identical to N400 potentials and reflect semantic processing during recognition testing. *Psychophysiology, 48*(4), 532–546. https://doi.org/10.1111/j.1469-8986.2010.01085.x.

Ward, E. V., Berry, C. J., & Shanks, D. R. (2013). An effect of age on implicit memory that is not due to explicit contamination: Implications for single and multiple-systems theories. *Psychology and Aging, 28*, 429–442. https://doi.org/10.1037/a0031888.

Wixted, J. T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review, 114*, 152–176. https://doi.org/10.1037/0033-295X.114.1.152.

Wixted, J. T., & Mickes, L. (2010). A continuous dual-process model of remember/know judgments. *Psychological Review, 117*, 1025–1054. https://doi.org/10.1037/a0020874.

Yonelinas, A. P. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*, 1341–1354. https://doi.org/10.1037/0278-7393.20.6.1341.

Yonelinas, A. P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1415–1434. https://doi.org/10.1037/0278-7393.25.6.1415.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441–517. https://doi.org/10.1006/jmla.2002.2864.

Zaki, S. R., Nosofsky, R. M., Jessup, N. M., & Unverzagt, F. W. (2003). Categorization and recognition performance of a memory-impaired group: Evidence for single-system models. *Journal of the International Neuropsychological Society, 9*, 394–406. https://doi.org/10.1017/s1355617703930050.