Journal of Experimental Psychology:
Learning, Memory, and Cognition
2008, Vol. 34, No. 3, 570–586

0278-7393/08/$12.00 DOI: 10.1037/0278-7393.34.3.570

A Simplified Conjoint Recognition Paradigm for the Measurement of Gist and Verbatim Memory

Christoph Stahl and Karl Christoph Klauer
University of Freiburg

The distinction between verbatim and gist memory traces has furthered the understanding of numerous phenomena in various fields, such as false memory research, research on reasoning and decision making, and cognitive development. To measure verbatim and gist memory empirically, an experimental paradigm and multinomial measurement model has been proposed but rarely applied. In the present article, a simplified conjoint recognition paradigm and multinomial model is introduced and validated as a measurement tool for the separate assessment of verbatim and gist memory processes. A Bayesian metacognitive framework is applied to validate guessing processes. Extensions of the model toward incorporating the processes of phantom recollection and erroneous recollection rejection are discussed.

Keywords: memory, conjoint recognition, multinomial modeling, source monitoring, metacognitive guessing strategies

In research on episodic memory, fuzzy trace theory (FTT) has recently received much attention. At its core lies the assumption of separate memory traces for the meaning or *gist* of an item (e.g., its semantic category) and its identity or *verbatim* detail (e.g., its exact wording) that can be retrieved independently from memory. This distinction has proven fruitful in a number of domains but especially so in research on false memory (e.g., Brainerd, Forrest, Karibian, & Reyna, 2006; Brainerd, Payne, Wright, & Reyna, 2003; Brainerd & Reyna, 2002; Brainerd, Wright, Reyna, & Mojardin, 2001; Odegard & Lampinen, 2005; Seamon et al., 2002; Wright & Loftus, 1998). To empirically separate verbatim and gist memory, Brainerd, Reyna, and Mojardin (1999) have proposed the conjoint recognition (CR) paradigm and multinomial model. However, only few studies have since used this relatively complex paradigm. In the present research, a simplified CR paradigm is introduced and validated.

In Brainerd et al.'s (1999) CR paradigm, participants are first presented with a study list. The test list contains three types of items: *target probes* (i.e., old items from the study list), *related distracters* that share a target's gist, and *unrelated distracters*. The memory test is administered to three groups of participants with different instructions: Under the T instruction, participants are asked to accept as old only targets; under the R instruction, participants are to accept as old only related distracters; and, finally, under the T + R instruction, participants are to accept both targets and related distracters. From the proportions of accepted targets, related distracters, and unrelated distracters obtained in the three different groups, the parameters of a model are estimated that provide measures of verbatim and gist memory.

In the following, a simplified CR paradigm and model are proposed that also provide valid estimates of verbatim and gist memory but on the basis of a much simpler procedure. In the simplified CR paradigm, all participants are presented with the same test list, and a single instruction is presented asking the same multiple choice question for all items. As we will show, in the simplified CR paradigm, valid measures of gist and verbatim memory can be obtained from a single group of participants.

The Original CR Paradigm and Model

In the original version of the CR paradigm, three different memory processes are empirically separated by way of a multinomial measurement model: an identity judgment based on the verbatim trace, a similarity judgment based on the gist trace, and a process of recollection rejection (Brainerd, Reyna, Wright, & Mojardin, 2003). FTT postulates that for each event, two different memory traces are created: First, a verbatim trace stores the event's perceptual detail. At test, when a match is found between a verbatim trace and the verbatim information present in the probe, an identity judgment is made that leads to the probe's acceptance. Second, independently, the gist trace stores the core meaning of an item. At test, a detection of similarity between the probe's gist and the gist traces stored in memory is assumed to lead to an acceptance response. Both memory traces are stored separately and can be retrieved independently.

According to FTT, false memories arise when the gist trace of an event is not integrated with its verbatim trace, that is, when gist memory is retrieved but verbatim memory is not. In this case, two opposing processes are postulated. First, a gist-based similarity judgment is thought to underlie false recognition of related dis-

Christoph Stahl and Karl Christoph Klauer, Institute for Psychology, University of Freiburg, Freiburg, Germany.

The research reported in this article was supported by Deutsche Forschungsgemeinschaft Grant Kl 614/31-1 to Karl Christoph Klauer. We would like to thank Stephan Stegt for providing the materials used in Experiment 2, and William Batchelder, Charles Brainerd, John Dunn, Stephan Lewandowski, and Timothy Odegard for their valuable comments on drafts of this article.

Correspondence concerning this article should be addressed to Christoph Stahl, Institute for Psychology, University of Freiburg, D-79085 Freiburg im, Breisgau, Germany. E-mail: stahl@psychologie.uni-freiburg.de

tracters. This process is in opposition to a second verbatim-based *recollection rejection* process that is the basis of rejecting related distracters. Recollection rejection is thought to be based on judging nonidentity between the verbatim information present in the probed related distracter and the verbatim trace of the corresponding target, and thereby reduces false recognition of semantically related lures based on item-specific verbatim traces (Brainerd et al., 1999; Brainerd, Reyna, et al., 2003). Considerable evidence for this process has accrued (e.g., Lampinen, Odegard, & Neuschatz, 2004; Lampinen, Watkins, & Odegard, 2006; Seamon et al., 2002).

To separate verbatim and gist memory traces empirically, Brainerd and colleagues have proposed the CR paradigm (Brainerd, Holliday, & Reyna, 2004; Brainerd et al., 1999, 2001; Brainerd, Stein, & Reyna, 1998). In a CR memory test, participants are presented with targets (i.e., items from the study list), related distracters (i.e., items that had not been presented on the study list but are related to a target via common gist), and unrelated distracters (i.e., new items that were neither part of the study list nor are related to a target). In the original CR paradigm introduced by Brainerd et al. (1999), participants undergo a memory test under one of three instruction conditions: the T, R, or T + R conditions that have already been introduced above. From the 3 (probe types) × 3 (instruction conditions) acceptance probabilities, the parameters of a multinomial model are estimated that provide measures of gist and verbatim memory as well as estimates of acceptance by guessing (for an introduction to multinomial models, see Batchelder & Riefer, 1999; Riefer & Batchelder, 1988).

Performance in the CR memory test is determined by the interplay of the identity, similarity, and recollection rejection judgment processes, as well as by guessing. The processing-tree representation of the original multinomial CR model (see Figure 1) illustrates how the postulated processes interact. Consider the first tree diagram that represents cognitive processes in reaction to a target probe. When the verbatim trace of the target can be retrieved (with probability V_t), comparison with the probe yields an identity judgment, which can be conceived of as a conscious and explicit recollection of the target episode. Participants can therefore correctly accept a target under the T and T + R instructions and correctly reject it under the R instruction.

With probability $1 - V_t$, the verbatim trace cannot be retrieved. In that case, the gist trace can still be retrieved with probability G_t, resulting in a judgment of similarity between the gist of the probe and the gist of a memory episode. Given this mnemonic state—the meaning of an item has successfully been retrieved, but no information is available that would support a classification of that item as a target or a related probe—which response should be selected? Brainerd et al. (1999) postulated that participants automatically attribute the detected similarity to the source specified in the instruction (i.e., under the T instruction, a similarity judgment would lead to a "target" response, whereas under the R instruction, it would result in a "related" response). Brainerd et al. (1999) reported indirect evidence to support this assumption. Thus, a similarity judgment is assumed to produce an acceptance response under all three instruction conditions. With probability $(1 - V_t)(1 - G_t)$, neither the verbatim nor the gist trace can be retrieved, and participants guess whether to accept (with a different probability b_i for each instruction condition) or reject the probe (with probability $1 - b_i$).

Next, consider the second tree diagram in Figure 1 for a related probe. First, it is possible that the related probe acts as a retrieval cue for the verbatim trace of the target that it is related to (with probability V_r). In this case, a judgment of nonidentity between the verbatim information of probe and target results in the rejection of the probe under the T instruction and acceptance under the R and

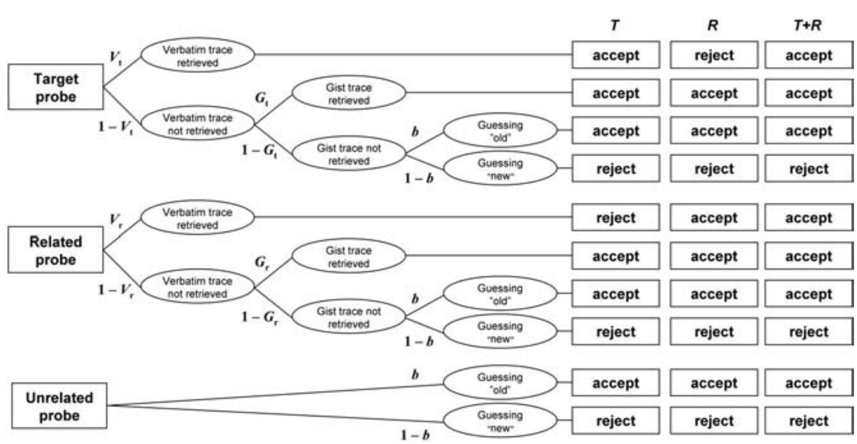

Figure 1. Processing tree model for the original conjoint recognition paradigm. Rectangles on the left denote probe type, rectangles on the right denote responses; columns represent different instruction conditions. Branches of the processing tree represent the combination of cognitive processes postulated by the model. V_t = probability of retrieving a target's verbatim trace given a target probe; V_r = probability of retrieving a target's verbatim trace given a related probe; G_t = probability of retrieving a target's gist trace given a target probe; G_r = probability of retrieving a target's gist trace given a related probe; b = probability of guessing that an item is old. Note that different parameters—b_T, b_R, and b_{T+R}—are used to represent guessing under the different instruction conditions (i.e., under the T instruction, participants are asked to accept as old only targets; under the R instruction, participants are to accept as old only related distracters; and under the T + R instruction, participants are to accept both targets and related distracters).

T + R instructions. With probability $1 - V_r$, the corresponding target's verbatim trace cannot be retrieved. In that case, its gist trace can still be retrieved with probability G_r, which would provoke a similarity judgment causing acceptance under all conditions. With probability $(1 - V_r)(1 - G_r)$, neither the corresponding target's verbatim trace nor its gist trace can be retrieved, and responses are again determined by guessing processes.

As illustrated by the third diagram in Figure 1, when presented with an unrelated probe, neither verbatim nor gist traces are available, and participants' responses are assumed to be based solely on guessing processes. This implies that the original CR model does not incorporate a process by which new items can be detected as new (e.g., a metacognitive process, such as the one suggested by Strack & Bless, 1994). Models of recognition memory differ with regard to the inclusion of such a process (for discussions of this issue for source monitoring [SM] models, see, e.g., Batchelder & Riefer, 1990; Batchelder, Riefer, & Hu, 1994; Bayen, Murnane, & Erdfelder, 1996).

The model's parameters can be interpreted as the probability of the cognitive processes that they represent. Thus, V_t represents the probability of retrieving a target's verbatim trace, given the target probe as a retrieval cue. V_r represents the probability of retrieving a target's verbatim trace, given a related probe as a retrieval cue. V_r is assumed to be smaller than V_t because a related probe likely does not constitute as good a retrieval cue as the target itself. G_t and G_r represent the probabilities that responses to a target probe and a related probe, respectively, are based on similarity ratings in the absence of retrieval of a verbatim trace. For these parameters, no order relation has been proposed a priori. Finally, b_T, b_R, and b_{T+R} represent the probabilities of accepting a probe via guessing processes in the T, R, and T + R conditions, respectively.

In a nutshell, in the original CR paradigm, participants' recognition memory for three types of probes (targets, related distracters, and unrelated distracters) is tested under three between-subjects conditions. A multinomial model is then fitted to the data, and its parameters provide estimates of verbatim and gist memory traces as well as guessing processes.

We endorse the model-based approach to process separation that was followed in the original CR paradigm. However, we believe that the same result—separate and uncontaminated measures of verbatim and gist memory as well as guessing processes—can be achieved more efficiently. Here, we propose a simplified CR paradigm in which a single memory test condition suffices to obtain those estimates.[1]

The Simplified CR Paradigm and Model

The CR procedure can be modified so that it is no longer necessary to administer a memory test in three separate groups, implying a considerable advantage in terms of the efficiency of data collection. In the present simplified CR paradigm, estimates for verbatim and gist memory can be obtained for a single group of participants, as compared with three groups of participants in the original paradigm, thereby reducing the costs in terms of the required number of subjects by two thirds. In the simplified paradigm, as in the original paradigm, participants are presented with targets, related probes, and unrelated probes, and they are informed as to the types of items that compose the test list. However, in contrast to the original paradigm, the simplified paradigm asks not for acceptance or rejection responses but for the identification of the probe's type. Participants are asked to classify targets, related probes, and unrelated probes in a single condition. They are instructed to respond "target" if they believe that the current probe has been presented in the learning phase. If they believe the current probe to be a related distracter, they are instructed to indicate this by selecting the "related" response. If they consider the probe to be an unrelated distracter, they are to select the "new" response. This procedure results in 3×2 independent empirical probabilities (two response probabilities for each type of probe are free to vary).

Before we introduce the multinomial model for the simplified paradigm, note that the procedural simplification affects the processes that are thought to occur given the mnemonic state of gist memory without verbatim memory. Given gist memory without verbatim memory, participants infer that the item is not an unrelated probe, but they are left to choose between the "target" and "related" responses. As mentioned above, in this case, it was assumed that the similarity judgment would automatically produce an "accept" response in all three conditions of the original CR paradigm (Brainerd et al., 1999). This assumption is no longer necessary in the simplified procedure. Instead, a new guessing parameter a is introduced to model the probability with which participants select the response "target" rather than "related" in this mnemonic state.

The modified multinomial model for the simplified CR paradigm has a total of six parameters. Given six empirical probabilities, it is a saturated model. The equations and a proof of identifiability are provided in the Appendix, in which we also discuss the relationship of the CR model to the related SM model. The processing-tree representation of the model is given in Figure 2. It is almost identical to the original model, with the exception of an additional parameter a that represents the process of guessing "target" or "related." This process is relevant in case of available gist but no verbatim memory, $(1 - V_t)G_t$, and in the case that a probe's meaning has been classified as old by way of guessing, $(1 - V_t)(1 - G_t)b$. Consider the first tree diagram that represents the cognitive processes occurring when a target probe is presented at test. In case of available verbatim memory, it is correctly identified as a target. Given no

[1] Despite its ability to separately assess verbatim and gist processes, the original CR paradigm and model has rarely been implemented (a database search for the term *conjoint recognition* yielded only 12 hits, as opposed to 129 hits for *fuzzy trace theory*; for exceptions, see Brown & Gorfein, 2004; Rotello, 2001; Stahl, 2004, 2006; Stahl & Klauer, in press). This might be due to the relatively costly design, requiring three between-participants test conditions for each single level of the independent variable. In addition, the relative novelty of the multinomial modeling approach and the potentially problematic assumption of the identity of cognitive processes across conditions (e.g., Rotello, 2001) might be relevant. Cowan (1998) suggested that the R instruction may be too complex and might therefore not yield reliable results. In line with this claim, in our own research using the CR paradigm (Stahl, 2004, 2006), we regularly observed that the R instruction posed a challenge for a (small) number of individuals from our sample of college students. If the R instructions proved to be too complex, data from the original CR procedure are likely to be unreliable. Note, however, that the CR procedure has been successfully applied not only for adults but also for children as young as 5 years of age (e.g., Brainerd et al., 1998, 2004), suggesting that, whereas specific realizations of it might be problematic, this is not true for the R instruction per se.

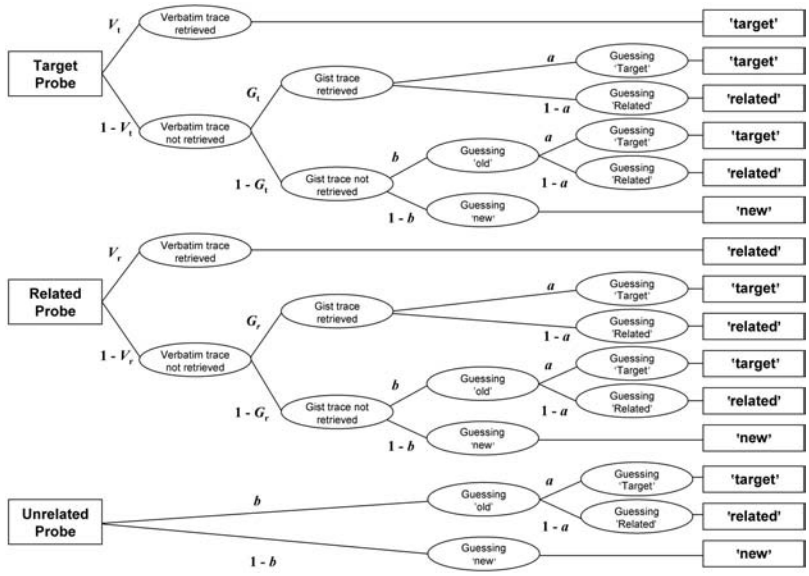

Figure 2. Processing tree model for the simplified conjoint recognition paradigm. Rectangles on the left denote probe type, rectangles on the right denote responses. They are connected by branches of the processing tree that represent the combination of cognitive processes postulated by the model. V_t = probability of retrieving a target's verbatim trace given a target probe; V_r = probability of retrieving a target's verbatim trace given a related probe; G_t = probability of retrieving a target's gist trace given a target probe; G_r = probability of retrieving a target's gist trace given a related probe; b = probability of guessing that an item is either a target or a related probe; a = probability of guessing "target."

verbatim memory but available gist memory—participants have identified the probe's meaning as old but cannot remember whether the probe itself or a related item with the same gist had been presented in the learning phase—a decision has to be made between the "target" and "related" response options. With probability a, the probe is classified as a target, and with probability $1 - a$, the probe is classified as a related distracter. Should neither memory trace be available, participants can still guess that the probe's meaning is old (with probability b). In this case, a choice between the "target" and "related" responses is again required, which is again modeled by the new parameter a as described above. The same decision process involving parameter a is postulated to occur for related probes, as can be seen in the branches $(1 - V_r)G_r$ and $(1 - V_r)(1 - G_r)b$ of the second diagram.

Classifications of unrelated probes are based on a combination of guessing processes a and b, as illustrated in the third diagram. Note that, as in the original CR model, we do not include a process of detecting new items as new in the simplified model. However, it is possible to do so, and we discuss this possibility in the General Discussion section.

Validation of a Measurement Model

A few words are in order to clarify the status of the models. In our view, the mathematical models discussed here should not be taken as precise and complete theories of recognition memory. In line with Batchelder and Batchelder (2008), we view models as measurement tools that are useful in empirically dissociating cognitive processes of interest. Such measurement tools are useful to the extent that they provide good approximations of the cognitive processes of interest. It is therefore necessary to demonstrate for each such model that its parameters are valid measures of the processes they represent. This can be achieved in an empirical validation program in which each of the processes represented in the model are separately targeted by manipulations that are hypothesized to selectively affect the given process. A model can be considered valid when its parameters are shown to selectively respond to manipulations targeted at the process that is measured by that specific parameter. In the following, we will present such a validation program for the model of the simplified CR paradigm to demonstrate that the model for the simplified CR paradigm is capable of separating verbatim and gist memory.

We conducted a series of six studies using the simplified CR paradigm. In these studies, experimental manipulations are implemented that have been shown to affect verbatim and gist memory parameters in the original CR paradigm by Brainerd et al. (1999), and results demonstrate that these manipulations have the same effects on the parameters for verbatim and gist memory obtained through the simplified CR paradigm.

General Method

Participants

Participants were sampled from the department's database of volunteers (mostly students from Freiburg's universities and colleges, as well as nonstudent citizens) and participated in exchange for a certificate of participation or monetary compensation (Experiments 1 and 2: 7 Euro; Experiments 3–6: 3.50 Euro). Each volunteer participated in only one of the reported experiments. Participants' native language was German.

Materials

Three sets of German word lists were used: synonym pairs, category lists, and Deese–Roediger–McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) lists. In all experiments, five items were added as primacy buffer, and five items were added as recency buffer to the study lists.

Synonym pairs. One set consisted of synonym word pairs with each pair denoting one occupation concept. Gist memory was defined as memory for the concept, and verbatim memory was defined as memory for the word that was presented. Twenty pairs of occupation names were taken from Stahl (2004), in which both members of a pair were synonyms for a single target occupation (e.g., *barber* and *hairdresser*). Pairs were randomly assigned to contribute either a target or a related probe to the test list, each synonym of a given pair being selected as a target or related probe with the same probability. Twenty additional occupation names were used as unrelated probes that denoted other occupations.

Category exemplars. The second set consisted of exemplar pairs of common categories. Gist memory was defined as memory for the category, and verbatim memory was defined as memory for the presented exemplar. Two exemplars were generated for 30 common categories by the authors (e.g., *hammer* and *saw* as exemplars of the category *tools*). Categories were randomly assigned to contribute either a target, a related probe, or an unrelated probe to the test list, and each exemplar was selected as a target, related, or unrelated probe with the same probability.

DRM lists. A third set was used in Experiment 2 and consisted of German DRM lists (Deese, 1959; Roediger & McDermott, 1995) that were taken from Stegt (2006). DRM lists consist of a number of to-be-presented items and a single critical lure item that is related to list items by common gist. The critical lure items are not presented but nevertheless often recalled and recognized with high probability. Gist memory was defined as memory for the list theme, and verbatim memory was defined as memory for the presented list item. Thirty DRM lists with four list items and one critical lure each were used. Lists were randomly assigned to contribute either a target, a related distracter, or an unrelated distracter to the test list. The first list items of presented lists were used as targets, the critical lures were used as related distracters, and the first list items of nonpresented lists were used as unrelated distracters.

Experiments 1 and 2: Gist Memory

In Experiments 1 and 2, we aimed to validate gist memory parameters. We implemented a manipulation that has been shown to affect gist memory in the original CR paradigm. Brainerd et al. (2001, Experiment 3) manipulated gist memory for a concept by presenting one versus multiple items related to that concept at study. When the concept was repeatedly activated at study by multiple items, memory for its gist was increased. In Experiment 1, gist memory was manipulated by varying the number of items (one vs. four) that were presented at study from a given target category. Gist memory was predicted to be greater for targets as well as for related items from categories from which more exemplars were presented. In Experiment 2, we attempted to selectively target the gist memory parameter for related items, as this process is thought to underlie many phenomena of false memory (e.g., Brainerd & Reyna, 2002). We again presented one versus four items from a given concept, but in this study, we used DRM lists (Deese, 1959; Roediger & McDermott, 1995) that are typically used in false memory research. DRM lists consist of semantic associates converging on critical, nonstudied lure words (e.g., butter, loaf, knife—*bread*). In this paradigm, false alarms for nonpresented, critical lure words (e.g., *bread*) are often as frequent or even more frequent than hits for presented items (e.g., butter). An important characteristic of these lists is that critical lures are strong semantic associates of list items, whereas strong semantic associations do not necessarily exist among list items. As a result, by presenting an increasing number of list items, the critical lure's gist becomes increasingly activated, whereas the gist of other list items is not or only weakly affected. We therefore predicted increased gist memory parameters for related probes (i.e., the critical lures) but not for target probes (i.e., the first item on each list).

Method

Participants. Twenty volunteers participated in Experiment 1. One participant had to be excluded because of prior participation in another experiment of the series. Nineteen participants remained (13 women; ages ranged from 20 to 39 years, $M = 23$). In Experiment 2, 21 volunteers participated (13 women; ages ranged from 18 to 23 years, $M = 20$).

Design. A 2 (Gist Activation: weak vs. strong) $\times$ 3 (Probe Type: target, related, unrelated) repeated-measures design was implemented.

Materials. Gist Activation was manipulated by presenting different numbers of items from each concept on the study list. In the weak condition, a concept was represented by a single item; in the strong condition, a concept was represented by four items. Experiment 1 used the 30 categories described in the General Method section. Two additional category exemplars were generated for each category so that there were five exemplars per category. The study list presented exemplars from 20 randomly selected categories that were randomly split into two halves of 10 categories each. Categories from the first half were represented by a single item randomly drawn from the five available exemplars; categories from the second half were represented by four items, again randomly drawn from the five available exemplars. In total, 50 items were thereby presented in random order.

At test, the 10 single-item categories were randomly split into five categories for which the single presented item was shown as target probe and five categories for which a randomly selected nonpresented exemplar was shown as related probe. Similarly, the 10 four-item categories were randomly split into five categories for which a randomly selected item from the four presented exemplars was shown as the target probe and five categories for which the nonpresented fifth exemplar was shown as the related probe. Ten unrelated probes were randomly selected from the exemplars of the remaining categories that were not presented in the study list, one unrelated probe representing each such category. Order of presentation of these 30 test list items was randomized. All randomizations were carried out for each participant anew.

In Experiment 2, we used 30 DRM lists (see the General Method section). The study list presented items from 20 randomly selected DRM lists that were randomly split into two halves of 10 DRM lists each. DRM lists from the first halves were represented by the first DRM list item; DRM lists from the second half were represented by the first four items. In total, 50 items were thereby presented in random order.

At test, the 10 single-item DRM lists and the 10 four-item DRM lists were each randomly split into five DRM lists for which the first DRM list item was shown as target probe and five DRM lists for which the critical lure was shown as related probe. Unrelated probes were the first DRM list items from 10 DRM lists that were not represented in the study list. In total, 30 items were thereby presented at test in random order. All randomizations were carried out for each participant anew.

Procedure. In Experiments 1 and 2, we used the simplified CR procedure in individual computerized sessions. Participants were instructed that they were to be presented with a list of items that they were to remember for a later test. Study items were presented sequentially for 4,000 ms in black Sans-Serif letters on a gray background in the center of the screen. After the study phase, participants solved arithmetic problems for a total duration of 5 min. The memory test was administered approximately 24 hr after the learning phase to minimize effects of verbatim memory. In the memory test, participants were presented sequentially with a list of probes, and they indicated their mnemonic state for each probe by selecting the appropriate response with a computer mouse. Specifically, they were to indicate whether the probe was identical to an old item (i.e., a "target"), "related" to an old item, or "new." On the basis of the results from two pilot studies, we chose to present participants with a simultaneous decision with these three response options (see also Marsh & Hicks, 1998). After completing the memory test, participants were thanked, debriefed, and dismissed.

Results

Response frequencies are given in Table 1. Parameter estimates and significance tests are given in Table 2. Guessing parameters were set equal across the Gist Activation factor for identifiability reasons.[2] Parameter estimation and hypotheses tests reported below were performed with the HMMTree software (Stahl & Klauer, 2007). Sensitivity power analyses (performed with G*Power 3; Faul, Erdfelder, Lang, & Buchner, 2007) assured high test power, $1 - \beta = .95$, for parameter comparisons across conditions. With $\alpha = \beta = .05$, we were able to detect small to medium effects ($.09 < w < .25$; see Cohen, 1988, chapter 5).

In Experiment 1, as predicted, gist memory (parameters G_t and G_r) was greater for items from categories from which four exemplars had been presented at study than for items from categories from which only a single item had been presented. In Experiment 2, again as predicted, gist memory for related probes (G_r) but not for targets (G_t) was greater for items from DRM lists from with four items had been presented at study than for items from DRM lists from which only a single item had been presented.

No other effects were significant. Verbatim memory was not affected. Although estimates of verbatim memory (V_t) are slightly increased for targets from four-item categories and DRM lists, this difference is not significant. Estimates of the recollection rejection process (V_r) appeared to be reduced for four-item categories in Experiment 1, but this is likely due to random error, given the large confidence intervals of V_r parameters.

Table 1
Observed Frequencies of Memory Judgments in Experiments 1–6

Experiment	Manipulation	Probe type	Response t	Response r	Response u
Experiment 1	Gist Activation: Weak	t	63	12	20
		r	9	45	41
	Gist Activation: Strong	t	75	15	5
		r	22	63	10
		u	8	25	157
Experiment 2	Gist Activation: Weak	t	57	19	29
		r	25	29	51
	Gist Activation: Strong	t	71	20	14
		r	40	50	15
		u	25	52	133
Experiment 3	Target Presentations: 1	t	80	11	9
		r	15	69	16
	Target Presentations: 2	t	96	3	1
		r	12	72	16
		u	4	30	166
Experiment 4	Target Presentations: 1	t	86	10	4
		r	4	59	37
	Target Presentations: 2	t	97	2	1
		r	6	72	22
		u	7	33	160
Experiment 5	Target-first	t	182	6	12
		r	14	162	24
		u	22	29	149
	Target-last	t	168	15	17
		r	41	117	42
		u	15	39	146
	Control	t	174	9	17
		r	29	136	35
		u	9	28	163
Experiment 6	Target-first	t	167	8	25
		r	11	163	26
		u	15	23	162
	Target-last	t	171	13	16
		r	16	115	69
		u	15	46	139
	Control	t	167	20	13
		r	18	132	50
		u	6	49	145

Note. t = target probe; r = related distracter; u = unrelated distracter.

Discussion

In Experiments 1 and 2, we examined the effects of a manipulation known to increase gist memory on the parameters of the CR model. In Experiment 1, activation of the gist of a category was

[2] It was necessary to introduce additional restrictions because the manipulation of number of presentations affected only targets and related probes but did not affect unrelated probes. Therefore, it yielded only four additional independent empirical probabilities, and as a result, it was possible to estimate only four additional parameters, making it necessary to equate two parameters across conditions. This is in contrast to the between-participants manipulations used in Experiments 5 and 6 that yielded six additional empirical probabilities for each condition.

Table 2
Estimates (and 95% Confidence Intervals) for the Parameters of the Simplified Conjoint Recognition Model for Experiments 1 and 2

Parameter	Gist activation: Weak	Gist activation: Strong	$\Delta G^2_{(df = 1)}$	p
	Experiment 1			
a	.25 (.11, .40)			
b	.17 (.12, .23)			
G_t	.33 (.10, .56)	.76 (.56, .96)	7.17	.01
G_r	.35 (.07, .63)	.87 (.76, .98)	22.15	<.01
V_t	.62 (.51, .73)	.74 (.62, .85)	2.20	.14
V_r	.20 (.00, .51)	.00 (.00, .62)	0.98	.32
	Experiment 2			
a	.41 (.30, .52)			
b	.37 (.30, .43)			
G_t	.25 (.00, .50)	.54 (.30, .77)	2.97	.08
G_r	.23 (.00, .47)	.77 (.65, .90)	26.45	<.01
V_t	.42 (.28, .56)	.55 (.40, .69)	1.80	.18
V_r	.00 (.00, .22)	.00 (.00, .31)	0.00	1.00

Note. a = probability of guessing "target"; b = probability of guessing that an item is either a target or a related probe; G_t = probability of retrieving a target's gist trace given a target probe; G_r = probability of retrieving a target's gist trace given a related probe; V_t = probability of retrieving a target's verbatim trace given a target probe; V_r = probability of retrieving a target's verbatim trace given a related probe.

manipulated by varying the number of items (one vs. four) that was presented from each category at study. Results indicate that this manipulation affected the gist parameters as expected: Both the gist parameters for targets and for related items were increased for categories from which four items had been presented. In Experiment 2, we aimed to discriminate between gist memory for targets and related probes by asymmetrically affecting the different types of gist memory. To accomplish this, we used DRM lists (Deese, 1959; Roediger & McDermott, 1995). These lists are characterized by strong associations from list items to a critical item, whereas associations between list items are weaker. To manipulate gist memory, we presented participants at study with either one or four DRM list items. At test, we presented the critical item as related probe, and we expected that gist memory for critical items (i.e., G_r) should be increased strongly by the presentation of three additional associates. On the other hand, gist memory should only show a modest increase for the first item from each list that was presented as a target probe (i.e., G_t). This pattern was observed in gist parameters for Experiment 2. In sum, gist memory parameters of the modified CR model responded to the manipulations as predicted, supporting the conclusion that they provide valid indicators of gist memory.

Experiments 3 and 4: Verbatim Memory

Next, we turned to validating the verbatim memory parameters, using the same manipulation that has been used to validate verbatim parameters in the original CR paradigm. In Experiments 3 and 4, half of the targets were presented repeatedly on the study list. This should increase verbatim memory for these targets, as an additional presentation doubles the time of exposition of the perceptual surface of that stimulus. Repetition has been shown to selectively affect verbatim memory in the original CR paradigm (Brainerd et al., 1999, Experiments 1 and 2), and we aimed to replicate this finding within the simplified CR paradigm.

We predicted that the V_t parameter would respond to the manipulation of verbatim trace strength by repeated presentation. Recall that this parameter is an indicator for the identity process by which targets' identity is verified with the help of a successfully retrieved verbatim trace. Gist memory should be affected by this manipulation only to a negligible extent because a single presentation of a word usually suffices to fully extract its meaning and activate a gist representation, and an additional presentation would not add to this activation.

Method

Participants. Twenty volunteers participated in each experiment (Experiment 3: 15 women; ages ranged from 18 to 26 years, $M = 21$; Experiment 4: 11 women; ages ranged from 19 to 45 years, $M = 25$).

Design. A 2 (Number of Presentations: 1 vs. 2) × 3 (Probe Type: target, related, unrelated) repeated-measures design was implemented.

Materials. Experiment 3 used the synonym pairs described in the General Method section. The study list presented one randomly selected item from each of the 20 pairs. These items were randomly split into 10 items that were presented once and 10 items that were presented twice. In total, 30 items were presented in random order.

At test, the 10 single-presentation items and the 10 repeated-presentation items were randomly split into five items each that were shown as target probe and five items for which their synonym was presented as related probe. Ten additional occupation names not presented at study were presented as unrelated probes. Order of presentation of these 30 items was randomized. All randomizations were carried out for each participant anew.

Experiment 4 used the categories described in the General Method section. The study list presented exemplars from 20 randomly selected categories that were randomly split into two halves of 10 categories each. Categories from the first half were represented by a single exemplar that was presented once; categories from the second half were represented by a single exemplar that was presented twice on the study list. In total, 30 items were presented in random order.

At test, the 10 single-presentation categories and the 10 repeated-presentation categories were randomly split into five categories each for which the presented exemplar was shown as target probe and five categories for which the nonpresented exemplar was shown as related probe. Ten unrelated probes were randomly selected from the exemplars of the remaining categories that were not presented at study, one exemplar representing each such category. Order of presentation of these 30 items was randomized. All randomizations were carried out for each participant anew.

Procedure. The procedures were identical to those of Experiments 1 and 2, with the exception that the memory test followed immediately after the arithmetic filler task.

Results

Parameter estimates and significance tests are given in Table 3. The predicted effect of repetition was observed on verbatim memory for targets, V_t. In both experiments, V_t was larger for twice-presented targets than for once-presented targets.

No other effects were obtained. As predicted, gist parameters were not affected by the presentation manipulation (all ps $> .18$). Recollection rejection (parameter V_r) was also not affected (both ps $> .72$).

Table 3
Estimates (and 95% Confidence Intervals) for the Parameters of the Simplified Conjoint Recognition Model for Experiments 3 and 4

Parameter	Target presentations: 1	Target presentations: 2	$\Delta G^2_{(df\,=\,1)}$	p
		Experiment 3		
a	.16 (.04, .28)			
b	.17 (.12, .22)			
G_t	.51 (.25, .77)	.74 (.27, 1.00)	0.58	.45
G_r	.81 (.64, .98)	.79 (.60, .98)	0.05	.82
V_t	.78 (.69, .87)	.95 (.91, 1.00)	12.36	<.01
V_r	.00 (.00, .76)	.09 (.00, .79)	0.13	.72
		Experiment 4		
a	.17 (.06, .29)			
b	.20 (.14, .26)			
G_t	.69 (.42, .96)	.63 (.01, 1.00)	0.03	.86
G_r	.23 (.00, .59)	.51 (.17, .85)	1.79	.18
V_t	.84 (.76, .92)	.97 (.93, 1.00)	8.39	<.01
V_r	.40 (.13, .67)	.44 (.08, .79)	0.04	.84

Note. a = probability of guessing "target"; b = probability of guessing that an item is either a target or a related probe; G_t = probability of retrieving a target's gist trace given a target probe; G_r = probability of retrieving a target's gist trace given a related probe; V_t = probability of retrieving a target's verbatim trace given a target probe; V_r = probability of retrieving a target's verbatim trace given a related probe.

Discussion

In Experiments 3 and 4, verbatim memory for target probes was manipulated. During the presentation phase, half of the items were presented once, whereas the other half were presented twice. As predicted, the repetition manipulation affected targets' verbatim memory (V_t) but left gist memory parameters unaffected. This replicates the finding by Brainerd et al. (1999) that verbatim memory is affected by repeated presentation in the original CR paradigm. It is concluded that the simplified CR paradigm and model is well capable of assessing targets' verbatim trace strength.

Recollection rejection (V_r) tended to be increased in Experiment 4 as compared with Experiment 3. This nonsignificant tendency cannot be attributed to differences in verbatim memory strength, as V_t was at comparable levels in both experiments. Instead, it might reflect differences in metacognitive strategy use in the recollection–rejection decision (e.g., Gallo, 2004). Recollection rejection also tended to be higher for repeatedly presented items but not significantly so. This is in line with previous research in which the probability of recollection rejection has been increased by repeated presentation of targets in some experiments but to a smaller extent than verbatim memory for targets (e.g., Brainerd et al., 1999; Lampinen et al., 2004). To demonstrate that the simplified CR paradigm is capable of measuring recollection rejection, we conducted two additional experiments using the priming manipulation introduced by Brainerd et al. (1999, Experiment 3).

Experiments 5 and 6: Recollection Rejection

In Experiments 5 and 6, the process of recollection rejection by way of a nonidentity judgment was examined. Recollection rejection occurs when participants are confronted with a related probe and succeed in retrieving verbatim information for the corresponding target. This verbatim information can then be compared with the probe and will result in a judgment of nonidentity and a correct rejection of the probe (or, in case of the present paradigm, in a correct classification as a related probe). To manipulate the probability of occurrence of this process, we replicated the priming manipulation that was used by Brainerd et al. (1999, Experiment 3) to demonstrate the recollection–rejection process. In the memory test, three conditions were realized: One third of participants were presented with the target probe just before the corresponding related probe was tested (target-first condition). This manipulation primes the verbatim trace of the target and should render a recollection rejection due to a nonidentity judgment more likely as compared with a second condition in which target items were tested after the corresponding related probe (target-last condition).

A Bayesian Framework for Guessing in the Simplified CR Paradigm

Experiments 5 and 6 also aimed at validating the guessing parameter a. For this purpose, a third condition was realized in which targets corresponding to related probes were not presented at all (control condition), and the base rate of targets was therefore only half that of the target-first condition. In this condition, estimates of parameter a were predicted to be smaller than in the target-first condition. For the target-last condition, intermediate values of a were predicted. These predictions were derived from a Bayesian framework of metacognitive guessing strategies (Batchelder & Batchelder, 2008) as explained in the following paragraphs.

In the simplified CR paradigm, two mnemonic states can be distinguished in which the target/related guessing process modeled by parameter a affects performance. Let m_1 be the state in which gist but no verbatim memory is available. Further, let m_2 be the state in which neither verbatim nor gist memory are available, but a given probe's gist has been guessed old as described by parameter b. In both states, participants have decided against the "unrelated" response but have to decide between the "target" and "related" response options. In the absence of relevant information in memory, this decision is guided by a guessing process modeled by parameter a.

In a given mnemonic state m_i, participants can determine the optimal response category k by computing the conditional probability $p(k|m_i)$, for the two response categories *target* and *related*. The optimal response is to select the response category k for which $p(k|m_i)$ is maximal. In states m_1 and m_2, a guessing-based decision between the response options *target* and *related* has to be made. This decision is modeled by parameter a in both states, and therefore, a joint strategy is determined for guessing in m_1 and m_2 to predict values of parameter a. Such a strategy would advise a *target* response when $p(\text{target}|m_1 \vee m_2)$ is larger than $p(\text{related}|m_1 \vee m_2)$, or stated differently, when the Bayes factor $BF_a = p(\text{target}|m_1 \vee m_2)/p(\text{related}|m_1 \vee m_2)$ is greater than 1. From $p(\text{target}|m_1 \vee m_2) = p(m_1 \vee m_2|\text{target})p(\text{target})/p(m_1 \vee m_2)$ and $p(\text{related}|m_1 \vee m_2) = p(m_1 \vee m_2|\text{related})p(\text{related})/p(m_1 \vee m_2)$, and the model equations, it follows that $BF_a = (p_{(t)}/p_{(r)})[(1 - V_t)/(1 - V_r)][G_t + (1 - G_t)b]/[G_r + (1 - G_r)b]$, with $p_{(t)}$ = proportion of targets in the test list, and $p_{(r)}$ = proportion of related distracters in the test list.

As indicated by the first term, the tendency to guess "target" is expected to increase with the proportion of target probes in the test list, and it is expected to decrease with an increasing proportion of related probes. This reflects a guessing strategy based on base rates. Note that in the case of zero memory, the Bayes factor simplifies to a simple base rate ratio, $BF_a = p_{(t)}/p_{(r)}$.

The second term indicates that the tendency to guess "target" is expected to increase when verbatim memory for target items (V_t) decreases and when the probability for recollection rejection (V_r) increases. As verbatim memory for target items increases (i.e., $V_t > V_r$), optimal guessing shifts toward related probes. Analogously, if $V_r > V_t$, then optimal guessing shifts toward the *target* response category. Thus, the ratio of verbatim memory parameters affects the optimal guessing strategy such that guessing is expected to be biased toward the class of items for which verbatim memory is weakest.

As indicated by the third term, gist memory exerts an opposite but somewhat weaker effect on the optimal guessing strategy. As the asymmetry between gist memory parameters for the two classes of items increases, so should the guessing tendency toward the class of items with stronger gist memory. The magnitude of this effect is comparable with that of the verbatim memory asymmetry only for $b = 0$, but it is rapidly attenuated as values of b depart from zero.

The first two terms are relevant for the conditions tested in Experiments 5 and 6, in which verbatim memory parameters and targets' base rates were varied but not gist memory. The second term indicates that the probability of guessing "target" is expected to be higher in conditions with higher values of V_r (i.e., the target-first conditions) than in conditions with low values of V_r (i.e., the target-last and control conditions). The first term indicates that the probability of guessing "target" is expected to be higher in conditions with higher proportions of target probes (i.e., the target-first and target-last conditions) than in conditions with equal proportions of target and related probes (i.e., the control condition). Participants' guessing tendency is expected to be influenced by a combination of both effects. Taken together, the magnitude of estimates of parameter a is expected to follow the order: target-first > target-last > control.

Method

Participants. Sixty volunteers participated in each experiment (Experiment 5: 42 women; ages ranged from 18 to 49 years, $M = 24$; Experiment 6: 32 women; ages ranged from 19 to 30 years, $M = 23$).

Design. A 3 (Priming: target-first, target-last, control) × 3 (Probe Type: target, related, unrelated) design was implemented with repeated measures on the last factor.

Materials. In Experiment 5, we used the synonym pairs; in Experiment 6, we used the categories described in the General Method section. In contrast to previous experiments, only one item was presented on the study list for each synonym pair and category, and items were presented only once. Order of presentation of the study list was randomized; order of presentation was also randomized for the test list but with the restrictions described below. Randomizations were carried out for each participant anew.

In Experiment 5, 20 synonym pairs were represented on the study list by a single item that was randomly selected. At test, these pairs were randomly split into two halves of 10 pairs each. For the first half, the item presented on the study list was shown as target probe. For the second half, the nonpresented item of the pair was shown as related probe and the presented item was used to implement the priming manipulation in the target-first and target-last conditions.

In Experiment 6, 20 categories were randomly selected to be represented on the study list by a single, randomly selected exemplar. At test, these pairs were randomly split into 10 categories for which the presented exemplar was shown as target probe and 10 categories for which the nonpresented exemplar was shown as related probe and the presented exemplar was used to implement the priming manipulation.

Procedure. Procedure was identical to that in Experiments 3 and 4 in the control condition. Departing from previous procedure, a priming manipulation was introduced in the memory test for the target-first and target-last conditions of Experiments 5 and 6. The priming manipulation was implemented as follows: In the target-first condition, before a given related probe was shown on the test list, the target to which it is related was probed. In contrast, in the target-last condition, the related probe was shown on the test list before the corresponding target was probed. In Experiment 5, target probes were either tested immediately before or after the corresponding related probe. In Experiment 6, target probes were presented before or after the corresponding related probe in a more unpredictable fashion, such that between zero and two probes taken from other lists were presented in between the target probe and the corresponding related probe. Data from the target probes used to implement the priming manipulation were discarded.

Results

Table 4 gives parameter estimates and significance tests. Across both experiments, as predicted, priming affected the recollection rejection process: V_r was improved in the target-first condition of Experiment 5, as compared with the target-last and the control conditions, $\Delta G^2_{(df = 1)} = 20.27$ and 11.70, respectively, both $p < .001$. No difference was observed between the latter two conditions, $\Delta G^2_{(df = 1)} = 0.38$, $p = .54$. This effect was replicated in Experiment 6: V_r was greater in the target-first than in the target-last condition, $\Delta G^2_{(df = 1)} = 11.49$, and greater than in the target-first and control conditions, $\Delta G^2_{(df = 1)} = 15.65$, both $p < .001$. However, there was no difference between the target-last and control conditions, $\Delta G^2_{(df = 1)} = 1.79$, $p = .18$.

No other effects on the memory parameters emerged. Neither verbatim memory for targets (V_t) nor gist memory parameters (G_t and G_r) were affected by the priming manipulation (smallest $p = .08$).

Guessing parameters. No effects were obtained on parameter *b* in Experiment 5. In contrast, this parameter was affected in Experiment 6. In the target-first condition, parameter *b* was depressed below the level observed in the target-last and control conditions, $\Delta G^2_{(df = 1)} = 7.15$ and 4.06, respectively, both $p < .05$. This reflects the fact that in the target-first condition, participants were less likely to accept an item as old for which they had no memory.

Planned comparisons were computed for parameter *a*. Remember that lower estimates of *a* were predicted in the control condition as compared with the target-first condition. These predictions were confirmed for Experiment 6, $\Delta G^2_{(df = 1)} = 10.41$, $p < .001$, and tended to be confirmed for Experiment 5, $\Delta G^2_{(df = 1)} = 3.40$, $p = .06$. In the target-last conditions, estimates of *a* had intermediate values ranging between those of the target-first and control conditions, and they did not differ significantly from the other two conditions. This pattern is consistent with the predicted rank order derived a priori, on the basis of the definition of the Bayes factor BF_a, as well as with the rank order of actual values of BF_a computed post hoc from parameter estimates. For Experiment 5, these values were .64, .28, and .10 for the target-first, target-last, and control conditions, respectively. For Experiment 6, values of BF_a were .94, .54, and .15 for the target-first, target-last, and control conditions, respectively.

Table 4
Estimates (and 95% Confidence Intervals) for the Parameters of the Simplified Conjoint Recognition Model for Experiments 5 and 6

Parameter	List condition: Target-first	Target-last	Control	$\Delta G^2_{(df = 2)}$	p
		Experiment 5			
a	.43 (.30, .57)	.28 (.16, .40)	.24 (.10, .38)	4.26	.12
b	.26 (.19, .32)	.27 (.21, .33)	.19 (.13, .24)	4.68	.10
G_t	.29 (.00, .63)	.38 (.14, .63)	.28 (.03, .53)	0.39	.82
G_r	.43 (.19, .67)	.70 (.55, .84)	.72 (.56, .88)	4.86	.09
V_t	.89 (.83, .94)	.81 (.75, .87)	.86 (.80, .91)	3.20	.20
V_r	$.72_a$ (61, .82)	$.05_b$ (.00, .42)	$.23_b$ (.00, .62)	20.89	<.01
		Experiment 6			
a	$.39_a$ (24, .55)	$.25_{ab}$ (.14, .35)	$.12_b$ (.03, .20)	10.50	<.01
b	$.19_a$ (.14, .24)	$.31_b$ (.24, .37)	$.28_b$ (.21, .34)	7.67	.02
G_t	.19 (.00, .43)	.31 (.03, .58)	.50 (.27, .73)	3.12	.21
G_r	.40 (.15, .66)	.26 (.00, .52)	.66 (.42, .89)	5.16	.08
V_t	.81 (.75, .87)	.83 (.78, .89)	.82 (.76, .88)	0.32	.85
V_r	$.73_a$ (.63, .84)	$.33_b$ (.12, .54)	$.00_b$ (.00, .64)	16.40	<.01

Note. Parameter estimates in each row that share subscripts do not differ significantly. a = probability of guessing "target"; b = probability of guessing that an item is either a target or a related probe; G_t = probability of retrieving a target's gist trace given a target probe; G_r = probability of retrieving a target's gist trace given a related probe; V_t = probability of retrieving a target's verbatim trace given a target probe; V_r = probability of retrieving a target's verbatim trace given a related probe.

Discussion

In Experiments 5 and 6, we successfully validated the model's V_r parameters as measures of the process of recollection rejection. We presented a target probe just before the corresponding related distracter was probed. As predicted, and as observed by Brainerd et al. (1999) using the original CR paradigm, the recollection rejection parameter was affected by the priming manipulation. It is concluded that the simplified CR paradigm adequately captured the effects of the priming manipulation and that the V_r parameter can be considered a valid indicator of the recollection–rejection or nonidentity process.

For the memory parameters, no other effects were significant. Brainerd et al. (1999, Experiment 3) have found that the priming manipulation reduced gist memory for related probes (G_r); in the present studies, a tendency toward such a reduction was also observed in both experiments, but this effect was not significant.

The guessing parameter *a* followed the pattern predicted by a Bayesian metacognitive account (Batchelder & Batchelder, 2008) as applied to the simplified CR paradigm. This framework can explain the pattern of guessing whether an item is a target or a related distracter (parameter *a*) on the basis of two factors: First, participants bias their responses toward the class of items that they are least likely to discriminate, that is, for which verbatim memory is weakest, on the basis of the ratio $(1 - V_t)/(1 - V_r)$.[3] According to this account, guessing is expected to be biased toward related probes (i.e., $a < .5$), and this holds true across all conditions of Experiments 5 and 6. In addition, it implies that guessing parameter *a* should decrease as V_r decreases relative to V_t as observed in the target-last and control conditions relative to the target-first condition. Second, the account implies that participants use base rate information to inform their guessing strategy, and this was reflected by the fact that participants were more likely to guess "target" in the target-first and target-last conditions that featured a greater proportion of targets in the memory test than in the control condition. Estimates of parameter *a* followed the rank order predicted by values of a Bayes factor representing a combination of both of these factors. In conclusion, the estimates of the guessing process captured by the new parameter *a* follow a pattern predicted by a Bayesian framework (Batchelder & Batchelder, 2008) that has received support in previous studies (e.g., Meiser, Sattler, & Von Hecker, 2007).

General Discussion

The simplified CR paradigm and model was successfully validated as a useful and an efficient measurement tool for researchers interested in separating memory for gist and verbatim detail. Furthermore, guessing processes were validated with a Bayesian framework.

First, the simplified CR paradigm and model provides valid indicators of gist memory: In Experiment 1, gist memory was manipulated by varying the number of items that were presented from a category (one vs. four items; cf. Brainerd et al., 2001, Experiment 3). As predicted, gist memory parameters were greater for categories from which more exemplars were presented. In Experiment 2, gist memory for related probes (i.e., gist-based false memory) was selectively manipulated by varying the number of items that were presented from a DRM list. The CR model's gist memory parameters adequately captured this selective influence, that is, G_r, but not G_t, was affected.

Second, the simplified CR paradigm also provides valid indicators of verbatim memory: In Experiments 3 and 4, targets' verbatim trace strength was increased by repeated presentation (cf. Brainerd et al., 1999, Experiments 1 and 2). This effect was adequately and selectively captured by the model's V_t parameters that represent indicators of verbatim-based identity judgments.

Third, the simplified CR paradigm provides valid indicators of recollection–rejection processes based on verbatim memory: In Experiments 5 and 6, targets' verbatim traces were primed before the corresponding related probes were tested (cf. Brainerd et al., 1999, Experiment 3), thereby strengthening the verbatim-based recollection–rejection process. This effect was adequately and selectively captured by the V_r parameters.

Guessing Processes

Batchelder and Batchelder (2008) suggested relevant metacognitive strategies for the guessing process modeled by parameter *a* (see also Meiser et al., 2007, for a similar discussion for the SM framework). Applied to the present model, their framework predicts that estimates of parameter *a* should be smaller than .5 whenever $BF_a < 1$. This was the case in all conditions in the present experiments. Furthermore, as predicted, estimates of *a* were monotonically related to BF_a in Experiments 5 and 6. The same monotonous relation can be found when estimates of *a* across Experiments 1–4 are considered.[4] Here, estimates of *a* follow the rank order Experiment 2 > Experiment 1 > Experiment 4 > Experiment 3 that was predicted by values of BF_a (0.45, 0.35, 0.23, and 0.12 for Experiment 2, Experiment 1, Experiment 4, and Experiment 3, respectively). We conclude that the guessing process measured by parameter *a* was influenced by metacognitive strategies taking into account base rates and the relative memory strengths for targets and related probes, as predicted by our application of Batchelder and Batchelder's metacognitive framework to the simplified CR paradigm.

Parameter Heterogeneity

In applying multinomial models to data aggregated across participants, it is assumed that parameters are homogeneous across participants. Violations of this assumption may lead to erroneous rejection of models as well as biased parameter estimates and confidence intervals, and significant results of parameter comparisons may be mere artifacts of data aggregation. We tested the homogeneity assumption and, when it was violated, we fitted latent-class hierarchical multinomial models (Klauer, 2006; Stahl & Klauer, 2007) that provide an extension of multinomial models accommodating parameter heterogeneity. Details of these analyses can be obtained from Christoph Stahl. Results reveal that, first, parameter heterogeneity across participants was present in 14 of 16 conditions; second, the latent-class hierarchical approach was successful in accounting for this heterogeneity; and, finally, the results reported above could be confirmed in these control analyses. Thus, parameter heterogeneity across participants did not affect the results reported or the conclusions drawn in the present research. Note that the present approach, as well as the original CR paradigm, relies on the assumption that parameters are homogeneous across items. It is desirable to investigate the present model's behavior under cases of considerable subject or item heterogeneity in more detail in future research.

A Repeated-Measures Approach

Recently, Brainerd, Reyna, Bellinge, and Myers (2007) also suggested a simplification of the CR paradigm using a repeated-

[3] Note that optimal guessing strategies as computed from Batchelder and Batchelder's (2008) framework may depend on whether an item has been judged as old by way of retrieving item or gist memory, or by guessing in the absence of memory retrieval (e.g., Meiser et al., 2007). In the present data, however, strategies consistently converge in favoring the "related" over the "target" response.

[4] As explained above, in Experiments 1–4, a single estimate of parameter *a* was obtained for both within-participant conditions (see Footnote 2), thereby assuming that guessing processes were identical for a single participant for probe items from both conditions. This assumption is plausible given the fact that guessing processes are relevant only when information from memory is not available to distinguish between probe items from both conditions. Furthermore, evaluations of model fit provided empirical support for this assumption. Consequently, given a single estimate of *a*, only one Bayes factor BF_a was computed for each experiment, on the basis of the means of V_t, V_r, G_t, and G_r across conditions.

measures procedure. In this approach, the T, R, and T + R instructions are administered to the same, single group of participants. After studying a list of words, participants were presented with target probes, related distracters, and unrelated distracters in a random order. Along with each probe, one of three questions was posed, representing the three instruction types. For the T instruction, participants were to indicate whether the probe was “a word on the list”; for the R instruction, they were to indicate whether the probe was “a synonym of a word on the list”, and for the T + R instruction, they were to indicate whether the probe was “a word on the list or a synonym of a word on the list.” Thus, in this approach, the instruction conditions from the original CR paradigm are realized in a within-subjects design, implying a considerable increase in efficiency. The original CR model is then applied to the data thus obtained, and it has been shown to provide adequate fits (Brainerd et al., 2007).

Two differences between the repeated-measurement approach and the present paradigm shall be addressed here. First, in the present paradigm, the same single question is asked for each probe, whereas different questions are asked in the repeated-measurement approach. Although there is only one group of participants in the repeated-measurement approach, which reduces the potential problems caused by different strategies across groups in the original CR paradigm, it is still possible that different response tendencies or even different retrieval strategies could be applied for different questions. This is not the case in the present approach in which a single question is asked, and differential strategies can be based only on different mnemonic states.

Second, in contrast to the present approach, the repeated-measures approach, as well as the original CR paradigm, is based on the logic of opposition as introduced by the process-dissociation procedure (Jacoby, 1991). In the opposition logic, performance in a condition in which the two processes of interest are thought to work in the same direction (i.e., the T + R condition) is contrasted with performance in conditions in which both processes work in opposition (i.e., the T and R conditions). In contrast, in the present approach, the frequencies with which the three different responses (i.e., “target,” “related,” “unrelated”) are compared for three types of memory probes in a single condition. We believe that this difference is not a fundamental one because data obtained from the process-dissociation procedure can be accounted for by models that were originally developed for a task not involving the opposition logic (the SM model; Buchner, Erdfelder, Steffens, & Martensen, 1997; Yu & Bellezza, 2000). The opposition logic was developed as a methodological improvement over the use of different tasks to separate underlying processes (i.e., “implicit” and “explicit” memory tests), and the multinomial modeling approach can be seen as an extension of the opposition approach that is capable of providing a more fine-grained measurement of the underlying cognitive processes within a single experimental paradigm (see Brainerd et al., 1999, for a related discussion).

To conclude, in our view, the repeated-measures paradigm (Brainerd et al., 2007) appears to be a promising alternative to the present approach, and we are confident that its validity can be demonstrated in future research.[5]

Extending the Model

The simplified CR paradigm and model presented here has been shown to provide valid indicators of gist memory, verbatim memory, recollection rejection, and guessing—the processes considered in the original exposition of the CR model by Brainerd et al. (1999). However, the present version of the model does not provide measures for some processes of interest to memory researchers. In the following, we discuss ways to extend the present model toward addressing the processes of phantom recollection, erroneous recollection rejection, and the detection of new items as new.

Neither the original CR model nor the present model considers the possibility that participants might detect new items as new. This might occur for example by way of autonoetic processes (e.g., “I would have remembered this item, had it been presented”; Strack & Bless, 1994). It is possible to extend the present model to include such a detection process by adding a parameter D_N to the third tree diagram in Figure 2 such that a new distracter would then be detected as new with probability D_N. With probability $1 - D_N$, no such detection as new would occur, and the response would be determined by guessing processes as already illustrated in the diagram.

Note that an extended model including D_N would need to incorporate at least one additional assumption to be identified for the present paradigm. For example, D_N could be set equal to some other memory parameter (e.g., V_r or G_r); similar restrictions are used to address the identifiability problem of the D_N parameter in SM models (e.g., Bayen et al., 1996). The present model is derived from the thus extended model by assuming that $D_N = 0$ after pretests revealed that a model incorporating this restriction provided better fit than one in which detection as new was equated to other memory processes. The assumptions underlying any such restriction can be questioned on theoretical grounds. However, for the purpose of using the model as a measurement tool, this is not crucial as long as its parameter estimates provide good approximation of the processes of interest. As an alternative to incorporating additional technical assumptions it is also possible to achieve identifiability of D_N by extending the empirical basis to include additional experimental conditions (e.g., Bayen et al., 1996).

The original CR model has been extended to accommodate two recently discovered phenomena of recollective experience and memory performance: *phantom recollection* and *erroneous recollection rejection* (Brainerd, Payne, et al., 2003; Brainerd et al., 2001). These phenomena have been found to occur in a relatively restricted set of situations (e.g., in studies that used DRM methodology to elicit high levels of gist-based false memory; Brainerd & Wright, 2005; Brainerd et al., 2001) and to be negligible in

[5] A validation is important because, just as in the present simplified CR paradigm, procedural changes might affect the cognitive processes contributing to performance. For example, in the repeated-measures procedure, participants are presented with targets as well as with their corresponding related distracters on the same test list. This resembles the priming procedure implemented in the present Experiments 5 and 6, in which probing the target before the corresponding related distracter affected the recollection–rejection process for that related distracter (see also Brainerd et al., 1999, Experiment 3).

memory test situations that lack such features (e.g., Brainerd et al., 2007). Therefore, these processes were not considered in the present work, as we aimed to provide and test a simple paradigm for the assessment of verbatim and gist memory that is applicable in a wide range of situations. However, investigations of these processes can provide important insight (e.g., Brainerd & Wright, 2005), and we would like to point out that the simplified CR paradigm can easily be extended to measure these additional processes. In the following, we illustrate how an integration of the phantom recollection and erroneous recollection rejection processes into the present model can be achieved.

Phantom recollection describes the phenomenon that when gist-based false memories arise at high levels, a subset of those false memories may be accompanied by false recollective experiences, as has often been reported in research that has used the DRM paradigm. Given that a related distracter at test elicits such vivid false memories, it is likely to be mistaken as a target. Phantom recollection can be modeled by an additional parameter P_r (Brainerd et al., 2001). In an extended version of our simplified CR model, phantom recollection would occur with probability $(1 - V_r)P_r$, describing the probability of a *target* response to a related distracter in the absence of recollection rejection. Gist-based false memories that lack such vivid recollective experiences as are required for phantom recollection would occur with probability $(1 - V_r)(1 - P_r)G_r$ and would be followed by a guessing process as modeled by parameter a. In the absence of memory, related distracters can be accepted as old on the basis of guessing processes with probability $(1 - V_r)(1 - P_r)(1 - G_r)b$, and guessing (parameter a) then determines classification as target or related probe.[6]

Erroneous recollection rejection can occur when a target probe for which no verbatim trace is available acts as a retrieval cue for a related target's verbatim trace. In this situation, participants might falsely classify the target probe as a related distracter by the same logic by which recollection rejection of related distracters is thought to operate. Erroneous recollection rejection can be modeled by an additional parameter E_t in the submodel for target probes. In an extended version of our CR model, erroneous recollection rejection of targets would occur with probability $(1 - V_t)E_t$, describing the probability of a *related* response to a target probe in the absence of retrieval of the target's verbatim trace. Gist-based acceptance of the target probe then occurs with probability $(1 - V_t)(1 - E_t)G_t$, followed by guessing as modeled by parameter a. In the absence of memory, target probes are accepted on the basis of guessing processes with probability $(1 - V_t)(1 - E_t)(1 - G_t)b$ and classified as targets or related probes by guessing (parameter a).

Applying an extended model with additional parameters requires additional degrees of freedom. They can be obtained by including additional experimental conditions that differ with regard to some but not all of the processes measured by the model. For example, a between-participants manipulation of the proportion of unrelated distracters on the test list could be implemented that would affect guessing processes, whereas memory parameters could be set equal across this additional factor. As a result of this restriction, sufficient degrees of freedom would be available for the addition of parameters for the processes of phantom recollection, erroneous recollection rejection, and detection of new items as new.

[6] To illustrate, we applied an extended version of our CR model to investigate phantom recollection in Experiment 2. Of the present experiments, this is the only one in which the phantom recollection process might have occurred (cf. Brainerd & Wright, 2005; Brainerd et al., 2001). Two P_r parameters were added to the simplified CR model, one for each level of the Gist Activation factor. The required degrees of freedom were obtained by restricting the V_r parameters to zero (recall that analyses with the simplified CR model have yielded zero values for the recollection rejection process measured by V_r; see Table 2). The model fitted the data well, $G^2 = 0$, and estimates of P_r were at .11 (.00, .23) for weak gist, and at .15 (.00, .32) for strongly activated gist (95% confidence intervals in parentheses). However, P_r parameters were not significantly different from zero, largest $G^2_{(df\ =\ 1)} = 2.56$, smallest $p = .11$, supporting our assumption that phantom recollection did not substantially affect performance in the present experiments. P_r parameters were also not affected by Gist Activation, $G^2_{(df\ =\ 1)} = 0.27$, $p = .6$.

References

Batchelder, W. H., & Batchelder, E. (2008). Meta-cognitive guessing strategies in source monitoring. In J. Dunlosky & R. A. Bjork (Eds.), *Handbook of metamemory and memory* (pp. 211–244). Mahwah, NJ: LEA Publishers.

Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review, 97,* 548–564.

Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin & Review, 6,* 57–86.

Batchelder, W. H., Riefer, D. M., & Hu, X. (1994). Measuring memory factors in source monitoring: Reply to Kinchla. *Psychological Review, 101,* 172–176.

Bayen, U. J., Murnane, K., & Erdfelder, E. (1996). Source discrimination, item detection, and multinomial models of source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 197–215.

Brainerd, C. J., Forrest, T. J., Karibian, D., & Reyna, V. F. (2006). Development of the false-memory illusion. *Developmental Psychology, 42,* 962–979.

Brainerd, C. J., Holliday, R. E., & Reyna, V. F. (2004). Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child Development, 75,* 505–522.

Brainerd, C. J., Payne, D. G., Wright, R., & Reyna, V. F. (2003). Phantom recall. *Journal of Memory and Language, 48,* 445–467.

Brainerd, C. J., & Reyna, V. F. (2002). Fuzzy-trace theory and false memory. *Current Directions in Psychological Science, 11,* 164–169.

Brainerd, C. J., Reyna, V. F., Bellinge, H., & Myers, J. (2007). *Conjoint recognition and the word-frequency mirror effect.* Manuscript submitted for publication.

Brainerd, C. J., Reyna, V. F., & Mojardin, A. H. (1999). Conjoint recognition. *Psychological Review, 106,* 160–179.

Brainerd, C. J., Reyna, V. F., Wright, R., & Mojardin, A. H. (2003). Recollection rejection: False-memory editing in children and adults. *Psychological Review, 110,* 762–784.

Brainerd, C. J., Stein, L. M., & Reyna, V. F. (1998). On the development of conscious and unconscious memory. *Developmental Psychology, 34,* 342–357.

Brainerd, C. J., & Wright, R. (2005). Forward association, backward association, and the false-memory illusion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 554–567.

Brainerd, C. J., Wright, R., Reyna, V. F., & Mojardin, A. H. (2001). Conjoint recognition and phantom recollection. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 307–327.

Brown, V., & Gorfein, D. S. (2004). A new look at recognition in the

Brown–Peterson distractor paradigm: Toward the application of new methodology to unsolved problems of recognition memory. *Memory & Cognition, 32,* 674–685.

Buchner, A., Erdfelder, E., Steffens, M. C., & Martensen, H. (1997). The nature of memory processes underlying recognition judgments in the process dissociation procedure. *Memory & Cognition, 25,* 508–517.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Cowan, N. (1998). Children's memories according to fuzzy-trace theory: An endorsement of the theory's purpose and some suggestions to improve its application. *Journal of Experimental Child Psychology, 71,* 144–154.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology, 58,* 17–22.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191.

Gallo, D. A. (2004). Using recall to reduce false recognition: Diagnostic and disqualifying monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 120–128.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language, 30,* 513–541.

Johnson, M. K., Hashtroudi, S., & Lindsay, D. S. (1993). Source monitoring. *Psychological Bulletin, 114,* 3–28.

Klauer, K. C. (2006). Hierarchical multinomial processing tree models: A latent-class approach. *Psychometrika, 71,* 7–31.

Lampinen, J. M., Odegard, T. N., & Neuschatz, J. S. (2004). Robust recollection rejection in the memory conjunction paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 332–342.

Lampinen, J. M., Watkins, K. N., & Odegard, T. N. (2006). Phantom ROC: Recollection rejection in a hybrid conjoint recognition signal detection model. *Memory, 14,* 655–671.

Marsh, R. L., & Hicks, J. L. (1998). Test formats change source-monitoring decision processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1137–1151.

Meiser, T., Sattler, C., & Von Hecker, U. (2007). Metacognitive inferences in source memory judgments: The role of perceived differences in item recognition. *Quarterly Journal of Experimental Psychology, 60,* 1015–1040.

Odegard, T. N., & Lampinen, J. M. (2005). Recollection rejection: Gist cueing of verbatim memory. *Memory & Cognition, 33,* 1422–1430.

Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review, 95,* 318–339.

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 803–814.

Rotello, C. M. (2001). Recall processes in recognition memory. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 183–221). San Diego, CA: Academic Press.

Seamon, J. G., Luo, C. R., Schwartz, M. A., Jones, K. J., Lee, D. M., & Jones, S. J. (2002). Repetition can have similar or different effects on accurate and false recognition. *Journal of Memory and Language, 46,* 323–340.

Stahl, C. (2004). *Enkodierung erwartungskongruenter und erwartungsinkongruenter Information* [Encoding of expectancy-congruent and expectancy-incongruent information]. Unpublished doctoral dissertation, Universität Freiburg, Freiburg, Germany.

Stahl, C. (2006). Multinomiale Verarbeitungsbaummodelle in der Sozialpsychologie [Multinomial processing tree models in social psychology]. *Zeitschrift für Sozialpsychologie, 37,* 161–171.

Stahl, C., & Klauer, K. C. (2007). HMMTree: A computer program for hierarchical multinomial processing tree models. *Behavior Research Methods, 39,* 267–273.

Stahl, C., & Klauer, K. C. (in press). The effect of incongruency and cognitive load on short-term verbatim and gist memory in the conjoint recognition paradigm. In F. Columbus (Ed.), *New research on short-term memory.* Hauppauge, NY: Nova Science.

Stegt, S. J. (2006). *Die Rolle impliziter assoziativer Reaktionen bei der Entstehung von Pseudoerinnerungen im DRM-paradigma* [The role of implicit associative responses in the generation of pseudomemories in the DRM paradigm]. Berlin, Germany: Logos.

Strack, F., & Bless, H. (1994). Memory for nonoccurences: Metacognition and presuppositional strategies. *Journal of Memory and Language, 33,* 203–217.

Wright, D. B., & Loftus, E. F. (1998). How misinformation alters memories. *Journal of Experimental Child Psychology, 71,* 155–164.

Yu, J., & Bellezza, F. S. (2000). Process dissociation as source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 1518–1533.

(*Appendix follows*)

Appendix

The Simplified Conjoint Recognition (CR) Model

Data Representation

In the simplified CR paradigm, the memory test list consists of target probes, related probes, and unrelated probes. The participant is required to classify each probe as either a target probe, a related distracter, or an unrelated distracter. This yields Table A1, which is a 3×3 frequency table, where F_{ij} is the frequency of event E_{ij}, that is, of a response of type j to a probe of type i.

Model Equations

A multinomial model is created by expressing the probability p_{ij} of an event E_{ij} as a function of latent parameters that represent the psychological processes that are to be measured. In a multinomial model processing tree model, the probability of an event is expressed as the sum of the probabilities of the branches that lead to that event. The following model equations can thus be easily derived from Figure 2. For target probes, the empirical probabilities are given by the following three equations:

$$p_{11} = V_t + (1 - V_t)G_t a + (1 - V_t)(1 - G_t)ba \quad (1a)$$

$$p_{12} = (1 - V_t)G_t(1 - a) + (1 - V_t)(1 - G_t)b(1 - a) \quad (1b)$$

$$p_{13} = (1 - V_t)(1 - G_t)(1 - b) \quad (1c)$$

For related probes, the empirical probabilities are modeled as follows:

$$p_{21} = (1 - V_r)G_r a + (1 - V_r)(1 - G_r)ba \quad (2a)$$

$$p_{22} = V_r + (1 - V_r)G_r(1 - a) + (1 - V_r)(1 - G_r)b(1 - a) \quad (2b)$$

$$p_{23} = (1 - V_r)(1 - G_r)(1 - b) \quad (2c)$$

For unrelated probes, the empirical probabilities are modeled as follows:

$$p_{31} = ba \quad (3a)$$

$$p_{32} = b(1 - a) \quad (3b)$$

$$p_{33} = 1 - b \quad (3c)$$

Together, these equations define the simplified CR model.

Identifiability

The simplified CR model is identified. That is, given a set of probabilities p_{ij} with $0 < p_{ij} < 1$ that conform to the model equations, simple algebraic manipulations reveal that the parameters can be determined as follows:

$$a = \frac{p_{31}}{p_{31} + p_{32}} \quad (4)$$

$$b = 1 - p_{33} \quad (5)$$

$$V_t = p_{11} - \frac{p_{31}}{p_{32}} p_{12} \quad (6)$$

$$V_r = p_{22} - \frac{p_{32}}{p_{31}} p_{21} \quad (7)$$

$$G_t = \frac{p_{12}p_{33}(1 - p_{33}) - p_{13}p_{32}(1 - p_{33})}{p_{12}p_{33}(1 - p_{33}) + p_{13}p_{32}p_{33}} \quad (8)$$

$$G_r = \frac{p_{21}p_{33}(1 - p_{33}) - p_{23}p_{31}(1 - p_{33})}{p_{21}p_{33}(1 - p_{33}) + p_{23}p_{31}p_{33}} \quad (9)$$

This demonstrates identifiability for the simplified CR model.

Model Fit

The model imposes inequality constrains on the probabilities p_{ij}. A set of probabilities p_{ij} that conforms to the model is associated with the parameter values given in Equations 4–9. It is readily seen that $0 < p_{ij} < 1$ immediately implies $0 < a < 1$, $0 < b < 1$, $V_t < 1$, $V_r < 1$, $G_t < 1$, and $G_r < 1$. Because the model parameters are themselves probabilities, a conforming set of probabilities must also lead to parameter values G_t, G_r, V_t, and V_r that are nonnegative. Using Equations 6–9, one can see that V_t, V_r, G_t, and G_r are nonnegative if the following inequality constrains are satisfied:

$$p_{31}p_{12} \leq p_{11}p_{32} \quad (10a)$$

$$p_{32}p_{21} \leq p_{22}p_{31} \quad (10b)$$

$$p_{13}p_{32} \leq p_{12}p_{33} \quad (10c)$$

$$p_{23}p_{31} \leq p_{21}p_{33} \quad (10d)$$

A set of probabilities p_{ij} with $0 < p_{ij} < 1$ conforms to the model equations if and only if it satisfies the above four inequality constrains. To see this, note first that a set of probabilities that conforms to the model is generated by the parameters given in Equations 4–9 with parameter values that fall between zero and one. Nonnegativity of V_t, V_r, G_t, and G_r implies in particular that the four inequality constrains must be satisfied. Conversely, if the constraints are satisfied for a set of probabilities p_{ij} with $0 < p_{ij} < 1$, the parameter values computed by means of Equations 4–9 all fall between zero and one; that is, they are admissible. Simple manipulations show that the p_{ij} are generated by means of the model equations with these parameter values.

Testing model fit therefore amounts to testing whether there are significant violations of the above four inequality constraints. Each constraint can be seen as stating that an odds ratio is not larger than one. For example, the first constraint states that the odds for

Table A1
Data Structure of the Simplified Conjoint Recognition Paradigm

	Response		
Probe	t	r	u
t	F_{11}	F_{12}	F_{13}
r	F_{21}	F_{22}	F_{23}
u	F_{31}	F_{32}	F_{33}

Note. t = target probe; r = related distracter; u = unrelated distracter.

responding "target" rather than "related" in reaction to unrelated probes (p_{31}/p_{32}) is smaller than, or at most equal to, the corresponding odds computed for responses in reaction to targets (p_{11}/p_{12}). For each constraint and experimental condition, we computed a standard chi-square test for the departure of the appropriate odds ratio from one. The test has one degree of freedom, and because the constraint specifies a direction, we conducted one-tailed tests. For experiments with two conditions (Experiments 1–4), this yields eight tests per experiment; for Experiments 5 and 6 with three conditions each, this yields 12 tests per experiment. In total, 56 tests were conducted, not all of which are statistically independent. In 50 cases, the constraint was satisfied empirically and therefore not significantly violated. In 6 cases, the estimated odds ratio departed from one in the violating direction, but in no case was the size of the departure significant, largest $\chi^2(1) = 2.57$, smallest p(one-tailed) = .054. All of these nonsignificant violations concerned the second constraint above that is to ensure that V_r is nonnegative; as can be seen in Tables 2–4, V_r was often small in the present data. Given that 56 tests were computed, one might have expected a few significant violations by chance alone; yet, none of the 56 tests revealed a significant violation of the model. Taken together, there was little evidence for misfit of the model.

CR and Source Monitoring (SM)

In this section, the simplified CR model is compared with the family of one-high-threshold SM models as presented by Batchelder and Riefer (1990), and the simplified procedure is shown to be very similar to the typical SM paradigm (e.g., Batchelder & Riefer, 1990; Bayen et al., 1996; Johnson, Hashtroudi, & Lindsay, 1993). In the following, a SM model (Model 6c; Batchelder & Riefer, 1990) is compared with the simplified CR model to demonstrate that both models are mere reparametrizations of each other.

In the study phase of a SM paradigm, participants are presented with items from two sources, A and B. In the memory test, old items from both sources are mixed with new items, and participants have to classify each probe as either from Source A, Source B, or new. The data structure of the SM model can thus be described by Table A2.

Table A2
Data Structure of the Source Monitoring Paradigm

	Response		
Source	A	B	New
A	F_{11}	F_{12}	F_{13}
B	F_{21}	F_{22}	F_{23}
New	F_{31}	F_{32}	F_{33}

It is easy to see the structural overlap between the SM and CR data structures. When Source A items are equated with target probes, Source B items are equated with related probes, and new items are equated with unrelated probes, the data structure is seen to be equivalent for the CR and SM models. Note at this point that there are also differences between the two paradigms, which are discussed below.

Next, we will introduce the SM model and show that the parameters of the CR model can be expressed as a function of the parameters of the SM model, and vice versa. Responses in the SM paradigm are modeled as a function of memory for the item itself (parameter D), memory for its source (parameter d), and two guessing processes (parameters a and b). Given item memory and source memory, an old item from Source A is correctly classified as a Source A item (with probability $D_A d_A$). If, with probability $D_A(1 - d_A)$, an item has been recognized as old but source memory is lacking, participants are left to guess the item's source; with probability a, the item is then classified as a Source A item, and with probability $(1 - a)$, it is classified as a Source B item. If, with probability $(1 - D_A)$, the item is not recognized (i.e., item memory is lacking), it can still be classified as old by way of guessing. When this happens (with probability b), a source is assigned by way of the source guessing process as described above that is captured by parameter a. With probability $(1 - b)$, the probe is classified as new. The SM model is given by the following equations. For items from Source A,

$$p_{11} = D_A d_A + D_A(1 - d_A)a + (1 - D_A)ba \qquad (11a)$$

$$p_{12} = D_A(1 - d_A)(1 - a) + (1 - D_A)b(1 - a) \qquad (11b)$$

$$p_{13} = (1 - D_A)(1 - b) \qquad (11c)$$

For items from Source B,

$$p_{21} = D_B(1 - d_B)a + (1 - D_B)ba \qquad (12a)$$

$$p_{22} = D_B d_B + D_B(1 - d_B)(1 - a) + (1 - D_B)b(1 - a) \qquad (12b)$$

$$p_{23} = (1 - D_B)(1 - b) \qquad (12c)$$

And for new items,

$$p_{31} = ba \qquad (13a)$$

$$p_{32} = b(1 - a) \qquad (13b)$$

$$p_{33} = (1 - b) \qquad (13c)$$

Given a set of probabilities p_{ij} generated from the CR model by Equations 1–3 with parameters a_{CR}, b_{CR}, V_t, V_r, G_t, and G_r, it is easy to see that the same set of probabilities is generated from the SM model with parameters

$$a_{SM} = a_{CR},$$

$$b_{SM} = b_{CR},$$

$$D_A = V_t + G_t - V_t G_t,$$

$$d_A = \frac{V_t}{V_t + G_t - V_t G_t},$$

$$D_B = V_r + G_r - V_r G_r, \quad \text{and}$$

$$d_B = \frac{V_r}{V_r + G_r - V_r G_r},$$

(*Appendix continues*)

where we assume that the denominators in the equations for d_A and d_B are positive. The case that either of these denominators is zero is trivial. For example, if $V_t + G_t - V_tG_t$ is zero, both V_t and G_t must be zero. The appropriate values for D_A and d_A are then that both are zero.

Thus, the parameters of the SM model can be expressed as a function of the parameters of the CR model. Conversely, given a set of probabilities generated by the SM model, it is easy to see that the same set of probabilities is generated from the CR model with parameters given by

$$a_{CR} = a_{SM},$$

$$b_{CR} = b_{SM},$$

$$V_t = D_A d_A,$$

$$G_t = \frac{D_A - D_A d_A}{1 - D_A d_A},$$

$$V_r = D_B d_B, \text{ and}$$

$$G_r = \frac{D_B - D_B d_B}{1 - D_B d_B}.$$

It follows that the CR model and the SM model generate the same set of probabilities, and the parameter values from both models are reparametrizations of each other (the case that one of the denominators in the equations for G_r and G_t is zero is again trivial).

Both models thus account for the same data but with different parameters. Just like the size and shape of a rectangle could be equivalently characterized by specifying height and width, or alternatively, by area and the ratio of height to width, the parameters are driven by the same underlying processes or dimensions, but combine them in different ways. Which of these models is more appropriate psychologically to account for the present data? Model fit must be equivalent and cannot help to discriminate between the models. In the following section, we describe the effects of the experimental manipulations on the SM parametrization to see whether it provides a more parsimonious and more easily interpretable account of the data.

Empirical Comparison of the CR and SM Models

How did the parameters of the SM model react to the experimental manipulations? Remember that for the CR model, gist memory parameters G_t (Experiment 1) and G_r (Experiments 1 and 2) were boosted by an increase in concept activation (four vs. one study items per concept). In the SM model, these effects were reflected by significant increases in both item memory parameters, D_t and D_r, in both experiments. In contrast to gist memory parameters, item memory parameters did not reflect the difference in structure between the category and Deese–Roediger–McDermott (DRM; Deese, 1959; Roediger & McDermott, 1995) lists used in Experiments 1 and 2, respectively.

In Experiments 3 and 4, the verbatim memory parameter, V_t, was expected and found to be greater for items that were presented repeatedly. In the SM model, effects of the repetition manipulation were observed on the d_t parameters in both experiments, reflecting an increase in source memory for repeated targets. In addition, two effects were observed on the item memory parameters. Item memory for targets, D_t, was significantly increased in Experiment 3 but not in Experiment 4. In addition, item memory for related probes, D_r, was significantly increased in the repeated-presentation condition in Experiment 4 but not in Experiment 3. Compared with the single and predicted effect on V_t in the CR model, this pattern is conceptually more difficult to explain.

Finally, in Experiments 5 and 6, the recollection rejection process was boosted by way of a priming manipulation, and this was reflected in the recollection rejection parameter, V_r. In the SM model, the priming manipulation affected both item and source memory for related items, D_r and d_r. In summary, the CR model provided a more parsimonious account of the effects of the manipulations than the SM model. Note, however, that the experimental manipulations were designed to dissociate verbatim and gist memory, and the measurement purpose was to find valid estimates for verbatim and gist memory. The parametrization implied by the SM model is likely to render the more easily interpretable results when experimental manipulations focus on dissociating item and source memory. A firm decision on whether item and source memory versus verbatim and gist memory are the "true" underlying factors cannot therefore be based on the measurement model per se. In the present case, a point in favor of the CR parameterization is, however, that the effects on the model parameters were predicted from the parent theory (i.e., from fuzzy trace theory), whereas it may be more difficult to deduce the effect patterns observed for the SM parameters from a SM framework.

The above comparison is based on the assumption that the task of discriminating between targets and related probes can be compared with the task of discriminating between items from two different sources. These discrimination tasks are comparable in that memory for the details of an item's presentation episode is required in both. However, there are also obvious differences between the discrimination tasks that participants perform in the CR versus the SM paradigms. First, whereas in the CR paradigm, participants discriminate between target probes that were presented at study and related probes that were *not* presented at study, the required discrimination in the SM paradigm is between two types of *presented* probes that differ with regard to the source of their presentation. Second, as a result of this fact, in the CR paradigm memory-based decision processes for target probes as well as for related probes are based on the target's memory traces only, as there can be no memory traces for the nonpresented related probe. In contrast, in the SM paradigm, memory information—item and source memory—is potentially available for both types of old items.

Given these differences, it may not be surprising that the effects obtained in the present experiments cannot readily be accounted for in terms of the parameters of the SM model. Nevertheless, the comparison of the two models illustrates that the effects cannot be explained equally well by *any* model, that is, there are many reparametrizations of each model that are statistically equivalent but differ with regard to psychological plausibility. This underlines the importance of demonstrating a model's psychological validity before using it as a measurement model. The fact that the findings can be accounted for in a plausible and parsimonious manner by the parameters of the simplified CR model provides support for the specific choice of parametrization implemented in the present model.

Received May 8, 2007
Revision received December 10, 2007
Accepted December 17, 2007 ■