

Conceptual and perceptual factors in the picture superiority effect

Georg Stenberg

Kristianstad University, Kristianstad, Sweden

The picture superiority effect, i.e., better memory for pictures than for corresponding words, has been variously ascribed to a conceptual or a perceptual processing advantage. The present study aimed to disentangle perceptual and conceptual contributions. Pictures and words were tested for recognition in both their original formats and translated into participants' second language. Multinomial Processing Tree (Batchelder & Riefer, 1999) and MINERVA (Hintzman, 1984) models were fitted to the data, and parameters corresponding to perceptual and conceptual recognition were estimated. Over three experiments, orienting tasks were varied, with neutral (Exp. 1), semantic (Exp. 2), and perceptual (Exp. 3) instructions, and the encoding manipulations were used to validate the parameters. Results indicate that there is picture superiority in both conceptual and perceptual memory, but conceptual processing makes a stronger contribution to the advantage of pictures over words in recognition.

The fact that pictures are generally better remembered than words has been known for a long time (Kirkpatrick, 1894). The picture superiority effect in memory applies to both recall and recognition (Madigan, 1983; Paivio, 1991). In picture recognition, performance can reach astounding levels. In one study (Standing, Conezio, & Haber, 1970), participants studied over 2000 pictures at a rate of 10 seconds each, and were over 90% accurate in a recognition test several days later. Although picture superiority over words is a reliable and reproducible phenomenon, it is constrained by some limiting conditions. Both encoding tasks (Durso & Johnson, 1980) and retrieval conditions (Weldon & Roediger, 1987) have been shown capable of abolishing or reversing picture superiority. The exploration of its boundary conditions, as well as recent neuroscience investigations of picture and word processing in the brain (Federmeier & Kutas, 2001; Kazmerski & Friedman,

Correspondence should be addressed to Georg Stenberg, School of Behavioural Sciences, Kristianstad University, SE-291 88 Kristianstad, Sweden. E-mail: georg.stenberg@home.sol.se

This study was supported by a grant from the Swedish Research Council. I am grateful to Sara Denward for her help in running the experiments.

1997; Köhler, Moscovitch, Winocur, & McIntosh, 2000; Schloerscheidt & Rugg, 2004), have renewed interest in the picture superiority effect. In spite of its long history, the phenomenon has not yet been given a universally accepted explanation (for reviews, see Engelkamp & Zimmer, 1994; Madigan, 1983).

Some theorising about the effect has drawn on the distinction between perceptual and conceptual processing in memory (Jacoby, 1983). Because the surface features of pictures are generally more varied and distinctive than those of words, it has often been hypothesised that the memory effect can be ascribed to those perceptual features. Drawings or paintings of, e.g., a horse can offer an endless variety of colour, texture, perspective, and size, while still preserving the meaning of the concept it illustrates, and all the different varieties can offer many opportunities for distinctive memory encoding. The possibilities of variation in pronunciation, accent, and timbre in a spoken word—or the varieties of typeface, style, and size in a printed word—seem meagre in comparison. The picture superiority effect could therefore be based on richer perceptual encoding.

PERCEPTUAL SUPERIORITY

The idea that picture superiority in memory is closely linked to the perceptual format in which study items are presented has been elaborated upon in several theories. In Paivio's dual-code theory (1971, 1986, 1991), two memory stores are assumed to handle verbal and pictorial input separately. The result is picture superiority, because pictures are often encoded in both memory stores whereas words are not; pictures invoke naming upon study more often than words invoke imagery. The use of two codes instead of one seems likely to cause better memory performance, but it cannot be the only explanation, because pictures are not always spontaneously named (Nelson & Brooks, 1973), and pictures preserve their memory advantage even without naming (Madigan, 1983). Therefore, a further assumption of Paivio's theory is that the image code is inherently better than the verbal code in producing memory performance. In Paivio's theory, the format of the perceptual input has a decisive influence on how well an item will be preserved in memory.

Other theories have not gone so far in connecting storage to perceptual form. The sensory-semantic model (Nelson, Reed, & McEvoy, 1977) posits a semantic store, in which the meanings of events are represented in a common code, regardless of input format. Additionally, perceptual aspects of the input are also stored, although in a code specific to the input format, be it pictorial or verbal. It is in this sensory code that pictures particularly excel; although picture superiority can be caused by both sensory and semantic

factors, sensory distinctiveness is suggested to be the more likely explanation.

In sum, it seems to be a widely held view that the special surface characteristics of pictures, in particular their perceptual distinctiveness, have a lot to do with their memorability. Also, outside of the particular research context of the picture superiority effect, recent memory research has come to pay attention to the special importance of perceptual features. As shown by findings in the field of implicit memory, retention of perceptual characteristics may be very long-lasting (Jacoby & Dallas, 1981) and effective in producing a sense of familiarity upon renewed acquaintance, even when conscious recollection has faded (Mandler, 1980). The contribution of perceptual fluency to the experience of memory can be substantial, and when manipulated at test, it can create deceptive illusions of memory (Johnston, Dark, & Jacoby, 1985; Whittlesea, 1993).

However, when applied to the picture superiority effect, the hypothesis of a special role for perceptual features encounters a few problems. First, pictures are not only better recognised, they are also better recalled (Madigan, 1983). In other words, the particular perceptual format need not be present at test for picture memorability to assert itself. Second, adding more perceptual detail and complexity to pictures does not add appreciably to memory performance in recognition. Nelson, Metzler, and Reed (1974) compared recognition memory for detailed colour photographs, embellished or simple line drawings, and words. The three types of pictures were all better recognised than words, but the different degrees of complexity did not make a difference. Third, in a direct test of the perceptual distinctiveness hypothesis, Kinjo and Snodgrass (2000) compared within-form priming for pictures and words. Using perceptual tasks (picture and word fragment identification, calibrated for equal difficulty), they expected study of a picture to facilitate identification of it in a degraded version, more so than the study of a word would prime an equally degraded word. Basing this expectation on the greater sensory distinctiveness of pictures, they nevertheless found only weak support for it. Over three experiments, which varied the encoding task, picture superiority in within-form priming was found in only one. The lack of an implicit superiority effect was brought into relief by an explicit picture superiority in the free recall tasks that concluded all three experiments.

CONCEPTUAL SUPERIORITY

The alternative view is that conceptual processing determines picture superiority. In this approach, it is assumed that pictures are processed semantically with a higher probability or higher efficiency than words are. Given the mostly conceptual nature of explicit memory tests such as recall

and recognition, processing at test is better matched by similarly conceptual processing at study, and this gives pictures an advantage, in accordance with the principle of transfer appropriate processing (Weldon, Roediger, & Challis, 1989).

A strength of this view is that it correctly predicts some constraints on the picture superiority effect. If the test requires perceptual processing of words, then studied words will be at an advantage relative to studied pictures because of the overlap between processes at study and at test. Such a reversal of the picture superiority effect has been observed in both implicit (Weldon & Roediger, 1987) and explicit (Weldon et al., 1989) memory tasks requiring word processing at test.

Analyses of the subjective experience of memory (Gardiner & Richardson-Klavehn, 2000) have indicated that it is sometimes accompanied by a vivid sense of the full context (“remember”), and sometimes only a vague sense of familiarity (“know”). The act of remembering is often thought to require conceptual processing, and the finding that picture recognition is accompanied by a higher proportion of “remember” responses than word recognition can be interpreted as consistent with a higher degree of conceptual processing for pictures (Rajaram, 1993).

The idea that conceptual processing of pictures is facilitated relative to that of words is supported by findings from semantic memory retrieval tasks. When pictures and words undergo speeded categorisation (e.g., as “living” or “dead”), reaction times for pictures are often faster (Potter & Faulconer, 1975; Smith & Magee, 1980). An important role for conceptual factors in episodic memory is indicated by data showing that when incomplete pictures are shown, and memory performance is conditionalised on whether the stimulus is given a meaningful interpretation by the participant or seen as a meaningless pattern, subsequent recognition is better for those stimuli that were given an interpretation (Wiseman & Neisser, 1974).

An opportunity for testing the conceptual-processing advantage of pictures is provided by conceptual implicit tests, such as category exemplar production or word association. The prediction of picture superiority in such tasks was tested by Weldon and Coyote (1996), who found no support for it. They concluded that conceptual processing does not provide pictures with their undeniable explicit memory advantage—instead, visual distinctiveness does.

There have been findings of picture superiority in conceptual implicit tests (Nicolas, 1995; Vaidya & Gabrieli, 2000; Wippich, Melzer, & Mecklenbraeuer, 1998), but they have mostly been confined to some encoding conditions and some implicit memory tasks, but not others. Nicolas (1995) found pictures to produce more priming than words in a category exemplar production task. The findings of Wippich et al. (1998) are many-faceted, but they include picture superiority on a conceptual implicit test. However, this

was found only with a shallow level of processing at study, and words produced approximately equal amounts of priming with pictures after a deep encoding task. Similarly, Vaidya and Gabrieli (2000) found picture superiority in a conceptual implicit test (category exemplar production), but only after one encoding task (naming) and not another, presumably deeper task (categorisation). Also, the type of implicit test proved to be important, with only a production task producing picture–word differences, and no differences emerging in a more passive task (verification).

In summary, implicit tests have not produced unqualified support for the idea that pictures undergo more conceptual processing. The expected pattern tends to be produced primarily by shallow encoding tasks, and the outcome of other encoding seems to be difficult to predict.

A MODELLING APPROACH

Computational modelling has been used to complement the information provided by memory tests. McBride and Doshier (2002) separated conscious from automatic contributions to the picture superiority effect by using a variation of Jacoby's (1991) Process Dissociation Procedure. A crucial assumption of the study was that the distinction between conscious and automatic processing coincides with that between conceptual and perceptual processing. Studied pictures were pitted against studied words in a picture fragment identification task (a perceptual, implicit test), a word fragment completion task (perceptual, implicit), and a category exemplar production task (conceptual, implicit). The purpose was to arrive at estimates of conscious and automatic memory for pictures and words separately. In an extension of the data analysis, multinomial models were fitted to the data. With these estimates in hand, the authors could indirectly infer the contributions from conceptual and perceptual processes to the picture superiority effect. The results showed superiority for pictures in conscious, hence conceptual, memory in all three tasks. Automatic, hence perceptual, memory varied expectedly with the degree of correspondence between encoding format and test format, being better for pictures in the picture fragment task, and better for words in the word fragment task. In the conceptual implicit memory task, there was an advantage for pictures in both the conscious and the automatic component.

A central inference in McBride and Doshier's (2002) study relies on the general correspondence between conscious and conceptual processes, on the one hand, and between automatic and perceptual processes, on the other. As the authors point out, this correspondence is far from perfect, and conceptual effects on some implicit tasks provide counterexamples (Toth & Reingold, 1996). Thus, some processes in memory are both conceptual

and implicit. Indeed, there are also other processes that are both perceptual and explicit, such as when a person is deliberately taking perceptual details into account when making a source judgement.

Leaving aside this particular interpretation, the McBride and Doshier study shows that computational modelling can separate factors in the picture superiority phenomenon that are entwined in the directly observable data. For present purposes, the hypothesised matching of perceptual and conceptual features is not directly observable, because both processes normally contribute to recognition. Modelling is one way of trying to separate those different but intertwined processes. The general class of Multinomial Processing Tree (MPT) models is applicable to this task, and it has proved to be useful in many applications in memory (Batchelder & Riefer, 1990; Riefer, Hu, & Batchelder, 1994) and in cognition generally (Batchelder & Riefer, 1999; Riefer & Batchelder, 1988). MPT models hold an intermediate place on a continuum of data analysis between, on the one hand, the general-purpose statistical models normally used for null hypothesis testing, and on the other hand, custom-built models, designed for a special field of application, such as the memory models SAM, MINERVA, and TODAM (Gillund & Shiffrin, 1984; Hintzman, 1984; Murdock, 1993). With the general-purpose models, MPT models share mathematical tractability and opportunities for relatively convenient hypothesis testing. With the special-purpose models, they share a capacity to be specific about cognitive processes, and a potential for making theoretical assumptions explicit. Although typically not as elaborate about underlying processes as the dedicated models, they still go some way towards specifying the cognitive processing that gives rise to the observable data. In this capacity, i.e., as a tool to discern latent variables, an MPT model was used in this article. As a complement, and for comparison, a memory model of the global matching kind (Clark & Gronlund, 1996) was applied to the same data. Any of the four dominant models (SAM, MINERVA, TODAM, and CHARM) could have served the purpose; MINERVA was chosen for reasons of simplicity and accessibility. “MINERVA 2 is perhaps the most impressive model, if only because it can do so much with so few assumptions and parameters” (Neath, 1998, p. 248). In this study, the models will be used solely as a data analysis tool; the purpose is not to make claims about the general value and applicability of MPT models and MINERVA, singly or in comparison.

TEST MANIPULATIONS

Implicit tasks have failed to provide a definitive answer to the question of what causes the picture superiority effect in memory. Furthermore, implicit

memory is of questionable relevance to the issue of effects in explicit memory. Therefore, the present study represents a return to an explicit task, i.e., recognition. The correspondence between study and test can be varied in a recognition task, for example by letting studied pictures be recognised as words, and studied words as pictures. This *crossover paradigm* has been tried in a few studies (Mintzer & Snodgrass, 1999; Stenberg, Radeborg, & Hedman, 1995), where the purpose has been to pry apart perceptual and conceptual influences by divorcing items at test from the format they had at study. Using these methods, both the cited studies found results that cast doubt on Paivio's *dual coding* account of picture superiority.

An even more desirable state of affairs would be to test both studied pictures and studied words on neutral ground, i.e., in a *third format* that shares no perceptual characteristics with any of the two. For bilinguals, the second language may offer such neutral ground. Testing studied concepts in participants' second language can give relatively unbiased estimates of the conceptual memory strength of pictures and first-language words. Studies of bilingual memory (reviewed by Francis, 1999) have indicated that first- and second-language words refer to a shared conceptual representation. In the present study, English-speaking Swedish students were given surprise memory tests in English after studying lists of Swedish words or pictures.

ENCODING MANIPULATIONS

A typical finding in the verbal memory literature is that deeper—i.e., more semantic—encoding leads to better retention. In picture memory this has not always been the case. The levels of processing effect seems to apply only in a limited sense to pictures as compared to words (D'Agostino, O'Neill, & Paivio, 1977).

Orienting tasks that direct attention to image qualities of a concept favour memory for words over pictures, and tasks that direct attention to verbal qualities of a concept favour memory for pictures over words (Durso & Johnson, 1980). This result has been replicated and located to the recollective component of recognition ("remember" as opposed to "know") by Dewhurst and Conway (1994). In other words, encoding that requires a mental transformation of the stimulus (naming a picture, or forming a mental image to a word) promotes better memory than passive encoding, apparently an application of the generation effect.

In the present study, encoding manipulations meant to enhance either conceptual or perceptual processing were used to verify that parameters in the model showed the desired kind of sensitivity. An effort was made to avoid confounds with the generation effect by using encoding instructions that could be applied with equal effort to pictures and words.

PURPOSE

The aim of this study was to assess the strength of conceptual and perceptual contributions to the picture superiority effect in explicit memory. Pictures and words were tested for recognition in both their original formats and translated into participants' second language. Models were fitted to the data, and parameters corresponding to perceptual and conceptual recognition were estimated. Over three experiments, orienting tasks were varied, so as to emphasise conceptual encoding (Exp. 2), perceptual encoding (Exp. 3), or neither (Exp. 1). The encoding manipulations were used to validate the parameters estimated from modelling. If parameters designed to measure conceptual recognition are enhanced by conceptual encoding, and if parameters designed to measure perceptual recognition are correspondingly enhanced by perceptual encoding, then more faith can be placed in the estimates.

A full (saturated) model was first fitted to the data. After that, restricted models were fitted, i.e., models in which certain parameters were forced to be equal to each other. In particular, we were interested in comparing the parameters (i.e., probabilities of recognition) for pictures with those for words. If probabilities of recognition for pictures and for words really were different, such a forced equality would put notable strain on the model, and the degree of misfit could be tested statistically. The main hypotheses were that (1) the probability of perceptually based recognition is higher for pictures than for words, (2) the probability of conceptually based recognition is also higher for pictures than for words, and (3) the difference between pictures and words is greater for conceptual recognition than for perceptual.

EXPERIMENT 1

The purpose of the first experiment was to examine memory performance for pictures and words, when encoding instructions were neutral, i.e., emphasising neither perceptual nor conceptual aspects.

Method

Participants. Participants were 19 students at Växjö University, who took part for a small monetary compensation. (Sex and age were not recorded in the data files of this experiment.) For the purposes of the study, it was crucial that the meanings of the English words be fully understood by the participants, although English was not their native language. The participants were all Swedish university students, who were judged to have an adequate command of English. Furthermore, the stimulus material covered mostly everyday objects, the names of which are easily understood.

Still, to avoid distortions of the results, subjects were always given the response alternative “I don’t understand the word”, whenever an English word was presented. This response option, which was used in 7%, 5%, and 7% of the trials in Experiments 1–3, respectively, automatically discarded the trial from further analysis.

Materials. The same stimulus material was used throughout all three experiments. It consisted of colour drawings similar to the Snodgrass and Vanderwart (1980) pictures, comprising 260 drawings of animals, tools, furniture, vehicles, etc. The picture set, which has been produced by Bruno Rossion and Gilles Pourtois (2004), is available on the Internet at <http://titan.cog.brown.edu:16080/~tarr/stimuli.html>. Swedish names were assigned to the pictures, and a subset of 200 items was selected, with a view to avoiding culture-specific items (such as an oval American football). Naming data for the original picture set are available in Snodgrass and Vanderwart (1980). Samples of the pictures are shown in Figure 1.

Because the Swedish and English names of the pictures were to be used as perceptually different labels for the same concept, an effort was made to avoid cognates, i.e., pairs of words in the two languages with a close phonetic and orthographic resemblance, for example *elefant* and *elephant*. However, given the fact that English and Swedish are related languages, it was not entirely possible to achieve that objective. In the selected set of 200 items, 22% can be judged to be cognates. The word pairs are listed in the Appendix, along with their judged cognate status. To eliminate the possibility that the results were produced by the cognate pairs, the results were recomputed with all cognate stimuli excluded. To anticipate, all important aspects of the results were retained in the subset of noncognate items.

The set of 200 items was divided into eight sets of 25 items each. An item in this context refers to a concept (such as “horse”), represented by three tokens: an English word, a picture, and a Swedish word. The eight sets were rotated in their assignments to use as target or as lure, or to being studied as picture or as word, so that each item was used in all assignments equally often across participants. Random order of presentation ensured that there were no systematic effects of primacy or recency. The experiments were programmed in E-prime (Psychology Software Tools, Inc.)

Procedure. The experimental session consisted of four study–test blocks, given in random order (see Figure 1). One block (PP) presented pictures during study, and the same pictures mixed with picture distractors during test. Another block (WW) presented Swedish words for study, and, likewise, Swedish words for test. Yet another block (PE) presented pictures during study, and the corresponding English words, mixed with other

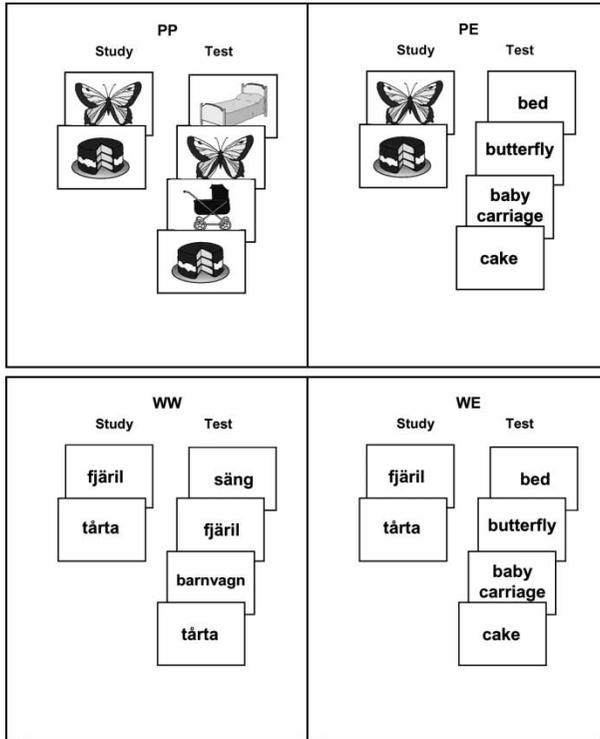


Figure 1. Conditions: PP: picture–picture, WW: (Swedish) word–(Swedish) word, PE: picture–English word, WE: (Swedish) word–English word. Samples of the stimulus material with pictures from the material of Rossion and Pourtois (2004), set 3 (colour pictures).

English distractors, during test. The fourth block (WE) presented Swedish words for study and English words in the test.

There was a 30 s pause between blocks. In each block, 25 items were presented for 2 s each. After an instruction screen, the test followed immediately, using 50 items, each presented until there was a response, or until 5 s had elapsed. Participants were to respond by pressing the “1” key for “old” or “2” for “new” (or, in the case of English words, “3” for “don’t understand”, see below). Feedback about the correctness of the response was given after each item. Instructions for the study subblocks were varied across experiments. In this experiment, instructions were just to watch the stimuli and try to remember as much as possible. Instructions for the PE and WE blocks explained that English words would be shown and that they were to be recognised as old if they corresponded to a previously shown picture (word). Examples were given to clarify the instruction. Participants were

tested in small groups (2–8), each seated individually in a booth with a computer.

Data analysis. Effect sizes for within-subjects tests are reported as partial eta squared, η_p^2 , which can be interpreted as the proportion of variance accounted for by the effect: $\eta_p^2 = SS_{\text{effect}} / (SS_{\text{effect}} + SS_{\text{error}})$.

Results and discussion

Memory performance is shown in Table 1. Sensitivity was computed as d' (Snodgrass & Corwin, 1988) for each of the four conditions and entered into a two-way repeated measures ANOVA, with study–test congruence as one factor (two levels: original and translated), and study format as the other (two levels: picture and word).

The results showed a powerful effect of study–test congruence, $F(1, 18) = 99.66$, $MSE = 0.25$, $p < .001$, $\eta_p^2 = .85$, and a strong effect of study format, $F(1, 18) = 41.47$, $MSE = 0.57$, $p < .001$, $\eta_p^2 = .70$, indicating better retention of items studied as pictures. This picture superiority effect was not modified by any interaction with study–test congruence, $F(1, 18) < 1$, *ns*. Still, for clarity, two separate t -tests were performed contrasting picture-studied and word-studied items. The blocks with original-form testing (PP and WW) showed picture superiority, $t(18) = 6.25$, $SEM = 0.19$, $p < .001$, and so did the blocks with translated testing (PE and WE), $t(18) = 4.08$, $SEM = 0.26$, $p = .001$.

To assess whether cognates contributed to the pattern of effects, all tests were recomputed with the cognate stimuli excluded. Both the effect of congruence, $F(1, 18) = 70.02$, $MSE = 0.35$, $p < .001$, $\eta_p^2 = .80$, and of study format, $F(1, 18) = 33.20$, $MSE = 0.54$, $p < .001$, $\eta_p^2 = .65$, remained highly significant. There was still no interaction, $F(1, 18) < 1$. The blocks with original-form testing showed picture superiority, $t(18) = 5.77$, $SEM = 0.18$,

TABLE 1
Experiment 1: Memory performance after a neutral orienting task

	<i>Original</i>		<i>Translated</i>	
	<i>Picture (PP)</i>	<i>Word (WW)</i>	<i>Picture (PE)</i>	<i>Word(WE)</i>
Hit rate	.96	.82	.83	.64
False alarm rate	.05	.14	.16	.25
d'	3.33	2.15	2.13	1.08
Effect size (η_p^2) of picture superiority	.68		.48	

PP: picture–picture; WW: Swedish word–Swedish word; PE: picture–English word; WE: Swedish word–English word.

$p < .001$, as did the blocks with translated testing, $t(18) = 3.58$, $SEM = 0.25$, $p = .002$.

The effect sizes (see Table 1) for the picture superiority effect were quite substantial for both the original ($\eta_p^2 = .68$) and translated ($\eta_p^2 = .48$) tests. Contrary to what could be expected if perceptual distinctiveness were the sole carrier of the picture superiority effect, the effect survived transformation into a new format having little in common with the original study format. Instead, and in conformity with the hypothesis of a semantic-access effect, studied pictures seemed to carry with them a richer memory representation, capable of being matched successfully against memory cues in a new format. The degree of semantic encoding was manipulated in Experiment 2, to examine whether picture superiority withstands the improvement in word encoding.

EXPERIMENT 2

The aim of Experiment 2 was to enhance semantic processing of the study items by giving a semantic orienting task, because earlier studies have shown smaller picture–word differences, when encoding is deep (D’Agostino et al., 1977).

Method

Participants. Twenty students at Växjö University participated and were given a small monetary compensation. Ten were male (age 24.6 ± 4.0), and ten were female (age 19.9 ± 7.6).

Procedure. Instructions in the study phase were to judge whether the item shown was living (plants, animals, body parts, etc.) or dead (tools, furniture, etc.). Participants responded by pressing the “1” or “2” keys. Unbeknownst to them, these responses were not recorded. Instructions were otherwise the same as in Experiment 1, as were all other aspects of the procedure.

Results and discussion

The results, which are presented in Table 2, were subjected to the same type of analysis as those of Experiment 1.

The results again showed a clear effect of study–test congruence, $F(1, 19) = 17.94$, $MSE = 0.65$, $p < .001$, $\eta_p^2 = .49$, and an equally clear effect of study format, $F(1, 19) = 25.88$, $MSE = 0.42$, $p < .001$, $\eta_p^2 = .58$, with better memory for items studied as pictures. The effect was not qualified by any interaction with study–test congruence, $F(1, 19) = 1.68$, *ns*. In the follow-up

TABLE 2
 Experiment 2: Memory performance after a semantic orienting task (abbreviations as in Table 1)

	<i>Original</i>		<i>Translated</i>	
	<i>Picture (PP)</i>	<i>Word (WW)</i>	<i>Picture (PE)</i>	<i>Word (WE)</i>
Hit rate	.96	.86	.84	.73
False alarm rate	.05	.12	.10	.13
<i>d'</i>	3.30	2.39	2.37	1.80
Effect size (η_p^2) of picture superiority	.49		.35	

t-tests, blocks with original-form testing (PP and WW) showed picture superiority, $t(19) = 4.24$, $SEM = 0.19$, $p < .001$, as did blocks with translated testing (PE and WE), $t(19) = 3.21$, $SEM = 0.21$, $p = .005$.

Recalculation with cognate stimuli excluded showed the same pattern of effects: Congruence, $F(1, 19) = 19.11$, $MSE = 0.68$, $p < .001$, $\eta_p^2 = .50$, and study format, $F(1, 19) = 26.46$, $MSE = 0.39$, $p < .001$, $\eta_p^2 = .58$, had strong effects, and there was no interaction, $F(1, 19) = 1.03$, *ns*. There was picture superiority in both original-format, $t(19) = 4.23$, $SEM = 0.20$, $p < .001$, and in translation, $t(19) = 3.23$, $SEM = 0.18$, $p = .004$.

As in the first experiment, picture superiority prevailed, in both original and translated forms. It did so even when the orienting task explicitly directed attention to the semantic properties of the stimuli. Although still considerable, the effect sizes were smaller than in Experiment 1. This seemed to depend not so much on what happened to the encoding of pictures, but rather on the sensitivity of *word* encoding to the orienting task. With a semantic task, word performance improved enough to narrow the gap somewhat.

To complete the manipulation of levels of processing, a further experiment (Experiment 3) was conducted, in which the orienting task instead promoted shallow, perceptual processing. This was expected to widen the gap between picture and word recognition.

EXPERIMENT 3

Shallow processing of words is often induced by directing attention to their orthographic or phonemic properties. However, in the present study, it was necessary to produce a task that could be applied to both pictures and words, in the same way and with roughly the same difficulty. Therefore, a new task was constructed that involved detecting changes in the spatial orientation of a minor part of the stimulus. This was meant to focus attention on the perceptual features of the stimuli.

Method

Participants. Nineteen students at Växjö University participated and were given a small monetary compensation. Eight were male (age 22.5 ± 2.3), and eleven were female (age 21.9 ± 2.6).

Procedure. The general method was applied as before, with the following exception. After the exposure of each stimulus in the study phase, the screen was erased, and a narrow horizontal strip of the stimulus was redisplayed, either mirror-reversed or not (see Figure 2). The task was to determine if the strip was reversed. The purpose of this task was to direct attention to the perceptual qualities of the stimulus during its initial, intact presentation (2 s, as in the other experiments).

Technically, a horizontal segment of the display, extending across its whole width (640 pixels), was saved to an off-screen memory buffer during the initial presentation, and copied back to the otherwise cleared screen for a 2 s poststimulus display, during which participants responded. The horizontal strip was centred on mid-screen; 50 pixels high for picture stimuli, and 10 pixels high for words (proportions that were found to make difficulty levels approximately equal). Strip height was chosen such as to make judgements about orientation very difficult without recourse to the memory image of the stimulus.

Results and discussion

The results, which are presented in Table 3, were submitted to the same type of analysis as the earlier experiments. The results again showed a very strong effect of study-test congruence, $F(1, 18) = 126.99$, $MSE = 0.34$, $p < .001$,

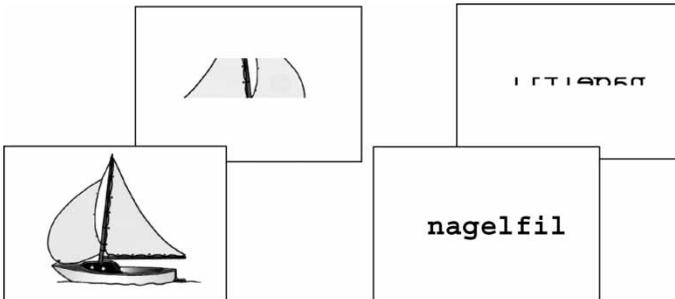


Figure 2. Examples of stimulus presentation in the study phase of Experiment 3. After the initial display (2 s) of a picture or a word, the screen was blanked (0.3 s), and then a narrow horizontal segment was redisplayed (2 s). The task was to determine whether the segment was mirror-reversed or not.

TABLE 3
 Experiment 3: Memory performance after a perceptual orienting task (abbreviations as in Table 1)

	<i>Original</i>		<i>Translated</i>	
	<i>Picture (PP)</i>	<i>Word (WW)</i>	<i>Picture (PE)</i>	<i>Word (WE)</i>
Hit rate	.96	.86	.82	.58
False alarm rate	.05	.21	.26	.37
<i>d'</i>	3.31	1.89	1.61	0.59
Effect size (η_p^2) of picture superiority	.84		.66	

$\eta_p^2 = .88$, and a similarly powerful effect of study format, $F(1, 18) = 95.04$, $MSE = 0.30$, $p < .001$, $\eta_p^2 = .84$, with better memory for items studied as pictures. There was a marginal tendency toward an interaction between the factors, $F(1, 18) = 3.81$, $MSE = 0.20$, $p = .07$. In the follow-up *t*-tests, blocks with original-form testing (PP and WW) showed picture superiority, $t(18) = 9.69$, $SEM = 0.15$, $p < .001$, as did blocks with translated testing (PE and WE), $t(19) = 5.87$, $SEM = 0.17$, $p < .001$. The effect sizes were larger than in the other experiments, mostly due to the poor performance for words. Shallow processing seemed to impair word retention disproportionately, widening the gap between pictures and words.

The tests were recomputed with all cognate stimuli excluded. The pattern of effects was unchanged. Congruence had a strong effect, $F(1, 18) = 116.06$, $MSE = 0.39$, $p < .001$, $\eta_p^2 = .87$, as did study format, $F(1, 18) = 72.25$, $MSE = 0.36$, $p < .001$, $\eta_p^2 = .80$. There was a tendency toward an interaction between the factors, $F(1, 18) = 3.42$, $MSE = 0.26$, $p = .08$. There was picture superiority in both original format, $t(18) = 9.13$, $SEM = 0.15$, $p < .001$, and in translated testing, $t(19) = 4.62$, $SEM = 0.21$, $p < .001$.

The encoding task seemed to be successful in directing attention to the perceptual features of the stimuli, because recognition performance deteriorated in the translated conditions, as would be expected if less attention could be devoted to semantic processing. It should be emphasised that the orienting task could not be performed without close attention to the perceptual qualities of the initial stimulus, because the redisplayed horizontal fragment was too narrow to be identified as mirror-reversed on its own.

In a comparison across experiments, the recognition of pictures seemed relatively resistant to changes in orienting task. Figure 3 shows memory performance, in *d'*-units, as a function of study–test condition (PP, WW, PE, or WE) and orienting task (perceptual, neutral, or semantic). In the PP condition, there was no increase (–1% change) from the neutral to the semantic encoding instructions, and no decrease (0%) from the neutral to the

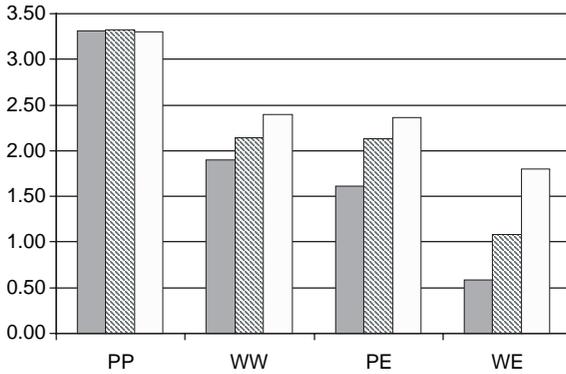


Figure 3. Recognition accuracy (d') across three experiments. Dark bars: Experiment 3 (perceptual orienting task). Hatched bars: Experiment 1 (neutral orienting task). Light bars: Experiment 2 (semantic orienting task). Abbreviations PP, WW, PE, and WE as in Table 1.

perceptual task. In contrast, the WW condition showed that the encoding of words was enhanced by semantic instructions (12%) and impaired by perceptual instructions (−12%), as could be expected from the vast literature on levels of processing.

Similarly, in the translated conditions, the retention of words proved to be greatly improved by the semantic task (+68% change in the WE condition), but the retention of pictures only modestly so (+11% in the PE condition). With a perceptual orienting task, performance for words declined drastically (−45%), whereas performance for pictures also declined markedly, but not nearly as much (−24%). The fact that items studied as pictures were encoded efficiently was manifest in both the original and the translated conditions. Picture recognition in the PP condition showed insensitivity to the orienting task that could be a consequence of performance having reached ceiling. In the PE condition, where there was no ceiling effect, picture encoding was predictably affected by instruction conditions, although less dramatically so than words.

A MULTINOMIAL MODEL

For present purposes, a model was constructed in which perceptually based recognition and conceptually based recognition were construed as separate processes. The model is illustrated in Figure 4 for the PP condition. An old (studied) item can be recognised by virtue of perceptual form with probability fp . Failing that, it can instead be recognised on the basis of its meaning, with probability sp . Failing that also, it can be called “old” based on guessing. The bias parameter, bp , determines how likely this is to happen.

Condition PP

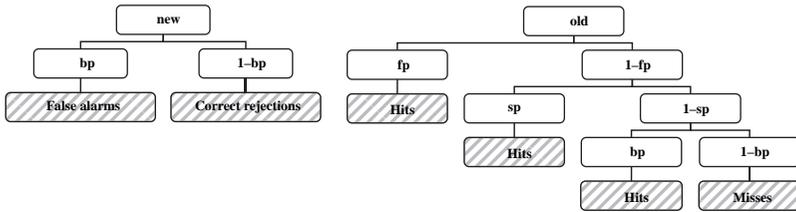


Figure 4. Multinomial Processing Tree models for the PP (picture–picture) condition. At the root are study conditions (studied/unstudied). From there on, branches marked by probability parameters (e.g., *sp*) lead to the leaves, i.e., the response categories, e.g., false alarms. Analogous trees were constructed for the WW condition.

All three of the described sequences lead to a *hit*, i.e., a studied item being called “old”. In contrast, an old item can be called “new” (a *miss*) if all three selection processes come out negative, and this happens with probability $(1 - fp) * (1 - sp) * (1 - bp)$.

For a new item, the process is simpler. The item can be called “old” only as a result of guessing, which leads to an incorrect response (a *false alarm*). Alternatively, it can be called “new” with probability $1 - bp$, a *correct rejection*. This high-threshold assumption (i.e., all false alarms being a result of guessing, never of, e.g., familiarity) is a typical part of MPT models of memory.

The model has been described here for the original-pictures condition, PP, but the case is exactly analogous for the WW condition, with the exception that the parameters for that case are called *fw*, *sw*, and *bw*, instead of *fp*, *sp*, and *bp*.

For the translated conditions, the recognition process is somewhat different (see Figure 5, where the PE condition is illustrated). An old item, i.e., an English word designating a studied picture, can be recognised based on meaning with probability *sp*. Form-based recognition is not possible in this condition. If semantic recognition fails, a correct response can be given as a guess, with probability *bpe*. (Note that the bias parameter in this case, *bpe*, is free to vary independently of *bp*, for the willingness to give an “old” response may very well be different in the more difficult PE condition from that in PP.) The *hit* probability is, in other words, $sp + (1 - sp) * bpe$. An incorrect response to an old item, a *miss*, can be given with probability $(1 - sp) * (1 - bpe)$.

For new items, an “old” response to a new item (a *false alarm*) is delivered with probability *bpe*. A *correct rejection* occurs with probability $(1 - bpe)$. The translated-word condition, WE, can be described in analogous terms, substituting *fw*, *sw*, and *bwe* for *fp*, *sp*, and *bpe*.

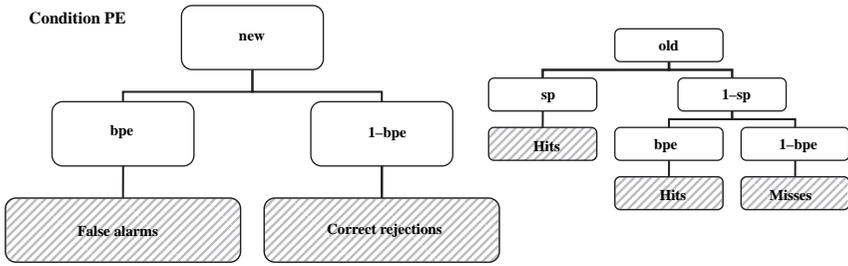


Figure 5. Multinomial Processing Tree models for the PE (picture-English word) condition. Branches leading from root to leaves are marked by transition probabilities—parameters of the models—as in Figure 4. Analogous trees were constructed for the WE condition.

The processing tree model was specified and tested in the program GPT (Hu & Phillips, 1999), available on the Internet at <http://xhuoffice.psc.memphis.edu/gpt/>. With all eight parameters free to vary, the model was fitted to the data, and the resulting parameter values are shown in Table 4. Because there are as many free parameters as there are independently observed categories, the model is saturated, and the fit was perfect, $G^2 = 0$.

For ease of inspection, the four most important parameters (fp , sp , fw , and sw) have been illustrated in Figure 6, where the experimental conditions have been arranged in an increasing order of processing depth, from Experiment 3 (perceptual), over Experiment 1 (neutral), to Experiment 2 (semantic).

The impression conveyed by Figure 6 is that the experimental manipulations of orienting task were effective. Semantic recognition increased from the perceptual to the semantic task, slightly for pictures (sp), and dramatically for words (sw). Conversely, form-based recognition decreased from the perceptual to the semantic task, somewhat for pictures (fp), and

TABLE 4
Fitted parameters of a full, and two restricted, Multinomial Processing Tree models across the three experiments

	<i>Full model</i>			<i>Restricted model $fp = fw$</i>			<i>Restricted model $sp = sw$</i>		
	<i>Exp. 1</i>	<i>Exp. 2</i>	<i>Exp. 3</i>	<i>Exp. 1</i>	<i>Exp. 2</i>	<i>Exp. 3</i>	<i>Exp. 1</i>	<i>Exp. 2</i>	<i>Exp. 3</i>
fp	0.791	0.743	0.809	0.628	0.559	0.750	0.872	0.812	0.893
fw	0.559	0.457	0.726				0.361	0.304	0.585
sp	0.798	0.828	0.755	0.824	0.851	0.769	0.671	0.765	0.565
sw	0.523	0.699	0.343	0.498	0.679	0.327			
bp	0.051	0.050	0.053	0.052	0.051	0.053	0.051	0.050	0.053
bpe	0.161	0.100	0.260	0.157	0.098	0.257	0.186	0.108	0.314
bw	0.139	0.120	0.206	0.136	0.117	0.204	0.139	0.120	0.206
bwe	0.245	0.128	0.368	0.251	0.131	0.372	0.215	0.120	0.314

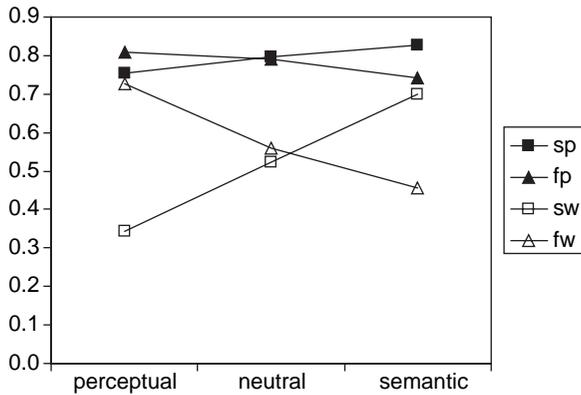


Figure 6. The four recognition parameters of the multinomial models, referring to form-based recognition of pictures (*fp*) and words (*fw*), and semantically based recognition of pictures (*sp*) and words (*sw*). The conditions have been ordered by increasing depth of processing: perceptual task (Exp. 3), neutral task (Exp. 1), and semantic task (Exp. 2).

drastically for words (*fw*). Thus, the figure contains two X-shaped patterns, one for pictures, and one for words. The X-shape for pictures is located above that for words, suggesting that probabilities of recognition for studied pictures were higher than those for studied words. This visual impression will now be tested more formally. Hypotheses about the model can be tested by the construction of submodels, where two parameters are constrained to be equal to one another. The logic is that if the restricted model shows a tolerable fit, the affected parameters were not reliably different from each other. If, on the other hand, the restricted model shows such strain as to be statistically rejected, we can conclude that the parameters were reliably different. By this reasoning, we will restrict the model in different ways in order to test two hypotheses concerning picture superiority.

Picture superiority

The picture superiority phenomenon has two aspects, which we are now in a position to tease apart. One is form-based recognition, which concerns the parameters *fp* and *fw*; the other is meaning-based recognition, which concerns the parameters *sp* and *sw*. In setting a criterion for accepting or rejecting the model, I have chosen $p = .01$ as the alpha level, because of the large number of observations ($N = 3800$). The critical $\chi^2(1)$ is therefore 6.63.

When a restricted model was tested, setting $fp = fw$, the model was rejected in Experiment 1, $G^2(1) = 7.97$, $p = .005$; again rejected in Experiment 2, $G^2(1) = 7.91$, $p = .005$; and accepted in Experiment 3, $G^2(1) = 1.86$, $p = .17$. Thus, perceptually based recognition was significantly better for pictures

than for words in the two first experiments. In the third experiment, where perceptual processing was encouraged, form-based recognition of words was raised almost to the level of pictures, resulting in no significant difference. Effect sizes were computed as w (Cohen, 1987) for χ^2 -tests, and judged by the convention that $w = .1$ is a small effect, $w = .3$ is medium-sized, and $w = .5$ is large. The size of the picture superiority effect in perceptually based recognition was small: .05 (Exp. 1), .05 (Exp. 2), and .02 (Exp. 3).

When a restricted model was constructed by setting $sp = sw$, we could test the hypothesis concerning differences between pictures and words in conceptually based recognition. The restricted model was rejected in all three experiments: $G^2(1) = 50.09$, $p < 10^{-11}$, in Experiment 1; $G^2(1) = 18.56$, $p < 10^{-4}$, in Experiment 2; and $G^2(1) = 71.55$, $p < 10^{-16}$, in Experiment 3. We can therefore conclude that sp is reliably higher than sw . Effect sizes were small by the w standard: .12 (Experiment 1), .07 (Experiment 2), and .15 (Experiment 3), yet in all cases larger than the corresponding effects for perceptually based recognition. The ratio of effect sizes (conceptual/perceptual) was 2.4, 1.5, and 6.3 in the three experiments, respectively, in each case a sizeable difference favouring the conceptual basis for the picture superiority effect over the perceptual basis.

A MINERVA-2 MODEL

Our focus is on the intuition that decisive differences between pictures and words lie in the ease and efficiency with which perceptual and, especially, semantic features are encoded during the study episode. For the present purposes, it is appropriate to use a model that articulates the mechanism of encoding and provides parameters by which it can be manipulated. Minerva 2 (Hintzman, 1984) is such a model.

In Minerva 2, stimuli are vectors of features, which are encoded into memory probabilistically. A parameter, the learning rate (L), specifies the probability with which a feature from the stimulus (held in a temporary primary-memory buffer) will be transferred to long-term memory. The higher the probability is, the more faithful the memory trace will be, and the higher the future chance of correct recognition.

Recognition is thought of as a matching of a test probe (again, a vector of features) against all traces in memory simultaneously. For each trace, a similarity measure with the probe is computed. If the probe has been studied, it will tend to evoke a strong response from at least one trace in memory, namely its more or less faithfully encoded replica. The similarity measures are summed over all traces, after first being raised to the third power, which has the effect of enhancing the contrast between strong similarities and weak ones. The resulting sum is called the echo intensity. The

echo can be thought of as an equivalent to the familiarity dimension that forms the basis of memory decisions in signal detection models of memory. The (simulated) participant calls the test probe old if the echo intensity exceeds a certain criterion value, which is also a parameter of the model.

Details of the implementation were as follows. Each stimulus was simulated as a vector of 40 features, half of which were semantic and half perceptual. Each feature could take on the values +1, -1, or 0, where the zero value means that the feature is irrelevant or unknown. When the stimulus material was assembled, which was done anew for each simulated condition and subject, values were assigned to the features independently, and with equal probabilities for the three possible values. In the simulated study blocks, features were encoded into memory with probabilities determined by four learning-rate parameters. These were the core of the model, and the purpose of the modelling exercise was to study whether variations in these four parameters alone could mimic the experimental results. Encoding of semantic features in pictures took place with probability L_{sp} , and encoding of perceptual features in pictures with probability L_{fp} . The corresponding learning rates for words were L_{sw} and L_{fw} . (The notation is analogous to that of the multinomial model.) The purpose of the present modelling was to see whether the learning rate parameters in this more elaborate memory model could carry out the same functions as the recognition probabilities in the multinomial model.

The picture–picture condition (PP) was simulated by encoding the whole 40-feature vectors; the first 20 features with learning rate L_{sp} , and the last 20 with learning rate L_{fp} . In the recognition phase, the whole 40-feature vectors were matched against the memory traces. The word–word condition (WW) used the same procedure, substituting the learning rates L_{sw} and L_{fw} .

In the two translated conditions (PE and WE), only the semantic features were meaningfully encoded; the perceptual features were set to 0. Consequently, only the semantic features were used in recognition.

For simplicity, the criterion in each condition (i.e., the cut-off limit used by the subject) was fixed so as to simulate bias-free responding, i.e., the criterion was set at the median value of the joint old + new distribution of echo intensities. This was a fairly good approximation to the actual behaviour of the experimental participants. It is to be noted that no bias parameters were fitted, nor were high-threshold assumptions made about the mechanism behind false alarms.

A number of simulations of the experiments were run with 300 simulated participants, each receiving four conditions with 50 test stimuli in each. The resulting hit rates and false alarm rates were compared with the actual ones, and the root mean squared (RMS) error was computed to guide the search for optimal parameters. The simplex algorithm (in function *fminsearch* of the

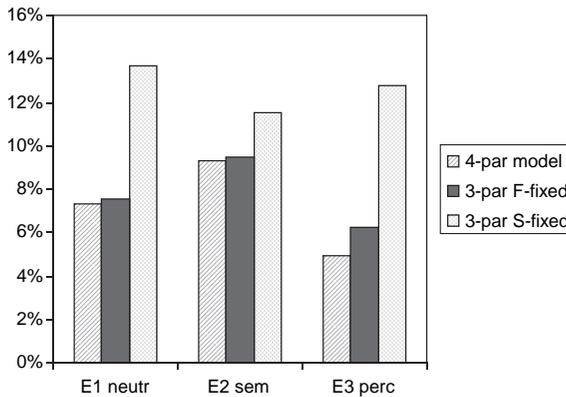


Figure 7. Fit of Minerva 2 models (root mean squared error as a percentage of mean expected value). The full four-parameter model shows the best fit. Nested models were also tested, where either form-based or semantically based recognition parameters were equated (pictures = words). The f-fixed model ($L_{fp} = L_{fw}$) increases misfit only slightly. The s-fixed model ($L_{sp} = L_{sw}$) fits significantly worse, demonstrating that L_{sp} and L_{sw} are reliably different.

Matlab program, Mathworks Inc.) was used to search the parameter space. The resulting best-fitting parameters are shown in Figure 7.

To determine the fit of the data statistically, Pearson's χ^2 was computed. Comparisons between observed and predicted frequencies were made for eight independent cells, i.e., hit rates and false alarm rates in the four conditions PP, WW, PE, and WE. Four parameters were free to vary, resulting in 4 degrees of freedom ($df = 8 - 4$). The critical $\chi^2(4) = 13.28$ at $p = .01$ ($N = 4000$). Frequencies were computed for a group size of 20 participants, for compatibility with the actual group size.¹

The model was accepted for Experiment 1 ($\chi^2 = 10.25$), rejected for Experiment 2 ($\chi^2 = 27.34$), and accepted for Experiment 3 ($\chi^2 = 7.54$). Thus, the fit of the model to the data was relatively good, resulting in acceptance in two cases out of three.

Our interest centres on the values of those four learning parameters that yielded the best fits. They are shown in Figure 8. The general pattern coincides with the recognition probabilities of the multinomial model (cf. Figure 6). First, it can be noted that the effect of orienting task is in agreement with expectations, i.e., the semantic learning rates increase with semantic tasks, and the perceptual learning rates increase with perceptual tasks. Second, it can be noted that learning rates for pictures are higher than

¹ However, the choice of group size in simulation studies contains an element of arbitrariness. Hence, the result of Pearson's χ^2 , which is size dependent, needs to be interpreted with caution.

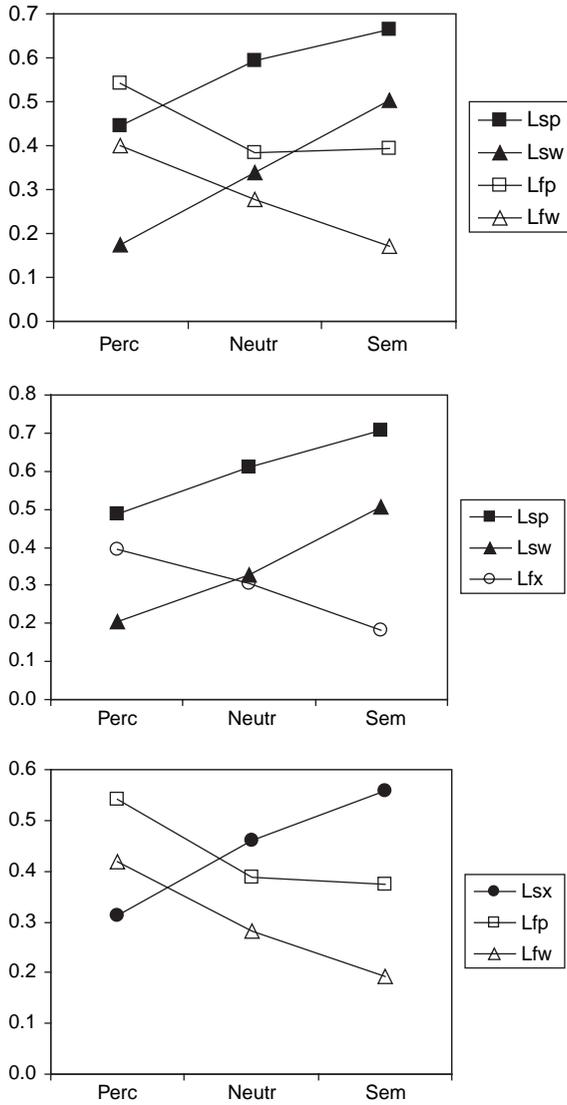


Figure 8. Best-fitting parameters from the Minerva 2 simulations. The four recognition parameters, referring to form-based recognition of pictures (*Lfp*) and words (*Lfw*), and semantically based recognition of pictures (*Lsp*) and words (*Lsw*). The conditions have been ordered by increasing depth of processing: the perceptual task (Exp. 3), neutral task (Exp. 1), and semantic task (Exp. 2). Top panel shows full model, middle panel the *f*-fixed model, and the lower panel the *s*-fixed model.

those for words, and this picture superiority effect seems to apply to both semantic features and perceptual ones.

The qualitative pattern of effects is in general agreement with those seen when analysing the multinomial model, as inspection of Figures 6 and 8 shows. The learning rates for semantic features of pictures are higher than those for words. A similar difference, although not as large, can be seen for the learning rates of perceptual features for pictures versus words. A more formal test of this impression was also conducted.

The likelihood of the data, given the set of parameters, was computed, using probability density functions derived by bootstrapping from a simulation of 300 participants. In a strategy similar to the treatment of the MPT models, the full (four-parameter) model was compared to restricted (three-parameter) models. First, perceptually based memory was set equal for pictures and words ($Lfp = Lfw$), and then conceptually based memory was treated similarly ($Lsp = Lsw$). The fits of the restricted models were compared with that of the full model, using the equation: $G^2 = -2 * (LL1 - LL2)$, where LL1 and LL2 are the log likelihoods of the restricted and the full model, respectively. G^2 is distributed as χ^2 with one degree of freedom. As Figure 7 suggests, the fit of the f-fixed model ($Lfp = Lfw$) was almost as good as that of the full model, resulting in nonsignificant tests: $G^2 = 0.73$, $p = .39$ in Experiment 1; $G^2 = 1.17$, $p = .28$ in Experiment 2; and $G^2 = 1.12$, $p = .29$ in Experiment 3.

The s-fixed model ($Lsp = Lsw$), on the other hand, fared notably worse than the full model. The tests showed reliable measures of misfit: $G^2 = 10.03$, $p = .002$ in Experiment 1, $G^2 = 5.11$, $p = .024$ in Experiment 2, and $G^2 = 9.00$, $p = .003$ in Experiment 3. Therefore, we can conclude that Lsp and Lsw are reliably different, or, in other words, that conceptual memory is reliably better for pictures than for words.

GENERAL DISCUSSION

This study examined the memorial superiority of pictures over words, using a method where pictures and (Swedish) words were studied, and recognition tests took place in a third format—English words—as well as in the original formats. The type of processing at encoding was varied over levels of processing: semantic, neutral, or perceptual. Picture superiority prevailed in all orienting task conditions, and when tested in translated as well as in original formats. The main purpose was to examine the part played by perceptual and conceptual factors in the picture superiority effect. Models were fitted to the data, and quantitative estimates showed that both perceptual and conceptual features contributed to the picture superiority effect, but the main contribution came from conceptual processing. This is in

some contrast to the view that perceptual distinctiveness is the major advantage accruing to pictures over words (Weldon & Coyote, 1996).

Some features of this study are new. A third format, in which concepts studied as pictures and words are to be recognised, has not been used before, to my knowledge. It serves to place both formats on an equal footing by depriving the test format of perceptual similarities with the studied format. In this study, participants' second language—English—served as the third format for Swedish-speaking students. Because of the similarities between the two languages, a certain proportion of the stimulus words showed orthographic similarities between (studied) Swedish and (tested) English words. However, elimination of these words from the computation of results did not alter the conclusions; in fact, all statistical inferences were left unaltered despite the reduction of the database.

Some of the assumptions about the neutral third format may be called into question. First, drawing a line between cognates and noncognates is somewhat arbitrary; inevitably, there are some partial resemblances even within the noncognate pairs. Second, English words may have been automatically activated by their Swedish counterparts at study, because bilinguals have been shown to undergo automatic activation of both languages even in a strictly monolingual task (Bijeljac-Babic, Biardeau, & Grainger, 1997). Third, fully voluntary activation of the English words may also have happened, when participants, after exposure to the first English-language test, caught on to the idea that more of the same may follow. In the first two cases, the result would have been a potential inflation of performance in condition WE, and as a consequence, an overestimation of the semantic encoding of words, i.e., parameters sw and Lsw in the models. Direct links between Swedish and English words would therefore, if they had any effect at all on performance, seemingly enhance conceptual memory for words relative to conceptual memory for pictures. But the main finding of this study is just the opposite, and these potential sources of error would, if anything, lead to overly cautious conclusions. The third source of error, i.e., voluntary translation, would be likely to affect picture and word encoding to the same degree, without distorting the relation between the two. Thus, there is nothing in these potential confounds that would work to enhance the main finding; more likely, they would detract from it.

The encoding manipulations varied levels of processing: from semantic over neutral to perceptual. Care was taken to treat pictures and words equally. In some earlier studies, levels of processing have been confounded with a generation effect, such as when one orienting task has been naming, i.e., generating a name to a picture as contrasted with just reading a word out loud. Another example is image production, which contrasts creating a mental image to a word with just perceiving a presented picture. The more active tasks unsurprisingly lead to better retention. In the present study, the

orienting tasks could be applied in a symmetrical fashion to pictures and words. They produced expected level of processing effects on retention of both pictures and words, and thus served to validate the models. Model parameters for conceptual and perceptual processing were affected in consistent ways by the orienting tasks: semantic encoding increased the probability of conceptually based recognition and decreased the probability of perceptually based recognition. Perceptual encoding had the opposite effect. Thus, the encoding manipulations dissociated perceptual and conceptual processing, and the parameters of the model captured this variation.

Earlier theorising has claimed an advantage in perceptual memory for pictures over words, but this has been relatively difficult to document empirically. Perceptual implicit tests favour studied words, if the test is verbal, and they favour studied pictures, if the test is pictorial. Because implicit tests tend to be format specific, results have not been comparable across formats. The only determined effort at comparing picture and word priming using a common currency is the study by Kinjo and Snodgrass (2000), which used picture and word fragment identification tests, calibrated for equal difficulty. The results did not, however, make a wholly consistent case for better perceptual memory for pictures, because only one experiment of three showed the expected picture superiority in within-format priming. The present study used explicit memory and showed estimates of perceptually based recognition to be higher for pictures than for words in two experiments out of three.

The conceptual memory advantage for pictures over words has often been proposed, but seldom demonstrated. Explicit recognition tests show a picture advantage, but attributing this to conceptual memory needs additional supporting data. Support has been marshalled from conceptual implicit tests, but no consistent picture superiority has emerged. Negative results have been found (Weldon & Coyote, 1996), along with positive (Nicolas, 1995) and mixed results (Vaidya & Gabrieli, 2000; Wippich et al., 1998). Weldon originated, with her collaborators, the idea that conceptual processing lies at the root of picture superiority in explicit tests, but after failing to confirm the predicted advantage in conceptual implicit tests, Weldon and Coyote reneged and instead proposed visual distinctiveness as the dominant contributor. A tendency in these diverse results seems to be that shallow encoding conditions result in picture superiority in conceptual implicit tests, whereas words benefit more from deep encoding, which tends to eliminate the picture advantage. In the present study, conceptual memory showed an advantage for pictures over words in all three experiments. In line with previous results, semantic processing diminished, but did not abolish, the picture superiority effect. Overall, memory for words stood more to gain from a deeper level of processing. Memory for pictures was high, in relative independence of levels of processing. However, the independence was only

relative, because studied pictures showed regular levels of processing effects when tested in the PE condition.

This study shares a modelling approach with McBride and Doshier (2002). McBride and Doshier used Jacoby's Process Dissociation Procedure to arrive at estimates of conscious and automatic uses of memory in three implicit tasks: picture fragment identification, word-stem completion, and category exemplar production. Multinomial Processing Trees were used for the modelling. By inference, conscious memory was equated with conceptual, and automatic with perceptual memory. Conscious memory was found to be higher for pictures than for words, whereas automatic memory was higher for studied pictures in the picture identification task, and higher for words in the word stem completion task. The results are consistent with the Transfer Appropriate Processing approach (Weldon & Roediger, 1987; Weldon et al., 1989).

The present study aimed to model an explicit memory task and to separate conceptual and perceptual memory directly in the process. There was therefore no need to take the inferential step of equating conscious with conceptual memory. The results showed a picture advantage in both conceptual and perceptual (explicit) memory. Results were consistent across two types of models with different methods and assumptions: Multinomial Processing Trees (Batchelder & Riefer, 1999) and Minerva 2 (Hintzman, 1984, 1988).

The two types of models confer different advantages. The MPT models are computationally tractable and efficient, but some of its parameters and constructs lack a clear psychological interpretation. In particular, the many different bias parameters seem unparsimonious. The processing of new items is described by the high threshold assumption, according to which a new item can be misjudged as old only as a result of guessing, never as a result of e.g., preexperimental familiarity. In comparison, the Minerva model is richer in psychologically plausible content. For example, false alarms are produced by the same familiarity mechanism that gives rise to hits. Both types of responses fall out of the few basic assumptions, using only a few free parameters. This realism comes at a cost, for it is computationally demanding and lacks closed-form solutions to hypothesis testing.

In conclusion, the conceptual and perceptual processing advantages for pictures over words have been debated, following inconclusive evidence from implicit tests. The present results indicate that both conceptual and perceptual factors confer advantages to pictures over words in memory. They also suggest that the main contribution comes from conceptual processing.

Original manuscript received October 2004

Revised manuscript received May 2005

PrEview proof published online month/year

REFERENCES

- Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, *97*(4), 548–564.
- Batchelder, W. H., & Riefer, D. M. (1999). Theoretical and empirical review of multinomial process tree modeling. *Psychonomic Bulletin and Review*, *6*(1), 57–86.
- Bijeljac-Babic, R., Biardeau, A., & Grainger, J. (1997). Masked orthographic priming in bilingual word recognition. *Memory and Cognition*, *25*(4), 447–457.
- Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin and Review*, *3*(1), 37–60.
- Cohen, J. (1987). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- D'Agostino, P. R., O'Neill, B. J., & Paivio, A. (1977). Memory for pictures and words as a function of level of processing: Depth or dual coding? *Memory and Cognition*, *5*(2), 252–256.
- Dewhurst, S. A., & Conway, M. A. (1994). Pictures, images and recollective experience. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(5), 1088–1098.
- Durso, F. T., & Johnson, M. K. (1980). The effects of orienting tasks on recognition, recall, and modality confusion of pictures and words. *Journal of Verbal Learning and Verbal Behavior*, *19*(4), 416–429.
- Engelkamp, J., & Zimmer, H. D. (1994). *Human memory: A multimodal approach*. Seattle, CA: Hogrefe & Huber.
- Federmeier, K. D., & Kutas, M. (2001). Meaning and modality: Influences of context, semantic memory organization, and perceptual predictability on picture processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 202–224.
- Francis, W. S. (1999). Cognitive integration of language and memory in bilinguals: Semantic representation. *Psychological Bulletin*, *125*(2), 193–222.
- Gardiner, J. M., & Richardson-Klavehn, A. (2000). Remembering and knowing. In E. Tulving & F. Craik (Eds.), *The Oxford handbook of memory* (pp. 229–244). Oxford, UK: Oxford University Press.
- Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1–67.
- Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments and Computers*, *16*(2), 96–101.
- Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*, *95*(4), 528–551.
- Hu, X., & Phillips, G. A. (1999). GPT.EXE: A powerful tool for the visualization and analysis of general processing tree models. *Behavior Research Methods, Instruments and Computers*, *31*(2), 220–234.
- Jacoby, L. L. (1983). Remembering the data: Analyzing interactive processes in reading. *Journal of Verbal Learning and Verbal Behavior*, *22*(5), 485–508.
- Jacoby, L. L. (1991). A process dissociation framework—separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–541.
- Jacoby, L. L., & Dallas, M. (1981). On the relationship between autobiographical memory and perceptual-learning. *Journal of Experimental Psychology: General*, *110*(3), 306–340.
- Johnston, W. A., Dark, V. J., & Jacoby, L. L. (1985). Perceptual fluency and recognition judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *11*(1), 3–11.
- Kazmerski, V. A., & Friedman, D. (1997). Old/new differences in direct and indirect memory tests using pictures and words in within- and cross-form conditions: Event-related potential and behavioral measures. *Cognitive Brain Research*, *5*(4), 255–272.
- Kinjo, H., & Snodgrass, J. G. (2000). Is there a picture superiority effect in perceptual implicit tasks? *European Journal of Cognitive Psychology*, *12*(2), 145–164.

- Kirkpatrick, E. A. (1894). An experimental study of memory. *Psychological Review*, 1, 602–609.
- Köhler, S., Moscovitch, M., Winocur, G., & McIntosh, A. (2000). Episodic recognition of pictures and words: Role of the human medial temporal lobes. *Acta Psychologica*, 105, 159–179.
- Madigan, S. (1983). Picture memory. In J. C. Yuille (Ed.), *Imagery, memory, and cognition: Essays in honor of Allan Paivio* (pp. 65–89). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87(3), 252–271.
- McBride, D. M., & Doshier, B. A. (2002). A comparison of conscious and automatic memory processes for picture and word stimuli: A process dissociation analysis. *Consciousness and Cognition: An International Journal*, 11(3), 423–460.
- Mintzer, M. Z., & Snodgrass, J. G. (1999). The picture superiority effect: Support for the distinctiveness model. *American Journal of Psychology*, 112(1), 113–146.
- Murdock, B. B. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review*, 100(2), 183–203.
- Neath, I. (1998). *Human memory: An introduction to research, data and theory*. Pacific Grove, CA: Brooks/Cole.
- Nelson, D. L., & Brooks, D. H. (1973). Functional independence of pictures and their verbal memory codes. *Journal of Experimental Psychology*, 98(1), 44–48.
- Nelson, T. O., Metzler, J., & Reed, D. A. (1974). Role of details in the long-term recognition of pictures and verbal descriptions. *Journal of Experimental Psychology*, 102(1), 184–186.
- Nelson, D. L., Reed, V. S., & McEvoy, C. L. (1977). Learning to order pictures and words: A model of sensory and semantic encoding. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 485–497.
- Nicolas, S. (1995). The picture-superiority effect in category-association tests. *Psychological Research/Psychologische Forschung*, 58(3), 218–224.
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart, & Winston.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford, UK: Oxford University Press.
- Paivio, A. (1991). *Images in mind: The evolution of a theory*. New York: Harvester Wheatsheaf.
- Potter, M. C., & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature (London)*, 253, 437–438.
- Rajaram, S. (1993). Remembering and knowing: Two means of access to the personal past. *Memory and Cognition*, 21(1), 89–102.
- Riefer, D. M., & Batchelder, W. H. (1988). Multinomial modeling and the measurement of cognitive processes. *Psychological Review*, 95(3), 318–339.
- Riefer, D. M., Hu, X., & Batchelder, W. H. (1994). Response strategies in source monitoring. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(3), 680–693.
- Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, 33(2), 217–236.
- Schloerscheidt, A. M., & Rugg, M. D. (2004). The impact of change in stimulus format on the electrophysiological indices of recognition. *Neuropsychologia*, 42(4), 451–466.
- Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture–word processing. *Journal of Experimental Psychology: General*, 109, 373–392.
- Snodgrass, J. G., & Corwin, J. (1988). Pragmatics of measuring recognition memory: Applications to dementia and amnesia. *Journal of Experimental Psychology: General*, 117(1), 34–50.
- Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 174–215.
- Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2560 visual stimuli. *Psychonomic Science*, 19, 73–74.

- Stenberg, G., Radeborg, K., & Hedman, L. R. (1995). The picture superiority effect in a cross-modality recognition task. *Memory and Cognition*, 23(4), 425–441.
- Toth, J. P., & Reingold, E. M. (1996). Beyond perception: Conceptual contributions to unconscious influences of memory. In G. Underwood (Ed.), *Implicit cognition* (pp. 41–84). Oxford, UK: Oxford University Press.
- Vaidya, C. J., & Gabrieli, J. D. E. (2000). Picture superiority in conceptual memory: Dissociative effects of encoding and retrieval tasks. *Memory and Cognition*, 28(7), 1165–1172.
- Weldon, M. S., & Coyote, K. C. (1996). Failure to find the picture superiority effect in implicit conceptual memory tests. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 670–686.
- Weldon, M. S., & Roediger, H. L. (1987). Altering retrieval demands reverses the picture superiority effect. *Memory and Cognition*, 15(4), 269–280.
- Weldon, M. S., Roediger, H. L., & Challis, B. H. (1989). The properties of retrieval cues constrain the picture superiority effect. *Memory and Cognition*, 17(1), 95–105.
- Whittlesea, B. W. A. (1993). Illusions of familiarity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19(6), 1235–1253.
- Wippich, W., Melzer, A., & Mecklenbraeuer, S. (1998). Picture or word superiority effects in implicit memory: Levels of processing, attention and retrieval constraints. *Swiss Journal of Psychology/Schweizerische Zeitschrift fuer Psychologie/Revue Suisse de Psychologie*, 57(1), 33–46.
- Wiseman, S., & Neisser, U. (1974). Perceptual organization as a determinant of visual recognition memory. *American Journal of Psychology*, 87(4), 675–681.

APPENDIX

Stimuli in English and Swedish, with cognate/noncognate status

<i>English</i>	<i>Swedish</i>	<i>Cognate</i>
airplane	flygplan	noncog
alligator	krokodil	noncog
anchor	ankare	cognate
apple	äpple	cognate
arrow	pil	noncog
ashtray	askkopp	noncog
asparagus	sparris	noncog
axe	yxa	noncog
baby carriage	barnvagn	noncog
ball	boll	cognate
balloon	ballong	cognate
banana	banan	cognate
barn	lada	noncog
barrel	tunna	noncog
basket	korg	noncog
bear	björn	noncog
bed	säng	noncog
bee	bi	cognate
bell	kyrkklocka	noncog
belt	livrem	noncog
bicycle	cykel	noncog
bird	fågel	noncog
book	bok	cognate
boot	stövel	noncog
bottle	flaska	noncog
bowl	skål	noncog
box	låda	noncog
bread	bröd	cognate
broom	kvast	noncog
brush	borste	noncog
butterfly	fjäril	noncog
button	knapp	noncog
cake	tårta	noncog
candle	stearinljus	noncog
cannon	kanon	cognate
car	bil	noncog
carrot	morot	noncog
cat	katt	cognate
chain	kedja	noncog
chair	stol	noncog
cherry	körsbär	noncog
chisel	stämjärn	noncog
church	kyrka	noncog
cigar	cigarr	cognate
cigarette	cigarrett	cognate
clock	klocka	cognate

APPENDIX (Continued)

<i>English</i>	<i>Swedish</i>	<i>Cognate</i>
clothespin	klädnyppa	noncog
cloud	moln	noncog
coat	rock	noncog
comb	kam	cognate
couch	soffa	noncog
cow	ko	cognate
crown	krona	cognate
cup	kopp	cognate
deer	rådjur	noncog
desk	skrivbord	noncog
dog	hund	noncog
doll	docka	noncog
donkey	åсна	noncog
door	dörr	cognate
doorknob	dörrhandtag	noncog
dress	klänning	noncog
drum	trumma	noncog
duck	anka	noncog
eagle	örn	noncog
ear	öra	noncog
envelope	kuvert	noncog
eye	öga	noncog
fence	staket	noncog
fish	fisk	cognate
flag	flagga	cognate
flower	blomma	noncog
flute	flöjt	cognate
fly	fluga	noncog
fork	gaffel	noncog
fox	räv	noncog
frog	groda	noncog
frying pan	stekpanna	noncog
garbage can	soptunna	noncog
glove	handske	noncog
goat	get	noncog
grapes	vindruvor	noncog
grasshopper	gräshoppa	noncog
guitar	gitarr	cognate
gun	pistol	noncog
hair	hår	cognate
hammer	hammare	cognate
hanger	klädhängare	noncog
harp	harpa	cognate
hat	hatt	cognate
heart	hjärta	cognate
horse	häst	noncog
iron	strykjärn	noncog
ironing board	strykbräda	noncog

APPENDIX (Continued)

<i>English</i>	<i>Swedish</i>	<i>Cognate</i>
jacket	jacka	cognate
kangaroo	känguru	cognate
kettle	kittel	cognate
key	nyckel	noncog
knife	kniv	cognate
ladder	stege	noncog
leaf	löv	cognate
leg	ben	noncog
lemon	citron	noncog
lettuce	sallad	noncog
light bulb	glödlampa	noncog
light switch	strömbrytare	noncog
lobster	hummer	noncog
lock	lås	noncog
monkey	apa	noncog
moon	måne	cognate
mountain	berg	noncog
mouse	mus	cognate
mushroom	svamp	noncog
nail	spik	noncog
nail file	nagelfil	noncog
necklace	halsband	noncog
needle	nål	noncog
nose	näsa	cognate
onion	lök	noncog
orange	apelsin	noncog
ostrich	struts	noncog
owl	uggla	noncog
paintbrush	pensel	noncog
pants	byxor	noncog
peach	persika	noncog
peacock	påfågel	noncog
peanut	jordnöt	noncog
pear	päron	cognate
pen	kulspetspenna	noncog
pencil	blyertspenna	noncog
penguin	pingvin	cognate
pig	gris	noncog
pineapple	ananas	noncog
pipe	pipa	cognate
pitcher	bringare	noncog
pliers	tång	noncog
plug	stickkontakt	noncog
pot	kastrull	noncog
rabbit	kanin	noncog
record player	grammofon	noncog
refrigerator	kylskåp	noncog
rhinoceros	noshörning	noncog

APPENDIX (Continued)

<i>English</i>	<i>Swedish</i>	<i>Cognate</i>
rocking chair	gungstol	noncog
roller skate	rullskridsko	noncog
rooster	tupp	noncog
ruler	linjal	noncog
sailboat	segelbåt	noncog
salt shaker	saltkar	cognate
sandwich	smörgås	noncog
saw	såg	noncog
scissors	sax	noncog
screw	skruv	noncog
screwdriver	skruvmejsel	noncog
seahorse	sjöhäst	noncog
seal	säl	cognate
sheep	får	noncog
shirt	skjorta	noncog
shoe	sko	cognate
skirt	kjol	noncog
sled	kälke	noncog
snail	snigel	noncog
snake	orm	noncog
snowman	snögubbe	noncog
sock	strumpa	noncog
spider	spindel	noncog
spinning wheel	spinnrock	noncog
spool of thread	trådulle	noncog
spoon	sked	noncog
squirrel	ekorre	noncog
star	stjärna	noncog
stool	pall	noncog
stove	spis	noncog
strawberry	jordgubbe	noncog
suitcase	resväska	noncog
sun	sol	noncog
sweater	tröja	noncog
swing	gunga	noncog
table	bord	noncog
tie	slips	noncog
toaster	brödrost	noncog
toothbrush	tandborste	noncog
traffic light	trafikljus	noncog
train	tåg	noncog
tree	träd	cognate
truck	lastbil	noncog
turtle	sköldpadda	noncog
umbrella	paraply	noncog
vase	vas	cognate
violin	fiol	noncog
wagon	kärra	noncog

APPENDIX (*Continued*)

<i>English</i>	<i>Swedish</i>	<i>Cognate</i>
watch	armbandsur	noncog
watering can	vattenkanna	noncog
watermelon	vattenmelon	cognate
well	brunn	noncog
wheel	hjul	noncog
whistle	visselpipa	noncog
windmill	väderkvarn	noncog
window	fönster	noncog
wineglass	vinglas	noncog
wrench	skiftnyckel	noncog