Psicothema 2007. Vol. 19, nº 1, pp. 163-170
www.psicothema.com
ISSN 0214 - 9915 CODEN PSOTEG

ROC parameters in item and context recognition

Salvador Algarabel and Alfonso Pitarque
Universidad de Valencia

Conflicting theories argue that recognition is achieved either by familiarity exclusively, or by a mixture of familiarity and recollection. We explore in three experiments the goodness of fit of both positions to experimental data in which context information is manipulated. In Experiments 1 and 2, we explore the availability of context information in recognition, testing the focus stimulus, its context, and their associative relation. In Experiment 3, participants were confronted with a plurality task in an attempt to force them to use the peripheral information in recognition. The results show that people acquire specific associative information, and although overall recognition performance was not affected by the use of context, receiver operating characteristic (ROC) analysis showed that people use a duality of processes in recognition.

Parámetros ROC en el reconocimiento de ítems y contextos**.** Las dos principales posiciones teóricas respecto al reconocimiento arguyen que en éste intervienen sólo la familiaridad, o la familiaridad y el recuerdo conjuntamente. En tres experimentos analizamos la bondad del ajuste de ambas posiciones respecto a datos experimentales donde manipulamos la información contextual asociada al ítem de estudio. En los experimentos 1 y 2 analizamos la disponibilidad de la información contextual tanto en el reconocimiento del ítem, como del contexto, como de su relación asociativa. En el experimento 3 sometimos a los sujetos a una tarea de pluralidad con el propósito de que se vieran forzados a usar la información contextual en el reconocimiento. Los resultados muestran que la gente adquiere información asociativa específica, aunque el reconocimiento global de los ítems no se ve afectado por el contexto. Los análisis ROC muestran que la gente usa dos procesos en el reconocimiento.

There is no single accepted view about how recognition proceeds. In fact there are two major group of theories, whose main difference is the postulated a different number of processes involved in recognition decisions. On the one hand, single process views (e.g., Donaldson, 1996; Glanzer, Kim, Hilford, & Adams, 1999; Hilford, Glanzer, Kim, & DeCarlo, 2002; Inoue & Bellezza, 1998; Ratcliff, Sheu, & Gronlund, 1992; Xu & Bellezza, 2001) whose major representative is the unequal variance signal detection model (e.g., Ratcliff et al., 1992) assumes that a single factor, usually called familiarity, is responsible for the decisions. This signal detection model assumes that there is a continuum of memory evidence over which the studied and non-studied items overlap and are normally distributed. The participant in a recognition experiment establishes a criterion in-between the two distributions that serves to decide whether a test item has been studied or not. Furthermore, when a receiver operating curve (ROC) relating hits (probability of saying old to studied items) to false alarms (probability of saying old to new items) is constructed, the theory (Hilford et al., 2002) predicts convex ROCs, their linearity when the ROCs are z-transformed, and slopes in the proximity of 0.80. These parameters have become standard «regularities» of recognition memory (Glanzer et al., 1999).

The other side of the explanatory accounts of recognition is taken by those who postulate the involvement of two processes underlying the decisions. Most theorists (e.g., Yonelinas, 1994, 2002) add to familiarity, the retrieval of specific contextual information (or recollection) as a second and differentiated mechanism contributing to recognition. These two process theories assume that recollection is slower, independent, has different neuro-anatomical substrate, and is more attention-demanding than familiarity. Although until recently, a dual model (Yonelinas, 1994, 2002) with signal detection behavior for familiarity and a threshold process for recollection was dominant, new two dimensional models (Dunn, 2004; Rotello, Macmillan, & Reeder, 2004; Rotello, Macmillan, Reeder, & Wong, 2005) using fully signal detection theory, offer also a promising account of recognition.

Both theoretical positions try to explain data obtained by process-dissociation (Jacoby, 1991), remember-know (Tulving, 1985), or classical receiver operating curve experiments (Hilford et al., 2002). Additionally, the recognition data generated with those three different methodologies have a strong relation to the investigation of the effect of context changes on memory. Research in context dependent memory is driven by the prediction that any change in context between the original and the current episode could produce unavailability of retrieval cues, with the subsequent decrement in performance (Smith & Vela, 2001). However, a major difference between both research fields is that

Fecha recepción: 18-12-05 • Fecha aceptación: 30-5-06
Correspondencia: Salvador Algarabel
Facultad de Psicología
Universidad de Valencia
46010 Valencia (Spain)
E-mail: Salvador.Algarabel@uv.es

the former tries to show detrimental effects, whereas in the main recognition research literature the opposite is the main goal.

Murnane and Phelps (1995) have set up a type of design for the study of context on recognition of interest for the present paper. They defined context as the combination of color (background & foreground), location, case, and font as displayed on the computer screen. When the test context was presented also at study time (AB-A design), the effect of context shift was not observed. But, when both were different, the detrimental effect was significant (AB-X design). We should expect that experimental manipulations in line with those of Murnane and Phelps, would affect either overall performance in hits, false alarms or discrimination, or/and the values of the estimated parameters according to signal detection analysis.

Having in mind the one versus two process explanations and the literature in the field of the effect of context changes on memory, our aim in this paper is twofold. First, to analyze the impact of specific as opposed to global contextual information for recognition in a situation in which the context is experimentally specified. In the current experiments we manipulate the availability of contextual information in such a way that we have conditions with and without it. Although the effect on recognition memory has traditionally been considered weak or nonexistent, and it continues to be controversial (see Alonso & Fernández, 1997; Fernández & Glenberg, 1985; Macken, 2002), a recent meta-analysis (Smith & Vela, 2001) shows that the effect size of context effects in recognition (.27) is significantly different from zero. This could mean that the context could come to mind as a co-associate of the item information, but not serve as a retrieval cue. That is, as in recognition the item itself is presented at test time, the influence of the specific contextual evidence could be restricted to a very narrow set of circumstances that depend on the specific context not being overshadowed by the global evidence associated with item presentation. The fact that in some experiments (e.g., Macken, 2002) differences in subjective judgments (remember/know) lead to an equality in performance is an indication of this possibility.

Our second goal is to compare ROC parameters between conditions with different level and type of contextual information. ROC analysis in recognition has a long tradition but its applications have not been extended to designs with explicit context manipulations.

In the current experiments, we define a specific context for each individually presented item (a word) as a combination of two colors. These two colors form a unique arrangement, distinct and easily treatable as a gestalt by the participant. To minimize the uncontrollable effect of conceptual processing at input, known to decrease the effect of physical context (Smith & Vela, 2001), we chose to present each item plus its context a number of times at a fast rate.

In the first experiment we contrast recognition performance on a list of words presented under different contextual conditions. Additionally, and without any previous indication, we tested also recognition performance of the background contexts. In one of the conditions (repeated), each word was presented on a specific and consistent background context. In a second condition (varied), a word was paired with a number of four different contexts. In a third condition (additional), additionally context-only trials were introduced to try to increase the recognition level of the context. Our goal is to compare one condition in which there is specific item-to-context associative information in memory against another in which there is not. In a previous paper (Algarabel & Pitarque, submitted), and using the remember/know methodology we have shown that both conditions evidence a similar level of remember and know responses, despite the differences in manipulation.

The repeated and additional study conditions of the second experiment were similar to the first one but here we changed the recognition instructions: participants were tested here for their knowledge of the individual items, for their knowledge of the context backgrounds (as in the former experiment), but also for their explicit knowledge of the association between the backgrounds and the items, following the methodology of the associative recognition experiments (see e.g., Hockley & Consoli, 1999).

Finally, in the third experiment we associated half of the presented words to the same background, and the rest to specific contexts. In this case, we tested the participants in a plurality task (Rotello, Macmillan, & Van Tassel, 2000), which maximizes the possibility of using the context to recognize the study words, given its high difficulty because the high similarity between old and new stimuli.

Data analysis is carried out looking out to the raw performance, discrimination indices (d'), and their relation to key ROCs parameters. In particular, we look at three parameters: the curvature of the untransformed ROCs, and the linearity and slope of the z-ROCs. According to past analyses (e.g., Hilford et al., 2002), ROCs must be curvilinear if the underlying mechanism based only in familiarity. On the other hand, the z-ROCs would be nonlinear if the dual models are correct, and linear, according to a signal detection based alternatives. Finally, slopes lower than one (around 0.80) are predicted by the signal detection models.

Given the requirements of design we could not gather enough individual data point from each participant to carry out parameter estimation from each subject. Correspondingly, the ROC analysis was carried out on group data (see e.g. Yonelinas, Kroll, Dobbins, & Soltani, 1999). As in previous papers (Glanzer et al., 1999; Hilford et al., 2002), we estimated the intercept and slopes of z-ROC by maximum likelihood and least square procedures for the group data. Given that the correlations between both estimations were 0.997 (intercepts) and 0.9983 (slopes) for the 16 conditions of our three experiments (Hilford et al., 2002), we used least square estimations throughout the paper.

Experiment 1

Method

Participants

Eighty-four Psychology students in the Psychology Department at the University of Valencia (Spain) participated voluntarily for extra course credit, and were randomly assigned to one of the three study conditions (28 subjects per condition). They completed the experiment in groups of 6 to 12.

Apparatus and materials

Two lists of 20 Spanish semantically unrelated words were prepared. The words were between 4 and 8 letters in length, with a frequency per two million (Alameda & Cuetos, 1995), imagery and concreteness (Bernia & López, 1985) of 42.8, 2.99 and 2.86 for list A, respectively, and 46.13, 2.99 and 2.84, respectively, for list B. Additionally, four sets of 20 frames measuring 102 (wide) $\times$ 59 (height) millimeters were created to serve as the contexts for the words presented. An example of word plus context can be seen in figure 1.

Each context frame was colored with a unique combination of two colors and presented on 17-inch computer screens at 1024 $\times$ 768 resolution. The inner area (78 $\times$ 35 millimeters) was chosen from (RGB values in parenthesis): maroon (128,0,0), green (0,128,0), olive (128,128,0), blue (0,0,255), purple (128,0,128), teal (0,128,128), silver (192,192,192), cyan (0,255,255), red (255,0,0), lime (0,255,0), yellow (255,255,0), magenta (255,0,255) and white (255,255,255). The remaining outer area was painted black (0,0,0), silver (192,192,192), red (255,0,0), yellow (255,255,0), blue (0,0,255), lime (0,255,0), or green (0,128,0). The combination of both sets of colors produced 91 unique compound frames. Six of the combinations placed the same color in both patches, and these were eliminated, and five additional context frames were used for practice. The 80 combinations of colors were grouped randomly in 4 sets of 20 for purposes of counterbalancing. Practice, study and test lists were randomized individually for each participant and presented on PC computer screens using *E-prime* software for experimental control (Schneider, Eschman, & Zuccolotto, 2002).

Design and procedure

The experiment consisted of four different tasks, preceded by a brief practice made up of 2 repetitions of 5 word plus context combinations (10 trials). During the study task, the participant studied 20 words written in Courier New font bold, 18 point in size and presented on a two-color background for 12 repetitions for a total of 240 trials. The instructions emphasized the need to study the words for a later unspecified memory test. Participants were also discouraged from forming images or inter-item associations among them. The trial presentation was randomized for each subject. Each stimulus was presented for 1 second, with 500 ms of inter-stimulus interval. There were three different study conditions: fixed, varied, and varied plus additional context-only presentations. In the fixed context condition, each word was associated with a unique context for the 12 presentations cycle, totaling 240 presentations of 20 words repeated 12 times each with the same context frame. This context was chosen from the 4 sets available and was counterbalanced across subjects. In the varied context condition, each word was presented 12 times paired with 4 different contexts (that is, 3 times associated to the same context), totaling 240 presentations of 20 words associated to 4 different contexts each. The contexts were chosen from the 3 blocks of context colors, and they were counterbalanced between subjects. The fourth set of colors was used for the recognition test. Finally, the varied plus additional context condition used the same pairing scheme as the varied group, but they also received context-only trials. In these trials, the context alone was presented for four trials interspersed among the word plus context trials. In this condition, the participants also received 240 word plus context (similar to the varied condition) plus 80 context only trials.

Figure 1. *Examples of a stimulus (word plus context) presented in the experiments (colours: blue —outer—, and lime –inner—)*

In a subsequent distractor task, every participant received 1000 trials (of 500 ms in duration plus 200 ms of inter-stimulus interval), in which five character strings (X, #, %, /, $) of 18-point size on a white background were presented. Some of these character strings included a single small letter. At the end of the section, subjects had to report the number of letters presented across the different trials.

The third task consisted of a common word recognition test of the 20 originally presented words mixed with 20 new ones, written in black characters on white background. For their response, the subjects had to indicate on a 6-point scale their degree of confidence in the yes/no response, from sure old to sure new. For this purpose, the keyword was covered leaving exposed the keys corresponding to the letters d, f, g, h, j, and k of the computer keyboard.

Finally, the fourth task was a «context» recognition test in which each participant was tested on a context presented alone and he had also to indicate on a 6-point scale their degree of confidence in the yes/no response, from sure old to sure new. For this task, the original 20 combinations of colors were mixed with 20 new ones not seen previously. In the case of participants who had received more than 20 different contexts at study, the choice of study background at testing was counterbalanced across subjects, in order to have all of them represented an equal number of times. The experiment was completed in about 40 minutes.

Results and discussion

The untransformed recognition scores for experiment 1 are presented in table 1. The significance level for all statistical tests was set at .05, unless otherwise noted. We analyze first the hits, false alarms and d' for words and for context judgments, and, secondly, we examine the parameters from the ROC and zROC quadratic and linear fits (Hilford et al., 2002).

Hits, false alarms and discrimination (d')

For word judgments (see table 1), the effect of condition was marginally significant on hits, $F(2, 81) = 2.98$, $MSe = 0.02$, $p = .06$. No significant effects were found on false alarms, $F(2,81) < 1$. Mean d' scores were calculated from individual participant d's, and they were significantly different, $F(2,81) = 3.62$, $MSe = 0.43$, $p = .03$. Newman-Keuls tests showed that the best performance was found on the additional condition, not having differences between the repeated and the varied conditions.

For context judgments there was a significant effect on hits, $F(2, 81) = 7.97$, $MSe = 0.02$, $p < .01$. Newman-Keuls tests showed that the worse performance was found on the varied condition, not having differences between the repeated and the additional conditions. With regard to false alarms, there were global differences between the three conditions, $F(2,81) = 5.71$, MSe=

0.03, p<0.01, showing the Newman-Keuls test that the repeated condition produced the smallest proportion of falses alarms. Context discrimination, as indexed by d', was significantly different in the three conditions; F(2,81)= 19.05, MSe= 0.23, p<0.01. Newman-Keuls tests showed that all contrasts were significantly different, except the comparison between the repeated and additional conditions. Subjects in the varied condition did achieve some context information because their d' of 0.20 was significantly different from zero, t(27)= 2.11, p<.05.

Overall, the results of this first experiment indicate that people acquired knowledge about the context, but they did not use it for word recognition, as if the global evidence associated with the item presentation overshadowed the contextual information. In the next experiment the instructions will ask explicitly for the subjects about if they are aware of the relationship between each word and its associated context.

Table 1
Mean proportions (and standard errors) for hits, false alarms (FA) and discrimination indices (d') for the conditions of the experiments 1, 2 and 3

	EXPERIMENT 1		
	Words		
	Hits	**FA**	**d'**
repeated	.78(.02)	.10(.02)	2.27 (.12)
varied	.75(.03)	.10(.02)	2.18 (.13)
additional	.84(.03)	.09(.02)	2.62 (.11)
	Contexts		
repeated	.70(.02)	.35(.03)	0.98 (.08)
varied	.58(.03)	.50(.03)	0.20 (.09)
additional	.71(.03)	.44(.04)	0.74 (.09)
	EXPERIMENT 2		
	Words		
	Hits	**FA**	**d'**
repeated	.79 (.04)	.16 (.03)	2.12 (.14)
additional	.71 (.05)	.14 (.02)	1.95 (.22)
	Contexts		
repeated	.79 (.03)	.47 (.05)	.99 (.20)
additional	.84 (.03)	.63 (.05)	.67 (.17)
	Associative		
repeated	.63 (.05)	.43 (.05)	.63 (.22)
additional	.71 (.04)	.71 (.04)	-.02 (.17)
	EXPERIMENT 3		
	Words		
	Hits	**FA**	**d'**
individual	.68 (.03)	.34 (.03)	.99 (.12)
common	.66 (.03)	.35 (.03)	.93 (.15)
	Associative		
individual	.76 (.04)	.72 (.04)	0.16 (.18)
common	.89 (.03)	.38 (.04)	1.85 (.18)

ROC and z-ROC analyses

Table 2 presents the parameters obtained in the transformation of the group's ratings into ROCs and z-ROCs by least square fitting.

In word recognition, although the three quadratic constants obtained in fitting a quadratic equation to the ROC data were clearly negative, they were not statistically different from zero, t(2)= 1.75, t(2)= 2.02, t(2)= 2.02, for the repeated, varied and additional conditions, respectively (see table 2), what means that these functions are not curvilinear. For context judgments, the equivalent quadratic constants did differ from zero, t(2)= 6.32, p<.05; t(2)= 58.88, p<0.01; t(2)= 5.55, p<.05, for the repeated, varied and additional conditions, respectively.

The z-ROC quadratic constants (word judgments) for the repeated, varied and additional conditions did not differ from zero, t(2)= 1.98, t(2)= 3.64, and t(2)<1, respectively (see table 2), what means that they were linear functions. For contexts, only the quadratic constant of the repeated condition did not differ from 0, t(2)<1, while the quadratic constants from the varied and additional conditions did, t(2)= 11.67, p<.01; t(2)= 4.65, p<.05, respectively.

All the z-ROC linear slopes (see table 2) for words were smaller than 1, t(3)= 3.27, p= .05; t(3)= 8.22, p<.01; t(3)= 5.14, p= .01, respectively, for repeated, varied and additional presentations. However, none of them differed from 0.80, t(3)= 1.22, t(3)= 2.67, t(3)= 1.57, respectively. For contexts, all the slopes did not differ from 1, t(3)= 2.58, t(3)= 1.52, t(3)<1, for the repeated, varied and additional conditions, respectively, reflecting the small context learning associate to these three conditions.

In sum, the ROC analysis indicates that the recognition models based on the signal detection theory have problems to fit our data: only three of the six ROC functions were curvilinear, two of the z-ROC functions were not linear, and half of the z-ROC slopes were not smaller than 1, as the signal detection models predict.

Experiment 2 is quite similar at the experiment 1 but we will explicitly ask each participant whether they recognized not only the studied words and their contexts but also of their relationship, following the procedure of the associative recognition experiments (see e.g. Hockley & Consoli, 1999). Only the repeated and additional study conditions will be manipulated here.

Experiment 2

Method

Participants

Forty Psychology students at the Faculty of Psychology, University of Valencia, participated voluntarily for extra course credit, randomly assigned to one of the two study conditions (20 subjects each).

Apparatus and materials

The same as the former experiment.

Table 2
Statistics (mean and standard error) for the ROCs and z-ROCs (group data) of experiments 1, 2 and 3

			ROC		z-ROC			
			quadratic constant	**R^2**	**linear slope**	**R^2**	**quadratic constant**	**R^2**
EXP. 1	words	repeated	-0.72 (0.38)	.86	0.68 (0.10)	.94	-0.29 (0.15)	.98
		varied	-0.85 (0.42)	.87	0.70 (0.04)	.99	-0.08 (0.02)	.99
		additional	-0.67 (0.33)	.85	0.71 (0.06)	.98	-0.04 (0.11)	.98
	contexts	repeated	-0.90 (0.14)	.99	0.94 (0.02)	.99	0.00 (0.04)	.99
		varied	-0.40 (0.01)	.99	1.05 (0.03)	.99	-0.07 (0.01)	.99
		additional	-0.87 (0.16)	.99	0.96 (0.04)	.99	-0.11 (0.02)	.99
EXP. 2	words	repeated	-1.03 (0.56)	0.83	0.79 (0.12)	0.93	-0.28 (0.08)	0.99
		additional	-0.89 (0.10)	0.99	0.65 (0.08)	0.96	0.28 (0.10)	0.99
	contexts	repeated	-0.47 (0.09)	0.99	0.73 (0.04)	0.99	0.05 (0.05)	0.99
		additional	-0.68 (0.06)	0.99	1.00 (0.03)	0.99	0.01 (0.04)	0.99
	associative	repeated	-0.77 (0.17)	0.98	1.16 (0.12)	0.97	0.39 (0.10)	0.99
		additional	0.16 (0.13)	0.99	1.03 (0.06)	0.99	0.11 (0.06)	0.99
EXP. 3	words	individual	-0.78 (0.11)	0.99	0.96 (0.03)	0.99	0.02 (0.06)	0.99
		common	-0.86 (0.03)	0.99	1.06 (0.06)	0.99	0.21 (0.01)	0.99
	associative	individual	-0.10 (0.12)	0.99	0.98 (0.04)	0.99	0.07 (0.07)	0.99
		common	-1.03 (0.28)	0.99	0.71 (0.13)	0.91	-0.20 (0.10)	0.97

Design and procedure

The experiment consisted of three different tasks each one preceded by a brief familiarization practice identical to the experiment 1. In the study phase the repeated and additional conditions were similar to the former experiment. The test phase consisted of 36 trials presented randomly. Participants were tested for their knowledge of the individual words (6 old plus 6 new), contexts (6 old plus 6 new), and contexts plus words (6 identical and 6 unpaired old context-word pairs). They were asked whether the word or the context was presented previously, or whether the context appeared with the same word as before. In any case, the response included a confidence response on a six point scale from sure old to sure new, as in experiment 1. The rest of the details were identical to the previous experiment.

Results and discussion

Hits, false alarms and discrimination (d')

The untransformed recognition scores for experiment 2 are presented in table 1 for words and contexts alone and for the combination of contexts plus words (associative).

For word judgments (see table 1), the effect of condition was neither significant on hits, $F(1,38)= 1.34$, $MSe= 0.04$, nor on false alarms, $F(1,38)<1$. Mean d' scores were also equivalent in the two conditions $F(1,38)<1$.

For context judgments, the effect of condition was not significant on hits, $F(1, 38)= 1.61$, $MSe= 0.02$, but it was on false alarms $F(1,38)= 4.57$, $MSe= 0.06$, $p<0.05$. No differences were found on d', $F(1, 38)= 1.48$, $MSe= 0.68$ between the two conditions.

Finally, for associative judgments, the effect of condition was not significant on hits, $F(1,38)= 1.67$, $MSe= 0.04$, but it was on both, false alarms, $F(1, 38)= 16.84$, $MSe= 0.05$, $p<0.01$, and d', $F(1, 38)= 5.45$, $MSe= 0.77$, $p<0.05$. Furthermore, with regard to this last analysis participants in the additional condition did not acquire any specific associative information because their discriminability index (~ 0) was not different from zero, $t(19)<1$.

Overall, the results of the second experiment indicate that participants achieved similar recognition levels across conditions in words as well as null context knowledge (Fernández & Glenberg, 1985; Macken, 2002). In the additional condition there was more false alarms in the recognition of the context and, as before, null associative word-to-context information.

ROC and z-ROC analyses

Table 2 presents the parameters obtained from the least square fit to group data. In word recognition the ROC quadratic constant for the repeated condition was not different from zero, $t(2)= 1.86$, but it was for the additional condition, $t(2)= 9.02$, $p<0.01$. For context judgments, the quadratic constants for both conditions did differ from zero, $t(2)= 5.55$, $p<0.05$, $t(2)= 11.13$, $p<0.01$, respectively, indicating that both functions were curvilinear. Finally, in associative judgments, the quadratic constant was also smaller than zero, $t(2)= 4.61$, $p<0.05$ in the repeated condition, but not in the additional, $t(2)= 1.30$.

With regards to the z-ROC quadratic constants (see table 2), the results indicated that for the word judgments the two conditions did not differ from zero, $t(2)= 3.51$ (repeated), $t(2)= 2.83$ (additional). The same pattern of results was found for context judgments, $t(2)= 1.06$ (repeated), $t(2)<1$ (additional) and for associative judgments, $t(2)= 3.74$ (repeated), $t(2)= 1.67$ (additional), indicating that the six conditions were linear functions.

With regard to the slopes of the z-ROC linear fits (words), in the case of the repeated condition they were not different from 1, $t(3)= 1.69$, but it was different in the case of the additional condition, $t(3)= 4.39$, $p<0.05$. For contexts, only the repeated slope differed from 1, $t(3)= 7.42$, $p<0.01$, whereas the slope of the additional condition did not, $t(3)<1$, nor the two slopes for associative judgments, $t(3)= 1.31$, and $t(3)<1$, respectively, for repeated and additional conditions. None of them were smaller than 0.80 by t tests.

Again, the ROC analysis indicates that the recognition models based on the signal detection theory have problems to fit our data: two of the six ROC curves were not curvilinear, and altought all zROC functions were linear, only two zROC slopes were smaller than 1.

In the third experiment, we will use a discrimination task to force subjects into using contextual information in the recognition of words. As in experiment 2 the task is one of item and associative recognition. However, at test participants have to discriminate between studied and plurality reversed distractors (Rotello et al., 2000). That is, if a singular item is studied, it must be distinguished from its plural and vice versa, presented as distractors. It is well documented (Rotello et al., 2000) that in these conditions, a recall-to-reject process may be used. With this strategy people try to remember the context to supplement item information and reach a more accurate response. In addition, more than using an individualized context for every item, we are going to simulate what probably happens in most experiments. That is, a common set of context cues will be associated to one specific item, whereas other specific items will be associated to different contexts, in a one to one way.

Experiment 3

Method

Participants

Thirty six Psychology students at the Faculty of Psychology, University of Valencia, participated voluntarily for extra course credit.

Materials

We constructed a list of 20 Spanish semantically unrelated words, half of them in singular and the other half in plural. All chosen words had a single dominant meaning, between 4 a 9 letters in length and with a frequency per two millions, imagery and concreteness of 63.40, 3.02, and 2.89 (Alameda & Cuetos, 1995; Bernia & López, 1985), respectively. From this base list, a second one was formed changing the plurality of the words, and converting singular into plural and vice versa. The four resulting lists were presented counterbalanced across subjects.

The backgrounds to be presented with each word were identical in size and RGB values as in the previous experiments. For the current experiment we prepared eleven two-colour backgrounds. Ten words in each list were presented with the same background (common condition), while the remaining 10 words were over imposed on a different background each (individual condition). The backgrounds used for the individual condition were counterbalanced across subjects, whereas the background used in the common condition was chosen randomly without replacement. The colours for the outer frame were either black, yellow, red, green, black, blue, yellow, blue, silver, silver and red, whereas for the inside could be red, olive, magenta, purple, teal, silver, blue, cyan, lime, yellow y green (see figure 1). From all combinations, those where no coincidence in color between outer and inner color and were clearly distinctive were chosen.

Design and procedure

The experiment consisted of three different tasks, each one preceded by a brief familiarization practice (10 trials). During the study task, the participant saw 20 words presented with their corresponding background for 12 repetitions for a total of 240 trials. Ten of these words were presented with ten different backgrounds (individual condition) whereas the remaining ten words were presented with a unique background each (common condition). The instructions emphasized the need to study the words for a later unspecified memory test. Additionally, participants were discouraged to form images or inter-item associations among words. Trial presentation was randomized for each subject. As before, each stimulus was presented for 1 second, with 500 ms of inter-stimulus interval.

The second task was a word recognition test of the 20 originally presented words mixed with 20 new ones in which the plurality was changed. For response, the participants had to indicate on a 6 point scale his degree of confidence on the yes/no response, from sure old to sure new.

Finally, without previous warning an associative recognition test was carried out in which subjects were ask to indicate whether the word plus the context were presented together at study or not. Twenty stimuli (word plus background) were presented for test. All of them had been seen at study, but half were unpaired, and the other half were intact. In both cases, half were presented with a common background and the other half with specific individual contexts, as in the study situation.

Results and discussion

Hits, false alarms and discrimination (d')

In the case of word recognition (see table 1), no differences were found among the two background conditions neither for hits, false alarms, or d's, $t(35)<1$ in all cases, indicating once again that possible differences in specific word-context associative information had no effect on word recognition. However, the results of the associative task were quite different (see table 1): the common context condition was more recognizable than the individual one, as indicated by its higher rate of hits, $t(35)= 3.00$, $p<.01$, its higher d', $t(35)= 6.30$, $p<.01$, and its lower rate of false alarms, $t(35)= 6.47$, $p<.01$, what means that subjects learned the specific item-context associations when they were repeated, but they do not use these associatins in recognizing the items.

ROC and z-ROC analyses

Table 2 presents the parameters obtained from the least square fit to group data. In word judgments, the ROC's quadratic constants of the individual and common conditions were different from zero, $t(2)= 7.29$, $p<0.05$, $t(2)= 34.23$, $p<0.01$, respectively. However in associative judgments both ROC's quadratic constants did not differ from zero, $t(2)<1$ (individual) and $t(2)= 3.65$ (common), although the last condition was marginally significant ($p= .07$).

With regards to the z-ROC quadratic constants (words), the common condition was different from zero, $t(2)= 17.33$, $p<0.01$, whereas the individual one was not, $t(2)<1$. In associative judgments the two z-ROC's quadratic constants were not different from zero, t(2)= 1.09, t(2)= 1.95, for the individual and common conditions, respectively, indicating linear functions.

Finally, the slopes of the z-ROC for words did not differ from 1, t(3)= 1.23 (individual), t(3)= 1.05 (common). For the associative task the same pattern of results was found, t(3)<1 (individual), t(3)= 2.24 (common).

This experiment has produced results in general congruent with the previous ones. Both conditions showed a similar level of recognition for words, but in the associative task subjects learned the item-context link in the condition of common context.

On the other hand, the ROC parameters show again that the recognition models based on the signal detection theory do not fit properly our data on the basis that half of the quadratic components obtained from the ROC quadratic fit are not different from zero, that z-ROC fit from the words common condition did differ from 0, that is, is not linear, and after all, that all the z-ROC slopes did not differ from 1.

General discussion

The three experiments show that people acquire information about the environment when their focus of attention is somewhere else, but within the present conditions, the process does not improve recognition. That is, the context appears to be a nonessential redundant piece of evidence for most recognition situations. To try to understand the reasons for this we have to look at the role of item and peripheral information in recognition. Whereas multidimensional item information (see Cleary, 2005) is present at test and the participant has only to identify it, the context has to be retrieved. From this point of view, the amount of item evidence available at test is much larger and has more weight than the specific information recollected, at least from the perspective of the requirements to be met for a recognition decision. Furthermore, as the prerequisite for the contextual information to act is the reliable presence of some degree of item knowledge, the presence of the latter is a requirement for the former. In conclusion, the context is overwhelmingly redundant with the item. The published studies with speed-accuracy trade-off (SAT) techniques prove this point (e.g., Boldini, Russo, & Avons, 2004): information given place to familiarity is active earlier than that resulting from the recollection process. Furthermore, as shown elsewhere (Algarabel, Pitarque, & Gotor, 2006) forgetting affects differentially the remember and know judgments: there is a greater probability with retention interval that remember responses changes to know than the other way around, indicating that the presence of contextual and item information define the global evidence at test time. Only in those situations in which there is item-to-context associative information, and the role of the item is very weak for whatever reasons, the former could help recognition above and beyond the item itself (see Algarabel & Pitarque, submitted).

Secondly, whereas in the first and second experiments participants had specific contextual details available, setting up the possibility that every item could be highly distinctive, in the third, the opposite was true in one of the conditions. In most recognition experiments with the remember/know methodology, the physical environment is very stable, despite of which, participants give a considerable high level of remember responses. There exists the possibility that people generate their mental and physical context that may tie to specific items, making them also distinctive and justifying the level of remember responses found in the published experiments. We think that the fact that participants may have specific discriminative cues associated to every item does not alter greatly their impact on performance, as can be concluded from the results of these experiments.

However, although it appears that in most situations the context is not a retrieval cue for the item, still the question remains whether recognition is or it is not a mixture of a two qualitative distinct or merely a one-dimensional process. The current data do not support the one process theory, at least within the current experimental conditions. As said in the introduction, the unequal variance signal detection model which is the most successful single process theory (Glanzer et al., 1999; Hilford et al., 2002) predicts convex ROCs, linear z-ROCs, and slopes lower than 1. In concrete, in item recognition slopes are found to average 0.80 for lower accuracy and 0.72 for higher accuracy. This difference in slopes is attributed to the effect of study process applied to the item list when there is at least certain level of performance (Glanzer et al., 1999; Hilford et al., 2002). This is what happens in the linear fits for words (experiment 1), but not in the evaluation of contexts. The same pattern of results are found in the second experiment, except for the repeated word condition which slope (0.79) was not smaller than 1, due probably to the great variability associated to this condition. Finally in the experiment 3 z-slopes for words were close to 1 because the level of performance was low, as was for the individual condition in the associative evaluation. However, in the common condition, the slope was 0.71 for the evaluation of the association. In this case, although all words and contexts have been presented for study, participants learnt the colours presented in the common condition, but not in the individual one. In conclusion, the slope of the adjusted z-ROC is lower than 1 in item recognition because the study episode usually affects only to the old item distribution.

The key prediction to distinguish among the different recognition theories is the linearity of the z-ROC data, as indicated elsewhere in this paper. The most up to date summary of published experiments (Hilford et al., 2002) indicates that z-ROCs are most of the time linear and not curvilinear. However, as we have argumented here and elsewhere (Algarabel & Pitarque, submitted), people use massively familiarity in recognition. The current experiments also show that in those conditions where there is no contextual evidence, the cuadratic component of the quadratic fit is not different from zero. If we do not take into account the particularity that recognition can be achieved most of the time by familiarity alone, we could interpret linearity as the general rule. In sum, our data, as other recent experimental data (see e.g., Boldini et al., 2004; Pelegrina & Tejeiro, 2006) supports the view that recognition is not only based on familiarity, but recollection must to be taken also into consideration. New two dimensional models (Dunn, 2004; Rotello et al., 2004, 2005) using fully signal detection theory, could offer a promising account for recognition.

Summarizing, words have colour contexts, but not the other way around, and the evaluation of associative information indicates that the first is no retrieval cue for the second, but that people evaluate the associative information from the point of view of familiarity.

Author's Note

This research was supported by Grant SEJ2004-02541 from The «Dirección General de Investigación Científica y Técnica» (Spanish Ministry of Education and Technology).

References

Alameda, J.R., & Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano.* Universidad de Oviedo, Departamento de Psicología, Oviedo.

Algarabel, S., & Pitarque, A. (submitted). Context, remember-know recognition judgments and ROC parameters.

Algarabel, S., Pitarque, A., & Gotor, A. (2006). Effect of the retention interval on the simultaneous cognate-noncognate and remember-know mirror effects. *Memory, 14,* 79-86.

Alonso, M.A., & Fernández, A. (1997). Contexto ambiental y memoria: ausencia de efectos en condiciones de procesamiento optimizado del entorno. *Psicológica, 18,* 209-225.

Bernia, J., & López, L. (1985). *Estudio normativo de vocabulario en siete dimensiones*. Unpublished manuscript.

Boldini, A., Russo, R., & Avons, S.E. (2004). One process is not enough! A speed-accuracy trade-off study of recognition memory. *Psychonomic Bulletin and Review, 11,* 353-361.

Cleary, A.M. (2005). ROCs in recognition with and without identification. *Memory, 13,* 472-483.

Donaldson, W. (1996). The role of decision processes in remembering and knowing. *Memory & Cognition, 24*, 523-533.

Dunn, J.C. (2004). Remember-know: A matter of confidence. *Psychological Review, 111,* 524-542.

Fernández, A., & Glenberg, A.M. (1985). Changing environmental context does not reliably affect memory. *Memory & Cognition, 13*, 333-345.

Glanzer, M., Kim, K., Hilford, A., & Adams, J.K. (1999). Slope of the receiver-operating characteristic in recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 25,* 500-513.

Hilford, A., Glanzer, M., Kim, K., & DeCarlo, L.T. (2002). Regularities of source recognition: ROC analysis. *Journal of Experimental Psychology: General, 131,* 494-510.

Hockley, W.E., & Consoli, A. (1999). Familiarity and recollection in item and associative recognition. *Memory & Cognition, 27*, 657-664.

Inoue, Ch., & Bellezza, F.S. (1998). The detection model of recognition using know and remember judgments. *Memory & Cognition, 26*, 299-308.

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

Macken, W.J. (2002). Environmental context and recognition: The role of recollection and familiarity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 28,* 153-161.

Mandler, G (1980). Recognizing: The judgment of previous occurrence. *Psychological Review, 87*, 252-271.

Murnane, K., & Phelps, M.P. (1995). Effects of changes in relative cue strength on context-dependent recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21,* 158-172.

Pelegrina, M., & Tejeiro, R. (2006). Parámetros ROC y z-ROC en memoria de palabras: efectos experimentales y preexperimentales. *Psicothema, 18,* 160-164.

Ratcliff, R., Sheu, C.F., & Gronlund, S.D. (1992). Testing global memory models using ROC curves. *Psychological Review, 99*, 518-535.

Rotello, C.M., Macmillan, N.A., & Van Tassel, G. (2000). Recall-to-reject in recognition: Evidence from ROC curves. *Journal of Memory and Language, 43,* 67-88.

Rotello, C.M., Macmillan, N.A., & Reeder, J.A. (2004). Sum-Difference theory of remembering and knowing: A two dimensional signal-detection model. *Psychological Review, 111,* 588-616.

Rotello, C.M., Macmillan, N.A., Reeder, J.A., & Wong, M. (2005). The remember response: Subject to bias, graded and not a process-pure indicator of recollection. *Psychonomic Bulletin & Review, 12,* 865-873.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide.* Pittsburg: Psychology Software Tools Inc.

Smith, S.M., & Vela, E. (2001). Environmental context-dependent memory: A review and meta-analysis. *Psychonomic Bulletin & Review, 8*, 203-220.

Tulving, E. (1985). Memory and consciousness. *Canadian Psychology, 26,* 1-17.

Xu, M., & Bellezza, F.S. (2001). A comparison of the multimemory and detection theories of know and remember recognition judgments. *Journal of Experimental Psychology: Learning, Memory and Cognition, 27,* 1197-1210.

Yonelinas, A. (1994). Receiver-operating characteristics in recognition memory: Evidence for a dual-process model. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20,* 1341-1354.

Yonelinas, A.P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory and Language, 46*, 441-517.

Yonelinas, A.P., Kroll, N.E.A., Dobbins I., & Soltani, M. (1999). Recognition memory for faces: When familiarity supports associative recognition judgments. *Psychonomic Bulletin and Review, 6,* 654-661.