Acta Psychologica 28 (1968) 45–62; © *North-Holland Publishing Co., Amsterdam*

THE GENERALITY OF WORD-ASSOCIATION HIERARCHIES

I. PRODUCTION AND RECALL OF ASSOCIATION RESPONSES

IDA J. VAN DER MADE-VAN BEKKUM

and

JITSKE PREVO–MOLENAAR

Psychological Laboratory, University of Utrecht, The Netherlands

ABSTRACT

*S*s responded to a stimuluslist in a free word association session (AS) and tried to recall their response to each stimulus in a recall session (RS). There was a significant positive correlation between commonality score (Co) in AS and number of correct reproductions in RS. All *S*s tended to gain in Co, but LC (i.e., *S*s of low Co in AS) more than HC (i.e., *S*s of high Co in AS).

Marbe's law applied both to the group of mean scores per *S* per session and (more or less) to the data of each *S* separately (no difference between HC and LC) in one or both sessions. Apparently the responses of LC really were, even for themselves, improbable ones likely to be forgotten and to be substituted by more probable ones (in terms of group norms).

It is generally supposed that associative commonality is based upon verbal habits shared by the members of a particular speech community. These verbal habits are assumed to be common rules of speech behavior for each member of the speech community. If each member really has these common verbal habits the consistently reported reliable individual differences in associative commonality require special attention.

These differences might imply that the *S*s who consistently react with only a few or no primary reactions (*S*s of low commonality or LC) have not the same verbal habits as the *S*s who consistently give many of them (*S*s of high commonality or HC). But equally well they might imply that LC – though having the same verbal habits – don't use them in the same degree as HC in a free association design. Such a design may differentially appeal to a *S*'s sense of originality. This explanation is the more plausible one, because we found, as an extension of Marbe's law, that LC tend to respond more slowly than HC.

In fact, we reported a significant negative product moment correlation between number of primary responses (i.e., the responses with highest association frequency) and the logarithm of the median reaction time of all responses of a *S* on a list of stimuli (VAN DER MADE–VAN BEKKUM 1966). Compared with HC, LC apparently tend to respond in a roundabout way if they have to give free-association responses.

We also found in accordance with JENKINS (1959) that commonality scores can be raised by an instruction 'to react with popular responses' (VAN DER MADE–VAN BEKKUM and VAN DER KAM, 1966). Our results indicated that the gain in commonality is not affected by a preceding free association instruction. Evidently *S*s can voluntarily adapt themselves to a definite level of commonality. This doesn't mean however that the differences between HC and LC disappear. Though they became smaller both JENKINS (1959) and we (VAN DER MADE–VAN BEKKUM and VAN DER KAM, 1966) reported a significant positive correlation between the commonality scores on 'free' and 'popular set' association. Thus even in the popular set design *S*s don't reach a common level of associative commonality. LC still succeed in responding in a roundabout way to quite a number of stimuli. Possibly they can easily do so because the top of the associative response hierarchy of a stimulus word often consists of opposites or synonyms of the stimulus. *S*s can avoid them categorically through the whole list and so even become a LC though they may still give highly probable other associations if there is a sufficient number of such stimuli in the list.

There is some evidence in support of this suggestion. CARROLL, KJELDERGAARD and CARTON (1962) reported that there is only a low (not significant) positive correlation between a *S*'s commonality score on the DOES (dominantly opposite evoking stimuli) and NOES (not dominantly opposite evoking stimuli) of the Kent–Rosanoff list. WALLENHORST (1965) found evidence for a number of idiodynamic sets (such as the set to give opposites) in the responses of his *S*s. His results also indicated a small but significant negative relationship between frequency of set reactions and mean reaction time (Rt) of a *S*, which may imply higher commonality of set reactions. Further research has to be done before it may be concluded – as CARROLL et al., (1962) did – that associative commonality is primarily based on the *S*'s set to give opposite or other set responses. Though it is not the principal principal object of our study we shall pay some attention to this point when we shall test the reliability of the commonality score used in this study.

First we shall test in this study the general hypothesis that though

LC probably always succeed in responding in a roundabout way a sample of responses of each *S* – including LC – will nevertheless show the function of the normative response hierarchy. This response hierarchy has to be taken – within the limits of its reliability – as a set of probabilities which applies to each *S*'s verbal behavior.

In this part of our study – which will consist of three parts in all – we will test the following specifications of this hypothesis:

I. If LC are to recall their association responses unexpectedly after some time they will perform poorer than HC, because their responses really were, even for themselves, improbable ones likely to be forgotten.[1] We predict that:

(1) There is a significant positive correlation between a *S*'s commonality score in a free association session and his number of correct responses in a recall session.

II. For all *S*s, incorrect responses in the recall session will tend to be more probable reactions than the original ones; restated in a testable prediction:

(1) There is a significant increase of a *S*'s commonality score in the recall as compared with the free association session.

Because all *S*s will be able to recall some of their original responses:

(2) There is a significant positive correlation between the two commonality scores though LC will probably gain relatively more than HC.

III. For each *S* individually responses of low probability will tend to be given slowly; responses of high probability tend to come fast. The prediction to be tested is:

(1) There is a significant negative correlation between association frequency and reaction time for each *S*. (Marbe's law, extended to individual scores). The association frequencies will be taken from the count for this group.

On average LC will have longer reaction times than HC. The prediction is:

---

1 For the present the word 'recall' will be used irrespective as to whether the process induced by the 'recall instruction' can adequately be described by the term 'recall'. We shall however return to the subject in the discussion of the experimental results.

(2) There is a significant negative correlation between commonality score and reaction time score on the whole stimulus list. (Marbe's law extended to a comparison of *S*s and verification of our own results (VAN DER MADE–VAN BEKKUM, 1966).

If prediction II.1. is verified a simple and straightforward application of Marbe's law would require that the recall responses, which tend to be more probable ones than the original ones, will be given faster than the association responses. The prediction is:

(3) There is a significant decrease of a *S*'s reaction time score on the whole list in recall as compared with the free association session.

If a *S* tends to respond slowly his individual correlation between association frequency and reaction time (prediction 1 of this section) may be lowered by the lack of reliability of his long associative reaction times. For a reaction time may be long because of a number of reasons besides low response probability. This is confirmed by the results of WALLENHORST (1965), who found a significant positive correlation between a *S*'s shortest reaction time and his mean reaction time but a non-significant correlation between a *S*'s mean and his longest reaction time. This might imply the following prediction:

(4) There is a significant correlation between a *S*'s associative commonality score and the degree of appropriateness of Marbe's law to these scores.

The first prediction of this section has already been tested by CASON and CASON (1925). They reported for all *S*s ($N = 28$) negative correlation coefficients ranging from –.11 till –.59, with a mean of –.33. They used the Kent–Rosanoff list which consists of 100 stimulus words which differ considerably in association value (AV) (defined as the relative association frequency of the primary response to that stimulus). Of course stimuli of low AV will tend to lower the reliability of the commonality score by the instability of their primary reaction words. Therefore – for scoring purposes – we shall use only part of the stimulus list, i.e. a number of critical stimulus words of reasonable high AV. We expect thereby:

(1) to achieve a higher reliability of both the commonality (Co) and the reaction time (Rt) score as compared with our earlier results for the same list (VAN DER MADE–VAN BEKKUM 1966).

(2) and – as a consequence of this prediction – to find also higher intercorrelations between these two scores and between the commonality scores on DOES and NOES stimuli.

METHOD

Subjects:

24 students of psychology, 11 females and 13 males, average age about 20 years.

Stimulus list:

The list consists of 48 short familiar Dutch words; 16 nouns, 16 adjectives and 16 verbs (see table 1). To give English readers an impression of the stimulus words the approximate English translation of the words has been stated below.[2]

For scoring purposes those 30 words have been chosen for which the following two statements are true:

(1) The words have the same primary response word in at least 2 out of 3 experiments recently performed at our laboratory. The experiment reported here is one of the three.

In fact 27 have the same primary response in 3 out of 3 experiments (the exceptions are marked in table 1).

(2) The relative association frequency of the primary response is at least 20 % in 2 out of 3 experiments. Again 27 stimulli reach this criterion in 3 out of 3 (see table 1 for the exceptions).

The critical words consist of 8 nouns, 12 adjectives and 10 verbs. Half of them are DOES. A DOES stimulus has been defined as a stimulus with a response given with a commonality of at least 70 % in another experiment – not yet published – in which *S*s were instructed to give opposites to the same stimuli. In table 1 the critical words are the stimuli for which primary reactions are stated. The DOES are printed in italics.

---

[2] The approximate English translation of the Dutch stimulus words is (in order of presentation in table 1):
giant, (to) sit, monkey, sour, shirt, (to) grow, (to) ask, black, sick, tree, (to) write, closed, hand, (to) tease, dirty, (to) fling, cold, man, (to) stick, beautiful, street, (to) sleep, good, train, house, clever, (to) have, stiff, mountain, (to) turn, hit (participle), fat, (to) work, spider, money, (to) be allowed to, bread, (to the) left (or left-handed), flame, old, (to) ring, (to) blow, soap, quick, (to) spash, time, (to) skip, small.

TABLE 1

The 48 stimuli in order of presentation with the primaries and association values in AS and RS of the 30 critical words. The 15 critical words which are DOES are printed in italics.

No	Stimulus	Primary Reaction	Ass. Value (number of *Ss*)		No	Stimulus	Primary Reaction	Ass. Value (number of *Ss*)	
			AS	RS				AS	RS
1	*reus*	dwerg	10	12	25	huis			
2	*zitten*	staan	10	13	26	slim			
3	aap				27	hebben	houden	6	7
4	*zuur*	zoet	15	16	28	stijf			
5	hemd	broek	8	7	29	berg			
6	groeien				30	draaien			
7	*vragen*	antwoorden	10	14	31	*raak*	mis	15	15
8	*zwart*	wit	16	21	32	*dik*	dun	19	21
9	ziek				33	werken			
10	boom				34	spin	web	10	13
11	schrijven	pen	11	11	35	geld			
12	*dicht*	open	15	17	36	mogen	kunnen*	8**	10
13	hand				37	brood	eten*	6	6
14	plagen	pesten	7**	8	38	*links*	rechts	21	22
15	vies				39	vlam	vuur	18	19
16	smijten	gooien	13	17	40	*oud*	jong	14	17
17	*koud*	warm	15	21	41	bellen			
18	*man*	vrouw	22	21	42	waaien	wind	15	15
19	kleven	plakken	15	16	43	zeep	wassen	11	11
20	*mooi*	lelijk	12	14	44	vlug	snel	11	14
21	straat				45	spatten			
22	slapen	bed	13	16	46	tijd	klok*	7**	6
23	*goed*	slecht	13	14	47	springen			
24	trein				48	*klein*	groot	15	17
								$\sum 381$	$\sum 431$

* secondary in 1 of the other 2 experiments

** less than 20 % in 1 of the other 2 experiments

Procedure:

There were two sessions: a free association session and a recall session. When participating in the free association part *Ss* didn't know that a recall part would follow. Recall was started exactly a week later, when all *Ss* had finished their free association session. In each session the stimuli were presented orally to each *S* individually. The order of the stimuli (according to their order in table 1) was rotated in such a way that the 24 *S*s each began with a different word (the first *S* with no 1, the second one with no 3, and so on). The *S*s responded orally. Their reactions were written down by E and their reaction times were measured by means of two microphones and an electric counter from the onset of the stimulus to the beginning of the response. Each session lasted for about 15 minutes on average. The instruction for the free association session was the usual one with time pressure, translated into Dutch. The recall instruction required the *S*s to give the same response as the week before, or if they couldn't recollect to try to guess it. The data of the association session have already been used with another objective in an earlier report (VAN DER MADE–VAN BEKKUM, 1966) to which we refer for more details about stimulus list and free association procedure.

RESULTS AND DISCUSSION

1. Reliability of the reaction time and commonality score

The critical words were divided in 2 × 15, in such a way that the halves were roughly matched to number of nouns, adjectives and verbs and to association value (AV) of the stimuli. For each *S* the median of the recorded reaction time (Rt) scores was determined, separately for the two halves and for the whole critical list, both for the association session (AS) and the recall session (RS). In the same way each *S* got 3 commonality (Co) scores on each of the sessions. The Co score is the number of primaries among the responses of a *S* on a list. The Rt and Co scores per S in AS and RS are presented in table 2.

To normalize the distribution of Rt's for all computations a log transformation has been applied to these scores.

The split-half correlation for the log Rt score in AS is .92 and in RS .93. The reliability of the score for the whole list estimated by the Spearman–Brown formula (GUILFORD, 1956: p. 452) is. .96 in both cases. For the Co score the split-half correlations are respectively .81 and .71 in AS and RS; The whole-test reliabilities estimated by the Spearman–Brown formula are .88 and .83 respectively.

The split-half correlation for the Co score on the whole list of 48 words for the same AS data only yielded a coefficient of .60 (VAN DER MADE–VAN BEKKUM, 1966). Apparently the reliability of this score gains much by omitting the stimuli of low AV. This is a verification of our prediction sub IV.1. It seems advisable to use in research studies only critical words of reasonable high AV for scoring purposes and to pay more attention to the Rt score because of its surprisingly high reliability. As a direct measure it has advantages over the more indirect and so less reliable Co score. With the critical Co scores we found a significant though low positive correlation between a *S*'s Co score on the 15 DOES and on the 15 NOES; r (Pearson) = .40 (df = 22; p = .05). CARROLL et al. (1962) who used the whole Kent–Rosanoff list reported a not significant coefficient of .26 (df = 44). Their conclusion that the widely used associative Co score is largely based on a *S*'s set to respond with opposites to DOES thus can't be verified. Nor can it simply be rejected because of the close relationship between reliability and association value. The mean Co for the DOES in our study is 9.2 (sd 4.0) and for the NOES 6.6 (sd 3.0). The difference is significant (t = 3.29; df = 23; p = .01). Perhaps the correlation between Co on DOES and NOES would be much better if the critical list were to consist of DOES and NOES of matched AV. This hypothesis has to be checked in further research.

II. *Commonality and recall score*

The mean number of primaries recalled in RS is 14.8 (sd 5.6) vs not recalled 3.2 (sd 2.1). The mean number of DOES recalled is 8.8 (sd 4.1) vs not recalled .4 (sd .8). The mean number of not-primary responses recalled in RS is 6.8 (sd 3.7) vs not recalled 3.5 (sd 2.5). These means added up only yield a total of 28.3 responses because on average there were. 1.7 response failures in RS. In AS only 1 failure occured (S no 1).

The mean number of all responses recalled in RS is 21.6 (i.e. 70 %) (sd 4.4). The r (*Pearson*) for the correlation between Co(AS) and the number of recalled responses in RS is .69 (df = 22; p = .01). This means a verification of our prediction I.1. LC perform poorer than HC in the recall session in such a way that the recall score depends on the associative commonality score.

The mean Co(AS) is 15.9 (sd 5.5) and the mean Co(RS) is 18.0 (sd 4.7). The difference in commonality between AS and RS is significant (t-test for dependent samples: t = 4.14; df = 23; p = .01). As we are especially interested in the relationship between the two Co scores we tried to determine the best fitting function for it. The F-test of linearity of regression yields a value of .316 (df = 9/13; n.s.), so the relationship may be represented by a straight line. The best fitting

TABLE 2

The scores per *S* in the association (AS) and recall session (RS)

	ASSOCIATION					RECALL					
Nr.	Comm. AS	MdRt /sec	N	r	p	Recall	Comm. RS	MdRt /sec	N	r	p
1	21	1.170	27	—.35	n.s.	27	22	1.120	28	—.72	.01
2	20	1.230	27	—.52	.01	22	20	1.240	28	—.53	.01
3*	17	1.480	29	—.72	.01	21	21	1.000	27	—.54	.01
4*	13	1.515	28	—.45	.05	20	14	1.470	25	—.28	n.s.
5	17	1.530	29	—.42	.05	25	20	1.425	28	—.17	n.s.
6	24	1.430	29	—.60	.01	26	26	1.160	30	—.49	.01
7	7	1.645	28	—.26	n.s.	14	15	1.870	30	—.32	n.s.
8*	21	1.370	29	—.67	.01	26	23	1.020	29	—.28	n.s.
9	18	1.430	29	—.54	.01	22	21	1.195	26	—.48	.05
10*	18	1.460	28	—.52	.01	17	19	1.260	25	—.60	.01
11*	11	1.310	27	—.60	.01	19	11	1.445	22	—.51	.05
12	11	1.800	27	—.51	.01	15	15	2.960	27	—.27	n.s.
13	22	1.265	30	—.59	.01	23	24	1.220	25	—.62	.01
14*	13	1.550	30	—.64	.01	19	18	1.600	29	—.55	.01
15*	4	4.175	30	—.50	.01	23	6	3.220	29	—.22	n.s.
16*	4	1.655	26	—.44	.05	11	12	1.420	29	—.32	n.s.
17*	22	1.110	29	—.75	.01	25	21	1.130	23	—.63	.01
18*	17	1.230	29	—.60	.01	28	17	1.185	28	—.43	.01
19	22	1.185	30	—.35	n.s.	26	23	1.025	28	—.34	n.s.
20	13	1.360	28	—.68	.01	20	15	1.405	28	—.38	.05
21	13	1.280	27	—.51	.01	18	10	1.885	24	—.48	.05
22	18	1.090	28	+.04	n.s.	20	19	1.210	25	—.31	n.s.
23	21	1.190	26	—.63	.01	26	23	0.960	26	—.46	.05
24*	14	1.570	29	—.50	.01	26	16	2.965	26	—.17	n.s.
Σ	381	36.030				519	431	36.390			
M	15.9	1.5012				21.6	18.0	1.5162			
sd	5.5	0.5866				4.4	4.7	0.6267			

* females

one by means of a least squares solution is: Co(RS) = .79 Co(AS) + 5.42 with a coefficient of determination (r^2) = .81.

This function implies that for a value of 25.8 the Co(AS) equals Co(RS). For lower values (i.e. the greater part of the range) the predicted value of Co(RS) is higher than Co(AS). Only for values of Co(AS) near the maximum of 30 the predicted value of Co(RS) becomes lower. Theoretically this is not unacceptable, because some amount of forgetting has always to be expected, which in the case of a Co score approaching the maximum can only affect primary responses.

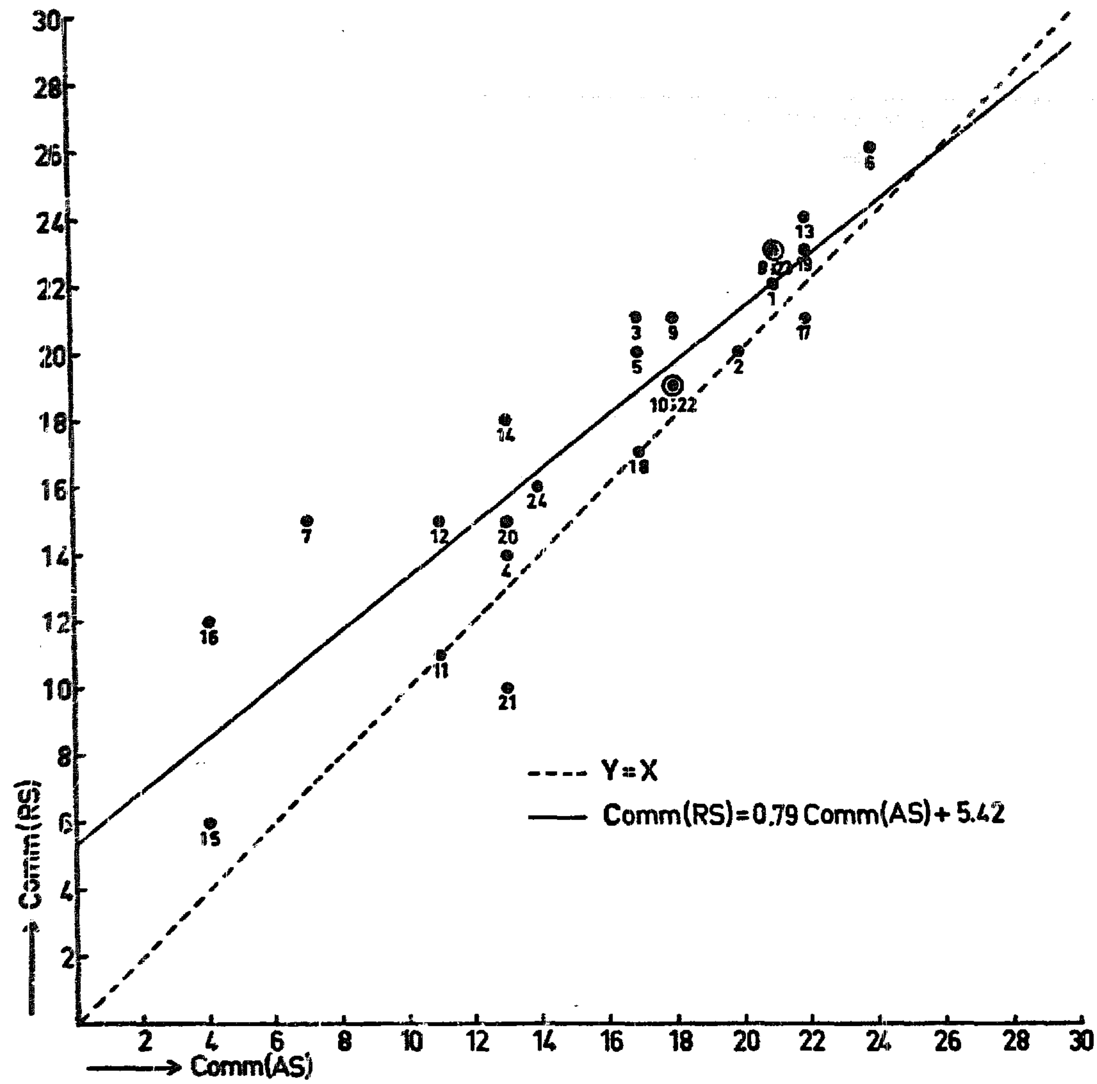

Fig. 1. The relation between commonality scores per S in free association and recall session (the numbers in the figure refer to the numbers of the Ss in table 2).

The function splendidly expresses the fact that LC tend to gain more than HC as JENKINS (1959) and we ourselves found before (VAN DER MADE–VAN BEKKUM and VAN DER KAM, 1966). To be able to compare the thus fitted function with another one for comparable data we computed the best fitting straight line function for our data of 1966. In that experiment the gain was induced by a 'popular set' instruction (PI) instead of the recall instruction of the present experiment. The resulting function is: Co(PI) = .64 Co(AS) + 7.75.

The point of intersection of this line with the line Co(PI) = Co(AS) is 21.53. The coefficient of determination (r^2) is .49 (N = 50 Ss). The mean Co for the 50 Ss of that study was 10.7 (sd 6.2) in AS on the same list of 30 critical words. This mean increased to 14.6 (sd 5.8) in PI. The gain in commonality in this experiment was much larger than in the present study, because this time we didn't demand higher commonality directly as we did in that experiment. However the mean level of Co in the present experiment is higher because of its individual and oral procedure which is known to increase commonality (see e.g. HORTON et al., 1963).

The point of equality of Co(AS) and Co(PI) again is higher than the maximum of the empirical range of Co(AS) scores, thus indicating a prediction of increase for all *S*s within the range. As we found in the present experiment the increase is larger for low values of Co(AS) as compared with high values of Co(AS). Hypothesis II.1 and 2 are thus both confirmed by our results. Incorrect responses tend to be more probable ones than the original reactions (I.1) and a *S*'s Co score in RS is predictable with reasonable accuracy from his Co score in AS (I.2) (see figure 1).

Related with the problem of this relationship is the question of the correlation between the association values of the stimuli in RS and AS. The significant increase in commonality from AS to RS may be equally spread over all stimuli or there may be a greater contribution of the stimuli of high AV. The latter is more likely if association frequencies can be taken as probabilities. The (Pearson) r for the correlation between the association frequencies of the primary reactions to the 30 critical words in AS and RS is .93 (df = 28; p = .01). Again the F-test of linearity yields a nonsignificant value (F = .96; df = 12/16) and the best fitting least squares solution is the function: AV(RS) = 1.03 AV(AS) + 1.24 (see figure 2).

For the data of VAN DER MADE and VAN DER KAM (1966) we found in the same way the function: AV(PI) = 1.25 AV(AS) + 2.06. The coefficients of determination (r^2) are respectively: .92 and .77. In both cases the slope of the line is larger than 1 so that AV(RS) and AV(PI) are always larger than AV(AS). The point of intersection with the line y = x has a negative value so that the increase of AV(RS) and AV(PI) is larger for words of high AV(AS). This is just what was to be expected if the value of AV(AS) is the empirical estimate of a probability. The function of the earlier experiment comes up to this expectation best because both AV(AS) and AV(PI) are based on free production whereas AV(RS) in the present experiment is based upon (correct and incorrect) reproduction.

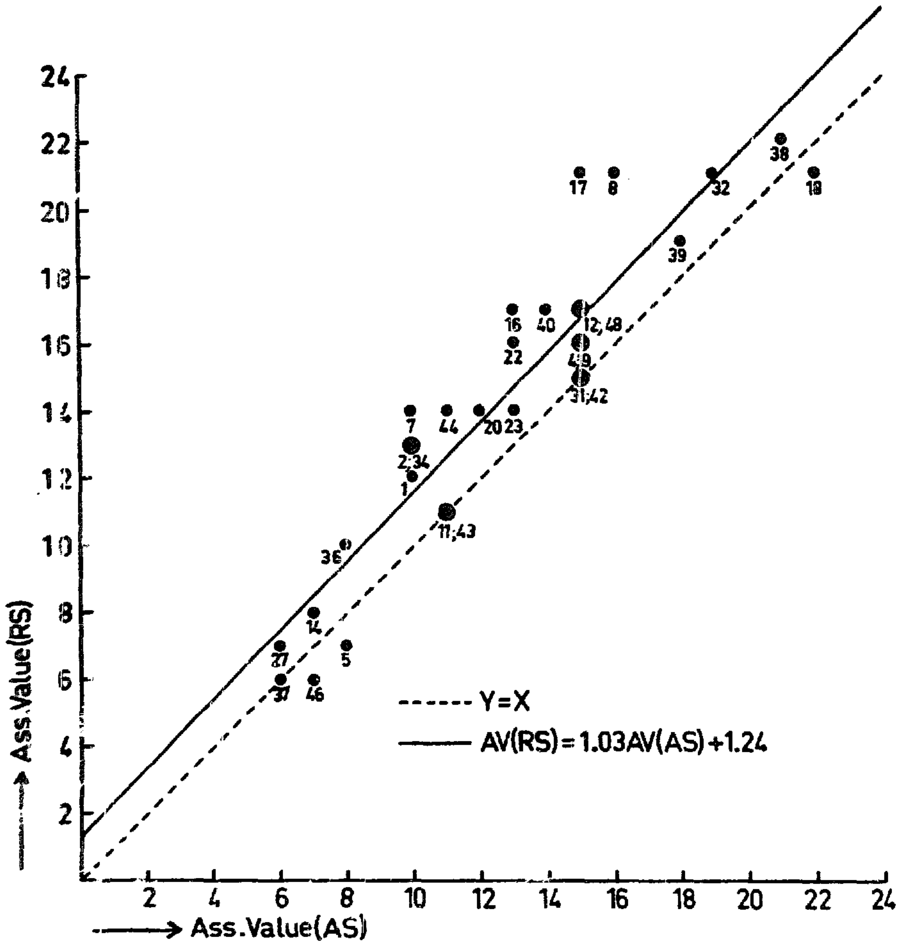

Fig. 2. The relation between association value per stimulus in free association and recall session (the numbers in the figure refer tot the numbers of the stimuli in table 1).

III. *Commonality and reaction time in AS and RS*

Table 2 gives the individual median Rt scores in AS and RS in seconds. The mean of the log median Rt's is 1.5012 (sd .5866) logsec in AS and 1.5162 (sd .6266) logsec in RS. The differences between AS and RS are not significant (t-test for dependent samples; t = .23; df = 23). Thus hypothesis III.3 has to be rejected: there is no systematic decrease in Rt in RS as compared with AS.

However from figure 3 it appears that HC – i.e., *S*s with an associative Co score of at least 16 (the mean score in AS is 15.9) – tend to respond faster in RS; whereas LC – i.e., *S*s with an associative Co score of 15 or less – tend to respond more slowly. Therefore the differences in Rt between AS and RS were tested separately for HC (N = 14) and LC (N = 10). The results are: HC are on average significantly faster in RS than in AS (t = 3.27; df = 13; p = .01). The mean Rt of HC in AS is 2.1106 logsec and in RS 2.0597 logsec. The differences for LC however are not significant ($t = -1.51$; $df = 9$). Their mean Rt in AS is 2.2236 logsec and in RS 2.2830 logsec. The coefficient for the correlation (product moment) between log MdRt in AS and RS

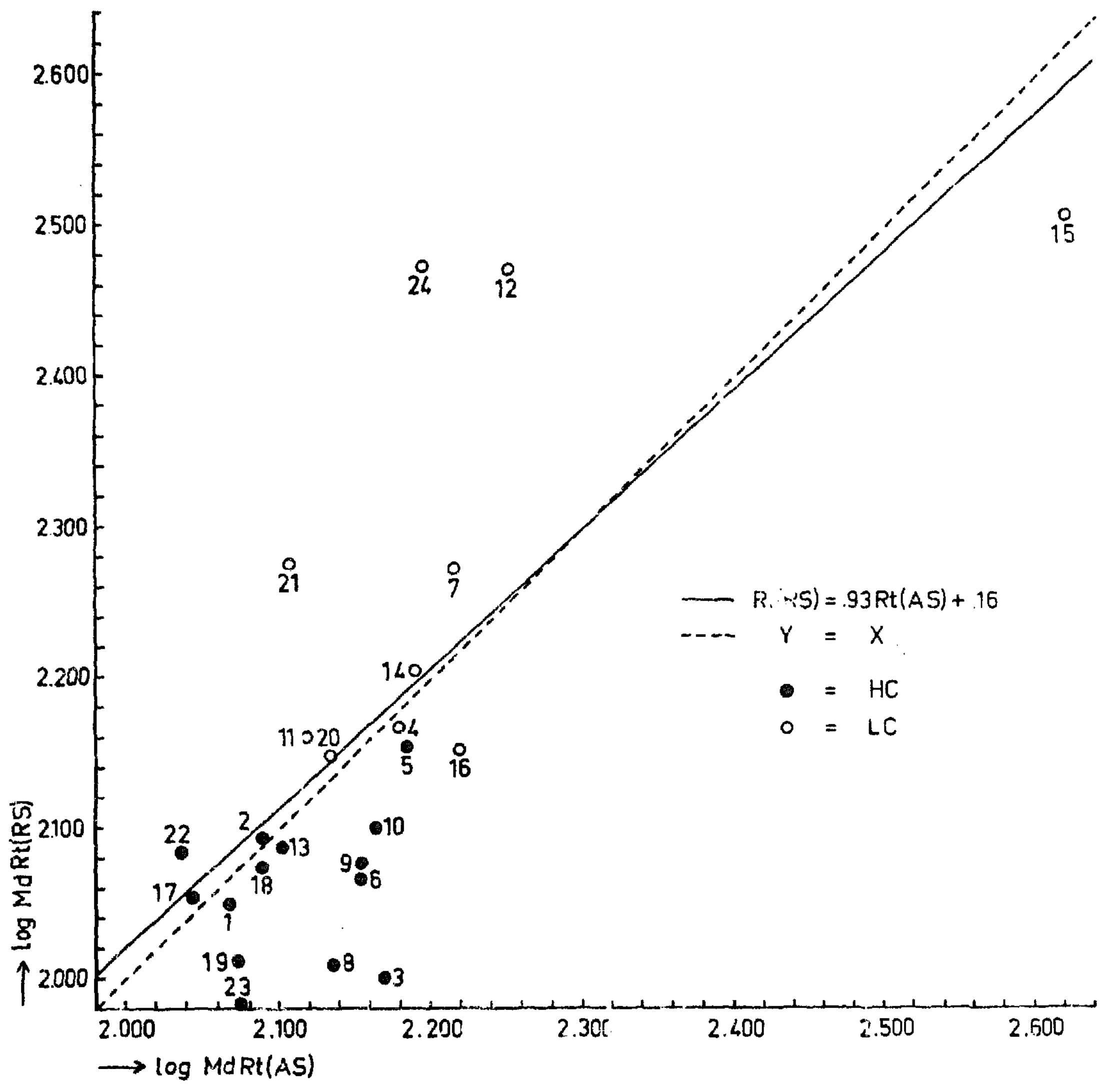

Fig. 3. The relation between the logarithm of the median reaction time per *S* in free association and recall session (the numbers in the figure refer to the numbers of the *S*s in table 2).

is .71 ($df = 22$; $p = .01$) for all *S*s. Separately for HC and LC respectively .26 ($df = 12$; n.s.) and .64 ($df = 8$; $p = .05$).

Apparently there are two subgroups of *S*s which behave differently. Almost all *S*s gain in commonality but only for HC does the gain result in a decrease in Rt as a simple and straightforward application of Marbe's law would predict. LC who gain relatively most in commonality don't show the decrease in Rt. After they had finished the RS task many *Ss* observed that often they had simply associated again to the stimuli, instead of recalling their original response. In that case they often wouldn't know if their response was a correct one until they had actually emitted it. The whole process tended to have an autonomous character. Though we can't deduct this from the *S*'s own remarks it seems highly likely that HC had this experience more often than LC. The former hardly need to recall their responses; they can confine themselves to associating again. Their Rt's decrease on account of familiarization (of stimulus and reaction words) and of gain in commonality. LC who generally will try to recollect their improbable original responses before daring to guess don't profit equally by familiarization and gain in commonality.

Both in AS and in RS Marbe's law applies to the individual Rt and Co scores of the *Ss*. In AS r (Pearson) is –.68 (df = 22; p = .01) and in RS r = –.72 (df = 22; p = .01). This is an improvement on our earlier results for the same data and a verification of our hypothesis III 2. In AS we then found only a coefficient of –.60 due to the use of all stimulus words including those of low AV for scoring purposes. The result for RS indicates an extension of Marbe's law to recall of response words. Both in free association and in the recall of free association responses there is a negative relationship between a *S*'s commonality and speed of reaction. This may be taken as an indication that in RS HC can respond directly whereas LC again (cannot but) take the roundabout way.

Marbe's law is also applicable to the log Rt and group frequency scores of the reactions of most of our *Ss*, taken individually. Table 2 contains the r's for each *S* in AS and in RS with the number of observations to which each coefficient refers. These numbers are often less than 30 and different for all Ss because of a different number of response (and instrumental) failures. In AS 20 out of 24 of the coefficients are significant: 17 at the 1 % and 3 at the 5 % level. In RS 14 out of 24 are significant: 8 at the 1 % and 6 at the 5 % level. Only three *Ss* (nos 7, 9 and 22) don't reach a significant level of r in at least one of the sessions. Yet each of them has at least once a coefficient of –.31. So all *Ss* tend to react more or less according to Marbe's law: frequent associations coming faster than infrequent ones. For most *Ss* our hypothesis III 1 can be maintained and our results seem to be somewhat better than those of CASON and CASON (1925), as we expected them to be, as a result of the use of a critical list.

In AS the mean Co of the *S*s with a significant correlation coefficient is 15.6 ($N = 20$). The mean of the others is 17.0 ($N = 4$). The biserial r for the correlation between commonality and significance of the Pearson r is only .144, which is not significant. In RS these means are resp. 19.1 ($N = 14$) and 16.3 ($N = 10$). Though this difference is, in contrast with the former one, in the expected direction it is not significant (r biserial is .368). In both cases we have to accept our null hypothesis: the associative commonality of a *S* doesn't have an effect on the applicability of Marbe's law to his scores. In a free association as well as in a recall task Marbe's law equally applies to HC and LC. Our hypothesis III 4 has not been verified, though there is a tendency in the expected direction, which may support our suggestion that HC can simply associate again in RS more frequently than LC.

Of our three main hypotheses the first two are clearly supported by the experimental results. The associative commonality score of a *S* is a good predictor of his performance in the recall of association reactions: *S*s of high associative commonality perform much better than *S*s of low associative commonality (hypothesis I).

All *S*s tend to gain in commonality – LC more than HC – and the associative commonality score is a good predictor of the commonality score in recall. Incorrect responses which are given more by LC, apparently tend to be more probable ones than the original associates (hypothesis II).

The third hypothesis concerning the relationship between Co and Rt requires some alternations. Though all *S*s tend to obey Marbe's law both in their individual scores and in their mean scores, they don't all of them decrease in Rt as their gain in commonality in RS would require. Only HC's scores showed such a decrease to a significant degree. There is some evidence that *S*s tend to associate again in the recall session which may explain the difference between HC and LC supposing that HC can associate more frequently and more directly than LC in RS. However the biserial r for the correlation between applicability of Marbe's law to the individual scores and the associative Co score was not significant. In RS there was a tendency in the expected direction, perhaps indicating that LC can indeed associate less in RS than HC; their r's tended to be more frequently nonsignificant. In AS there was no tendency in the expected direction which we take to mean that in free association the applicability of Marbe's law is really independent of a *S*'s score. Even the low association frequencies and long Rt's of the LC which may be much less reliable than the higher scores of the HC thus represent the negative correlation. Stated otherwise this may also be taken as an indication that for many research purposes there is no need for association norms for large samples. The count of the experimental group perhaps will suffice in many cases. The good reliability of our Co score is another support for this assertion.

Both the reliability of the Co and of the Rt score appeared to increase considerably by reducing the stimuluslist – for scoring purposes only – to a number of critical words of sufficient associative stability. Especially for research problems this reduction is a point to be considered again in later studies. For only by constructing systematically a stimulus list according to the relevant variables the question whether associative commonality is largely caused by the set to give opposites on a stimulus list or whether it has a wider significance can be answered.

On the whole our present results are good enough to conclude that the group response hierarchy of free associations more or less applies to the responses of all *S*s of our sample; i.e., in the specific way to which we restricted our predictions in this study. This means support for our general hypothesis that normative associative response hierarchies have a general validity for all members of a particular speech community. In studies to be published hereafter two more specifications of the same general hypothesis will be tested: the first concerning repeated associations, the second concerning paired-associate learning.

Conclusion

In this experiment 24 *S*s took part in a free association session (AS) and – a week later – in a recall session (RS). In the RS they were instructed to recall – or to guess at – their own association responses of the week before. The *S*s were not prepared to the recall session and had all finished their AS before RS was started.

The prediction that the number of correct reproductions would depend on a *S*'s Co score (i.e., his number of primary reactions) in AS was verified (r = .69). It was also found that the Co score in RS was significantly higher than the Co score in AS, whereas there was a significant positive correlation between the scores ($r = .90$). The function which fits the data best is: $Co(RS) = .79\ Co(AS) + 5.42$, which indicates that all *S*s gain in Co but *S*s of low Co (LC) more than *S*s of high Co (HC). The reverse holds for the association value of the stimuli (AV i.e., the association frequency of the primary reaction in this experiment): $AV(RS) = 1.03\ AV(AS) + 1.24$ ($r = .93$), which indicates that the gain is highest for high AV(AS). Both results are in agreement with the general hypothesis that association frequencies are empirical estimates of probabilities which roughly apply to all *S*s; the difference in Co between *S*s being primarily one of set. The 'set' hypothesis is also supported by the result that there is both in AS and RS a significant negative correlation (resp. $r = -.68$ and $r = -.72$) between a *S*'s Co score and his log MdRt score (Marbe's law applied to mean scores per *S*). LC tend to be slow. HC tend to be fast, both in free association and in recall. Apparently LC tend to react in a roundabout way; they avoid the top of the response hierarchy.

In RS the Rt score was significantly shorter than Rt(AS) only for HC, suggesting that HC hardly need to recall but can frequently simply associate again, whereas LC frequently have to search their memory for their original improbable responses. Their Rt neither increased nor decreased significantly. For all *S*s Marbe's law more or less applied to their individual scores in AS and RS. Each S reached at least once a coefficient of –.[illegible]1. With the exception of 3 *S*s all correlations in AS were significant. The prediction that the applicability of Marbe's law to individual scores would depend on the Co(AS) score has not been verified.

The Co and Rt scores per *S* in this study were not based on the whole stimulus list of 48 words but only on 30 critical words of reasonable associative stability. Split-half correlations for Co and Rt are resp. .81 and .92 in AS. In RS they are resp. .71 and .93. This is much better than the coefficients for the whole list and suggests that it would be advisable to use, at least in research studies, critical words for scoring purposes.

Our critical words consist of 15 DOES (dominant opposite evoking stimuli) on 15 NOES (not-DOES). The correlation between the Co scores and DOES and NOES is .40, which indicates a low though significant relationship. The set of a *S* to react with responses of high or low probability thus is not simply a set to react with opposites or not. Further research has to be done on this point because our DOES yielded a significantly higher Co score than the NOES, which results – as we know now – in a lower reliability of the Co(NOES) score.

References

CARROLL, J. B., P. M. KJELDERGAARD and A. S. CARTON, 1962. Number of opposites vs number of primaries as a response measure in free association tests. J. verb. Learn. verb. Behavior **1**, 22–30.

CASON, H. and E. B. CASON, 1925. Association tendencies and learning ability. J. exp. Psychol. **8**, 167–189.

GUILFORD, J. P., 1956. Fundamental statistics in psychology and education. New York: McGraw–Hill.

HORTON, D. L., D. MARLOWE and D. P. CROWNE, 1963. The effect of instructional set and need for social approval on commonality of word association responses. J. abnorm. soc. Psychol. **66**, 67–72.

JENKINS, J. J., 1959. Effects on word-association of the set to give popular responses. Psychol. Reports **5**, 94.

VAN DER MADE–VAN BEKKUM, I. J., 1966. Commonality and reaction time for free word associations of children and adults. (Commonaliteit en reactietijd bij de vrije woordassociaties van kinderen en volwassenen). Ned. T. Psychol. **21**, 102–135.

——— and P. VAN DER KAM, 1966. Effect on commonality of word-associations of instruction to give popular responses. Psychol. Reports **19**, 357–358.

WALLENHORST, R., 1965. Some relations between reaction time and choice of response in word association. Psychol. Reports **17**, 619–626.