APPLIED COGNITIVE PSYCHOLOGY, VOL. 4, 33–46 (1990)

Imagery Experts: How Do Expert Abacus Operators Process Imagery?

SHINSUKE HISHITANI

Department of Literature, Seinan Gakuin University, Japan

SUMMARY

This study examined how expert abacus operators process imagery. Without imagery instructions a digit series was auditorily presented as one whole number (WHL list) or separate digits (SEP list). RT from offset of the probe to onset of the response was measured. The main findings were as follows: experts showed no difference in RT between the two lists, while significant differences occurred in non-experts; non-experts' RT increased with probed position, while experts' RT was flat if the series size was within their image capacity; experts' RT increased with probed position when the series size was longer than their image capacity, but its rate of increase was smaller than that of non-experts; and the smaller the image capacity, the steeper the slope of the RT function. It was concluded that experts spontaneously encode the digit series into an imaged abacus, while non-experts encode it verbally; that experts directly access the probed position within their image but serially process the verbally coded overflowed part; and that non-experts search the digit series serially.

When we experience visual imagery, we gain the subjective impression that we can simultaneously see different objects contained in an image, and access directly at least some part of it. Many subjects report such introspections when they are asked to construct a visual image. This impression leads us to consider that visual imagery is quasi-perceptual with regard to the simultaneous attention to groups of objects and the free access to part of them.

However, some researchers criticize the attempt to infer the function and nature of imagery on the basis of the subjects' introspections (e.g., Pylyshyn, 1973). Certainly introspection may be fallible, but imagery is a private mental experience and, therefore, introspection must provide a necessary source of evidence (Marks, 1983). Furthermore, Miyazaki (1983), accepting Anderson's (1978) statement that plausibility is one of the criteria for deciding whether an image theory is appropriate, argued that if there is no essential principle for judging plausibility, to the extent that a plausible theory can be constituted, there is no reason why introspection cannot be a source of data. Therefore, the important point is that we must try to confirm introspection by behavioural data.

Two decades ago, Hebb (1966) attempted to show the picture-like nature of visual imagery using experimental data. If imagery is picture-like it is expected that imagined words can be spelled in a backward direction as fast as in a forward direction. Unfortunately, however, Hebb found that the forward spelling of imagined words is much faster than the backward spelling. Hebb concluded that whatever the subjective impression, a visual image is not of a picture-like nature.

0888–4080/90/010033–14$07.00

Received 16 March 1988
Revised 14 July 1988

About one decade later, Weber and Harnish (1974) re-examined Hebb's hypothesis. They contended that 'a spelling test is not necessarily visual', and that the subject 'may indeed have a visual image representation of the word but may still use his familiar verbal/speech representation for spelling in a forward direction, while possibly dropping back to his visual representation for the unfamiliar backward spelling' (pp. 409–410). A rask requiring the subjects to use visual imagery to generate responses was devised. In their experiment the subjects imagined three- or five-letter words, and had to say whether the letter at the probed position in an imaged word is vertically large lower case (e.g., b, d, . . ., y) or small (e.g., a, c, . . ., z). The reaction time (RT) from onset of the probe to onset of the response was measured. If the letter image can be directly accessed, RT should not vary with probe position. However, unfortunately, probe position effects were significant.

The above two studies seem to show a fatal flaw of introspection-based studies. However, an important point at issue may be what kind of people are employed as subjects. It is not always so easy for ordinary people to freely generate, maintain, and control imagery. Both Hebb (1966) and Weber and Harnish (1974) may not have confirmed their predictions because of instability in image-generation and control, even if their hypotheses were correct. It is said that expert abacus operators can spontaneously encode digit series into an abacus image and easily maintain and operate it (Hishitani, 1988). If abacus experts are employed as subjects, the above two hypotheses may be confirmed.

In fact, Hatano and Osawa (1983), and Hishitani (1987) found that RTs for reproducing digits were almost the same for backward and forward directions in the expert abacus operators. These results seem to support Hebb's (1966) original hypothesis. However, it may be impossible for the reproduction method to prove conclusively no difference between forward and backward reproduction times, because it is difficult to measure reproduction time precisely. Reproduction time can easily change according to pronounceability of the digit in a series. Subjects often fall silent during reproduction and it has not yet been decided how the silent time should be treated. Furthermore, reproduction time was measured manually by stopwatch in the experiments of Hatano and Osawa (1983) and Hishitani (1987). Therefore, if the difference in time between forward and backward reproduction is less than the precision of measurement, an erroneous conclusion may be drawn.

The Weber and Harnish (1974) technique does not have the problem of imprecision associated with the reproduction method, so their task is better than the reproduction method for examining whether imagery is directly accessed. The present study investigates how expert abacus operators process their images using a modified form of Weber and Harnish's technique. However, special imagery instructions were not given to the subjects, because some researchers have criticized previous imagery experiments for giving subjects too much instruction which may over-explicitly direct them as to how to perform (Pylyshyn, 1979, 1981; Yuille, 1983). Even with no instruction, experts can be expected to generate spontaneously the abacus image for the digit series. Furthermore, non-experts were employed as a control group. They must verbally encode digit series when imagery instruction is not given. Therefore, it will be clearly shown how experts process imagery by making a comparison between experts' and non-experts' performance.

In this experiment, the digit series was auditorily presented and the subjects responded with the number at the probed position in a digit series. Two memory lists were prepared. Each digit of a series in one list (SEP list) was separately pronounced (e.g., 1234 was read as 'one, two, three, four'), but each series in another list (WHL list) was presented as a whole number. For example, 5678 was read as five thousand six hundred and seventy-eight. *Kurai* is the general term for 'hundred', 'thousand' and so on in Japanese.

If non-experts encode the digit series into a verbal code and rehearse it as inner-speech, according to the theory of working memory (e.g., Baddeley, Thomson and Buchanan, 1975), they must search the series from the most significant position to the probed position in sequence. Non-experts have to pronounce *kurai* internally in addition to each digit in the case of whole numbers. Therefore, the following predictions were made for non-experts:

Prediction 1: Retrieval time will be longer in the WHL list than in the SEP list.
Prediction 2: The later the probed position, the longer the retrieval time.

On the other hand, if experts encode the digit series into an imaged abacus, they can directly access the probed position. Therefore, the following predictions were made for experts:

Prediction 3: There will be no difference in retrieval time between WHL and SEP lists.
Prediction 4: Retrieval time will not vary with probe position.

If the probed position is searched in parallel by experts, and in series by non-experts, it follows that:

Prediction 5: Retrieval time will be shorter in experts than in non-experts.

Is is assumed that there are individual differences in image capacity, i.e., the number of digits that experts can simultaneously imagine as a mental abacus. Since the experts are able to do concurrent verbal tasks by using the rehearsal buffer while holding the digit series (Hatano and Osawa, 1983), the expert with a relatively small capacity must treat the items which overflow this capacity in a verbal code. This means that retrieval speed should correlate positively with image capacity. Therefore, the last prediction is that:

Prediction 6: Retrieval speed will be faster in experts with a large capacity than those with a smaller capacity.

METHOD

Subjects

Nineteen skilled abacus operators, whose grades of expertise were higher than third *kyu* and lower than third *dan*, took part in the experiment as members of the expert group. *Dan* are classes for masters, and tenth *dan* is the highest. *Kyu* are classes for beginners and intermediates, and first *kyu* is the most advanced. Six of the experts were junior high school (age range 13–15 years), and 13 of them were high school students (age range 16–18 years). The non-expert group consisted of six junior high, and 13 high school students who had very limited experience in abacus operation.

Materials and apparatus

Two memory lists were prepared. Both lists consisted of eight series of digits. Two series were constructed for each length of 3, 4, 5 and 6 digits respectively by random selection from zero to nine. Both lists were presented auditorily. As stated earlier, each digit of a series in the SEP list was pronounced separately, but each series in the WHL list was pronounced as a whole number (e.g., in hundreds, tens and units). Therefore, for comparability, the series in both lists were constrained to have a most significant digit other than zero.

The subjects were required to make 72 responses for each memory list, because each serial position in a series was probed twice. These 72 cases were randomized, and SEP and WHL lists were tape-recorded separately at 2 digits a second. Both lists were presented to each subject, and the order of the lists was counterbalanced between subjects.

Subsets of SEP and WHL lists served as practice lists. Each subset consisted of three 3-digit series. RT was measured using a multi-purpose latency measurement system (Hishitani and Yumino, 1976). This system could measure the time from the offset of the tape-recorded digit series to the onset of the verbal response fed from a microphone, and the precision of measurement was 1 ms.

Procedure

Each digit series in a list was presented after a warning signal ('ready'). A bell marked the end of the series and a probed position was then announced. The first such position was the most significant digit, in the case of the WHL list, or the digit pronounced first, in the case of the SEP list. The subjects were asked to state orally the item at the probed position as quickly and as accurately as possible and the response time (PP–RT) from the offset of the probe to the onset of the response was measured.

The subjects were instructed that each series of digits in the WHL list would be presented as a whole number, so that a zero in the series would not be overtly pronounced. They had, however, to give zero as an answer if one was implied at the probed position.

In the first session, the subjects were given either a block of 72 WHL items (the WHL list) or a block of 72 SEP items (the SEP list) depending on the counterbalancing. This block was preceded by a practice list. Following a 5 minutes' rest, each subject received a further block of 72 items of the contrary kind preceded by an appropriate practice list.

At the end of the second session, both groups were asked to report their coding strategies, and the skilled abacus operators provided an index of their image capacity in the form of the number of digits in the largest numbers on which they could perform mental addition.

RESULTS

According to the introspections, as expected, non-experts verbally encoded input items, and experts used a verbal code together with imagery on demand. Their mean image capacity was 3.9 digits. The mean RTs in experts and non-experts are shown in Tables 1a and 1b respectively.

Analysis 1

In order to examine the relation between RT and list type in non-experts (prediction 1) and in experts (prediction 3), and RT difference between experts and non-experts (prediction 5), a 2 × 2 × 3 ANOVA was carried out with grade of expertise (expert and non-expert), list type (WHL and SEP) and series length (3, 4, 5 and 6 digits) as factors. Only the first was a between-groups factor. The results showed that non-experts' RT was longer than experts' RT ($F(1,36)=18.22$, $p<.001$), that RT for SEP lists was shorter than RT for WHL lists ($F(1,36)=27.21$, $p<.001$), and that RT increased with series length ($F(3,108)=79.08$, $p<.001$).

The interaction between grade of expertise and list type was significant ($F(1,36)=15.50$, $p<.001$). Further analysis of simple main effects revealed that there was no difference in RT between WHL and SEP lists in experts ($F(1,36)<1$), but in the non-experts RT was longer for WHL lists than for SEP lists ($F(1,36)=41.91$, $p<.001$).

The interaction between grade of expertise and series length ($F(3,108)=12.00$, $p<.001$) showed that the slope of the RT function of non-experts was steeper than that of experts (Figure 1).

Table 1a. Mean reaction times (ms) in experts

Series length	List type	Probed position						Mean
		1	2	3	4	5	6	
3	WHL	498	549	504				517
	SEP	592	613	531				579
	Mean	545	581	518				548
4	WHL	615	601	595	711			631
	SEP	621	657	635	638			638
	Mean	618	629	615	675			635
5	WHL	618	642	666	755	1070		750
	SEP	583	640	675	752	764		683
	Mean	601	641	671	754	917		717
6	WHL	675	635	752	1062	1496	995	936
	SEP	750	692	652	890	915	1062	827
	Mean	713	664	702	976	1206	1029	882
Mean	WHL	602	607	629	843	1283	995	827
	SEP	637	651	623	760	840	1062	762
	Mean	620	629	627	802	1062	1029	795

Table 1b. Mean reaction times (ms) in non-experts

Series length	List type	Probed position						
		1	2	3	4	5	6	Mean
3	WHL	665	756	749				723
	SEP	619	735	744				699
	Mean	642	746	747				711
4	WHL	775	1033	1432	1369			1152
	SEP	751	872	853	983			865
	Mean	763	953	1143	1176			1009
5	WHL	789	1000	1424	1557	2008		1356
	SEP	776	846	1146	1410	1316		1099
	Mean	783	923	1285	1484	1662		1228
6	WHL	892	945	1436	1866	2293	1769	1534
	SEP	761	882	986	1893	1658	1777	1326
	Mean	827	914	1211	1880	1976	1773	1430
Mean	WHL	780	934	1260	1597	2151	1769	1415
	SEP	727	834	932	1429	1487	1777	1198
	Mean	754	884	1096	1513	1819	1773	1307

Further analysis of the interaction of series length and list type ($F(3,108)=7.15$, $p<.001$) was not done, because this result had no relation to the predictions, and was not theoretically important.

Analysis 2

Nineteen experts were divided into three subgroups on the basis of their image capacity to examine prediction 6, that retrieval speed depends on the experts' image capacity. Mean image capacity was 5.2 digits in six high-capacity subjects, 3.7 digits in seven medium-capacity subjects, and 2.8 digits in six low-capacity subjects. The slope of the function relating RT to series size, and the mean RT over all sizes of series, were calculated for each subgroup. Slopes and mean RTs were plotted as a function of image capacity. Figure 2 shows that they decrease linearly with image capacity.

Analysis 3

In order to examine whether or not the experts can directly access the probed position on an imaged abacus (predictions 2 and 4), four ANOVAs were performed on RT with grade of expertise, list type and probed position as factors.

The main effects of expertise and probed position were significant for all series lengths. The former showed that RT was shorter in experts than in non-experts ($F(1,36)=9.33$, $p<.005$ in 3-digit series; $F(1,36)=12.44$, $p<.005$ in 4-digit series; $F(1,36)=22.73$, $p<.001$ in 5-digit series; $F(1,36)=17.73$, $p<.001$ in 6-digit series).

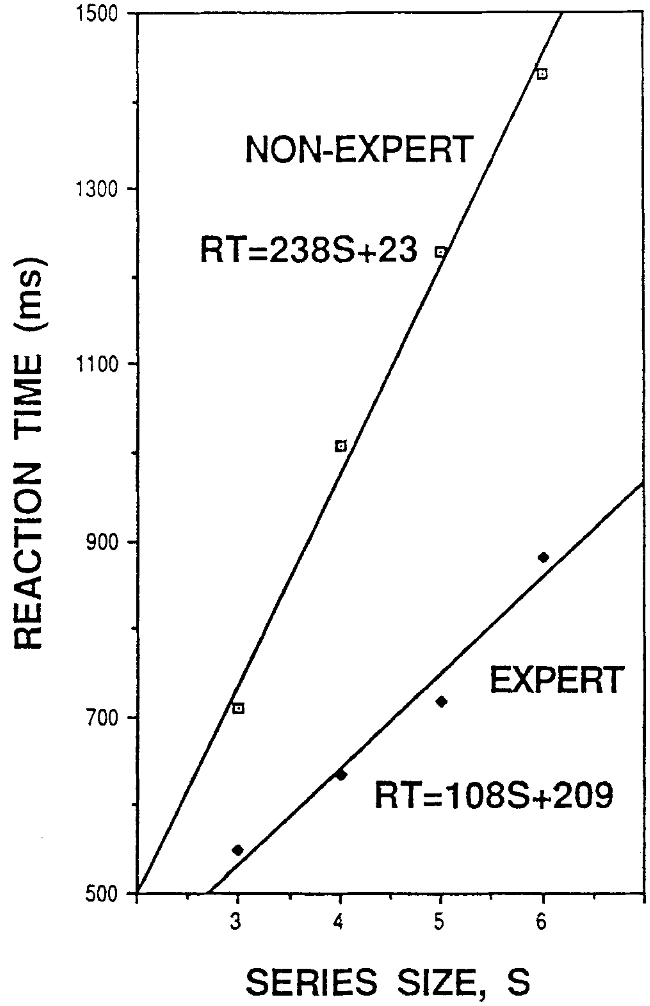

Figure 1. Reaction time of experts and non-experts as a function of series size.

The latter indicated that RT depended on the probed position ($F(2,72)=4.09$, $p<.025$ in 3-digit series; $F(3,108)=5.09$, $p<.005$ in 4-digit series; $F(4,144)=24.22$, $p<.001$ in 5-digit series; $F(5,180)=31.71$, $p<.001$ in 6-digit series). The effect of list type was not significant in 3-digit series ($F(1,36)=1.11$), but RT for SEP lists was shorter than that for WHL lists in the other series ($F(1,36)=13.59$, $p<.001$ in 4-digit; $F(1,36)=12.98$, $p<.005$ in 5-digit; $F(1,36)=16.08$, $p<.001$ in 6-digit).

The interaction between expertise and list type was significant in 3- ($F(1,36)=5.47$, $p<.05$), 4- ($F(1,36)=14.97$, $p<.001$) and 5-digit series ($F(1,36)=4.41$, $p<.05$), but not in 6-digit series ($F(1,36)=1.54$), where RT for the WHL list was longer than that for the SEP list in both groups. The tests of simple main effects showed that in 4- and 5-digit series there was no difference in RT between the WHL and SEP lists in experts ($F(1,36)<1$ in the 4-digit series, and $F(1,36)=1.13$ in the 5-digit series, respectively), but RT for the WHL list was longer than that for

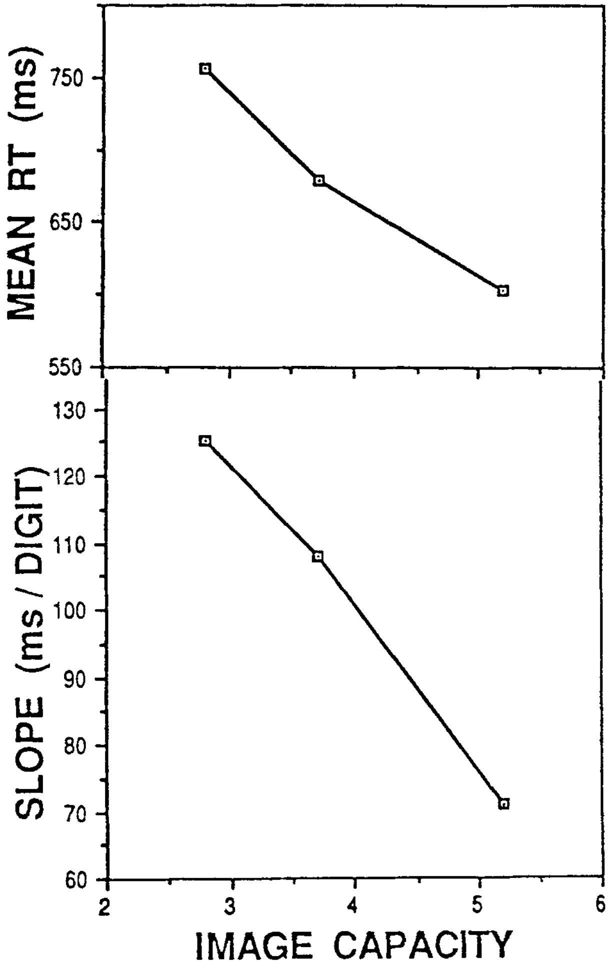

Figure 2. Mean reaction time and slope as a function of image capacity.

the SEP list in non-experts (*F*(1,36)=28.55, *p*<.001 in the 4-digit series, and *F*(1,36)=16.26, *p*<.001 in the 5-digit series). In 3-digit series, however, non-experts did not show the difference in RT between list type (*F*(1,36)<1), and RT for the WHL list was shorter than RT for the SEP list in experts (*F*(1,36)=52.80, *p*<.001).

The expertise by probed position interactions were significant in all series (*F*(2,72)=3.61, *p*<.05 in 3-digit; *F*(3,108)=3.74, *p*<.025 in 4-digit; *F*(4,144)=7.00, *p*<.001 in 5-digit; *F*(5,180)=6.05, *p*<.001 in 6-digit). Further analysis of simple main effects showed that in 3- and 4-digit series the probed position effect was significant in non-experts (*F*(2,72)=6.09, *p*<.005 and *F*(3,108)=8.69, *p*<.001 respectively), but not in experts (*F*(2,72)=1.04 and *F*(3,108)<1 respectively). Both groups showed significant simple main effects of probed positions in 5-digit series (*F*(4,144)=3.23, *p*<.025 in experts, and *F*(4,144)=28.01, *p*<.001 in non-experts), and in 6-digit series (*F*(5,180)=5.97, *p*<.001 and *F*(5,180)=31.81, *p*<.001

respectively). As shown in Tables 1a and 1b, however, RT increases more slowly with the probed position in experts than in non-experts, especially RTs in the first three or four positions, which seem almost equal in experts.

These results indicate that the function of RT against probe position is flat at least within the first three or four positions for experts, while the function increases linearly in non-experts. This is clearly confirmed in Figure 3, where RT averaged across all series is plotted from position 1 to 3.

All three factors significantly interacted only in the 4-digit series ($F(3,108)=5.67$, $p<.005$), but not in 3- ($F(2,72)<1$), 5- ($F(4,144)<1$) and 6-digit series ($F(5,180)<1$). Further analyses of the three-factor interaction in 4-digit series revealed that there was no interaction between list type and probed position in experts ($F(3,108)=1.83$), but the simple interaction for non-experts was significant ($F(3,108)=33.84$, $p<.001$). This means that RT for SEP lists increases more slowly with probed position than RT for WHL lists in non-experts relative to experts.

The list type by probed position interaction was significant in 4-digit ($F(3,108)=6.18$, $p<.001$), in 5-digit ($F(4,144)=4.92$, $p<.005$) and in 6-digit series ($F(5,180)=5.13$, $p<.001$) but not in 3-digit series ($F(2,72)<1$). Further analyses of the above

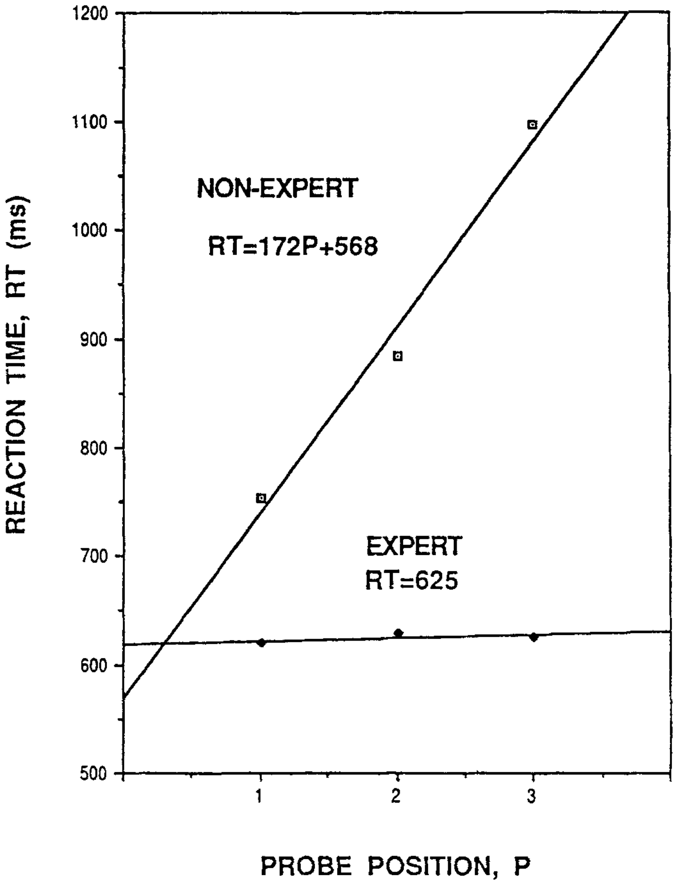

Figure 3. Combined reaction time of experts and non-experts from positions 1 to 3.

significant interactions were not performed, because they had no relation to the predictions, and were of no theoretical importance.

DISCUSSION

Significant interactions of expertise with list type resulted from analysis 1 and analysis 3 for 4- and 5-digit series. There was no difference in RT between WHL and SEP lists in experts, but RT was longer in the WHL list than the SEP list in non-experts. These results are consistent with the subjects' introspections about their coding strategies and imply the following: non-experts needed longer retrieval time for WHL lists than SEP lists, because they verbally rehearsed input digit series, and had to pronounce internally additional *kurai* in comparison with SEP lists; on the other hand, retrieval time in experts was not affected by whether or not the digit series had *kurai*, because they held input series as imagery.

In the 3-digit series, however, the mean RT for the WHL list was shorter than that for the SEP list in experts, and there was no difference between list types in non-experts. Even with series from the WHL list, non-experts may be able to neglect *kurai* and rehearse them in the same way as those from the SEP list if they consist of strings not exceeding three digits. *Kurai* seems to associate closely with its position on an abacus in experts, for they are very familiar with numbers. Therefore, *kurai* may function as an effective retrieval cue for a short digit series which does not exceed 3 digits, so that RT is shorter for WHL than for SEP lists.

In 6-digit series, the mean RT for WHL lists was longer than the mean RT for SEP lists in experts as well as in non-experts. A possible reason for this is that experts verbally rehearse the parts of input series which overflow image capacity (on average, 3.9 digits, according to introspections from experts).

Although the 5-digit series exceeded image capacity and mean RT was longer in the WHL list than the SEP list, there was no significant difference between them. Since the capacity was exceeded by only one digit, the overflow effect may not have been so large enough to show.

The above results do not unconditionally support prediction 1, that in the non-experts RT is longer in WHL lists than SEP lists, and prediction 3, that in the experts there is no difference in RT between WHL and SEP lists. However, at least it can be concluded that if the size of the input digit series is moderate, RT is longer in WHL lists than SEP lists in non-experts, and that in addition to the series size condition, if the input series is within image capacity, RT for WHL lists is equal to RT for SEP lists in experts.

Analysis 3 showed that in 3- and 4-digit series, RT in experts is constant across different probe positions, while RT in non-experts increases with it. These results correspond to the subjects' introspective reports that experts and non-experts encode input series into imagery and verbal codes respectively, and that the image capacity of the former group is four digits. According to the above behavioural data and introspections, it is considered that non-experts search the input digit series from the most significant position to the probe position in sequence, and that experts directly access the probe position, if the series size does not exceed image capacity. This is clearly demonstrated by Figure 3. Therefore, predictions 2 and 4, which are respectively that RT for non-experts is longer for later probed positions and that experts' RT does not vary with probed position, appear to hold good for 3- and 4-digit series.

In 5- and 6-digit series, however, even in experts, RT increased with probe position, though the rate of increase of RT was lower in experts than non-experts. This may imply that when series size exceeds image capacity, experts directly access probe position within their capacity and process the overflowed part sequentially as verbal code. In other words, experts might employ a hybrid method between parallel and serial processing, and could process digits faster in parallel than serially so that, on the whole, the PP–RT function increases more slowly in experts than non-experts.

The above interpretation leads to the following expectations: experts will show no difference in the slope of the PP–RT function between WHL and SEP lists in 3- and 4-digit series which are almost within image capacity, but their slope for WHL lists will be steeper than that for SEP lists in the 5- and 6-digit series, because the *kurai* of the overflowed part must be verbally rehearsed in WHL lists. On the other hand, in non-experts, the slope will be steeper in WHL lists than SEP lists in all series, because of the need to rehearse *kurai* in WHL lists. Consequently, in 3- and 4-digit series, the simple interaction between list type and probe position will be significant in non-experts but not in experts; in 5- and 6-digit series it will not be significant for either group. Therefore, the three-way interaction of expertise, list type and probe position will be significant only in 3- and 4-digit series. Except for 3-digit series, this expectation was confirmed. In the 3-digit series there was no difference in the slope of the PP–RT function in non-experts or experts, because, as stated earlier, *kurai* of series from WHL lists can be neglected, and the series can be rehearsed in the same way as for one from SEP lists. The mean RT in experts was shorter than the mean RT in non-experts, and retrieval speed increased linearly with image capacity. These results confirmed that RT is shorter in experts than in non-experts (prediction 5) and that retrieval speed is faster in experts with a large capacity than in those with a small capacity (prediction 6).

The facts can also be explained in terms of the information-processing methods employed by experts and non-experts stated above; namely that the former use parallel together with serial processing, whereas the latter use only serial processing. Since serial processing is relatively slower than parallel, information-retrieval speed should be faster in the experts than in the non-experts (Figure 1).

The main point of this study is that, although no special instructions to use imagery were administered, experts reported that they did use an abacus image to perform the task, indicating that the abacus image is a usual way for them to represent numbers.

A second major point is that the abacus image is so stable that it can be accessed reliably and freely. In the experiment by Weber and Harnish (1974) there was nothing to help the subjects stabilize their images, and there is no evidence that the letter images were stable and familiar to their subjects. This may be the reason that a flat PP–RT function was not found. On the other hand, Hishitani (1983) asked non-experts to project *katakana* (Japanese alphabet) letter-images into a matrix on a screen in order to stabilize their images, and obtained flat PP–RT functions like the ones in this experiment. It is supposed that the RT functions were flat, because even non-experts could maintain the images stably with the help of an external perceptual framework.

The third important conclusion from this study is that valid and precise introspective reports are obtainable from experts. This can be supported by the fact that an average image capacity of 3.9 digits reported by the experts agrees with the directly accessible figure of 4 digits objectively confirmed from the PP–RT functions. Since imagery appears to be highly stable in abacus experts, and they have many opportunities for using and observing it, they appear able to inspect their imaged abacus with great precision and accuracy, in a way similar to actual visual perception.

As stated above, overall, RTs produced by experts were significantly shorter than those of non-experts, and in some cases there was no difference in RT between conditions for experts. It may therefore be asked whether the RT of experts might have reached a floor, in which case some important results in the present experiment could be artefacts of such floor effects.

Since there has been no previous experiment in which exactly the same task was used as in the present study, it is difficult to show directly that the RT even of non-experts can sometimes be as short as, or shorter than, that of experts—that is, that the present data gained from the experts do not reach a floor. Previous studies, however, seem to show indirectly that floor effects have not occurred in this experiment.

Using the classic short-term retrieval task similar to the present task, Sternberg (1966) showed that the mean retrieval times for 3- and 6-digit series were between 500 and 525 ms, and between 600 and 625 ms, respectively. In the present experiment the mean RT of experts was 548 ms for 3-digit series and 882 ms for 6-digit series. Kristofferson (1972) required the subjects to perform the same task as Sternberg's (1966) for 144 trials per day for 30 days. After the practice the mean RT for 4-digit series was slightly less than 450 ms. In the present experiment the mean RT of experts was 635 ms for 4-digit series. In another experiment, Seamon (1972) asked subjects to relate three concrete nouns to each other and construct an image of a scene. Following that, Seamon presented a target item and measured the time which the subjects needed to decide whether or not the target was in the previously presented three nouns. The mean RT was slightly less than 540 ms.

The above studies which used a similar short-term retrieval task to the present one imply that experts' RT in the present study did not approach close to the minimum. Moreover, the experts in this experiment were highly skilled abacus operators whose grades of abacus operation are between second *kyu* and second *dan*, but their expertise was a long way from the level of grand master whose grade is tenth *dan*. Therefore the performance by the experts in the present experiment cannot be considered to have reached a ceiling.

Nevertheless the RT pattern of the experts was very different from that of the non-experts. This means that the expertise of abacus operation affected task performance. It has been argued already that this effect is due to the ability with which the experts can encode a digit series into an abacus image.

In conclusion, the results of the present experiment suggest that the imagery system creates an imagery space in which multiple images can be maintained and freely accessed. Parallel processing of imagery resembles the simultaneous processing of groups of objects in perception. In this respect imagery really does seem to be quasi-perceptual. Furthermore, it can be said that the study of imagery experts can provide some important clues to the understanding of imagery processes which may not be so obvious in studies using only non-experts as subjects.

However, it may be considered that one serious issue for the present study is whether the type of imagery studied is domain-specific, namely images of an abacus, and that the imagery abilities of experts are very high in contrast with those of non-experts. More investigation in this area is necessary in order to understand to what extent the results from studies comparing imagery experts and non-experts can be applied more generally.

However, it is a debatable issue in itself whether or not general imagery processes exist in our information-processing system. This issue has been discussed by Hishitani (1988). For example, Hatano and Osawa (1983), and Hishitani (1980) showed that expert abacus operators performed well on tasks in which abacus images could be used, but their imagery abilities did not transfer to objects other than the abacus. Those findings indicate that the experts' imagery abilities are domain-specific. Hunter (1986) has drawn attention to domain specificity in another cognitive skill—namely, memory expertise. Therefore it seems that the domain-specific characteristics of expertise are to be found in any field, probably because many parts or domains of our cognitive processes are formed by our past experiences. Even if general imagery processes exist in our information-processing system, they must be specialized and optimized to function most effectively for each person on the basis of his or her past experience, and imagery processes must, to a greater or lesser degree, assume a domain-specific aspect. Therefore, the study of extremely domain-specific imagery, such as that of an abacus, helps us to understand more deeply the functions and nature of imagery processes in general. For this reason we should study how novices become imagery experts. As pointed out by Hishitani (1988), our imagery processes can vary and develop in a flexible fashion. Therefore, comparative studies of experts and non-experts, or longitudinal studies of developing experts, will provide an effective paradigm for understanding the flexibility and variety of the imagery system.

ACKNOWLEDGEMENT

I would like to thank Professor D. F. Marks of the Middlesex Polytechnic for his helpful advice on this article; Dr E. R. Valentine and Mr J. D. Valentine of the Psychology Department, Royal Holloway and Bedford New College for their invaluable help during its preparation; and Professor M. W. Eysenck of the same department who provided a stimulating environment for the writing of this paper during a sabbatical leave. I also thank anonymous reviewers for their valuable comments on an earlier version of the manuscript.

REFERENCES

Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, **85**, 249–277.

Baddeley, A. D., Thomson, N. and Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, **14**, 575–589.

Hatano, G. and Osawa, K. (1983). Digit memory of grand experts in abacus-derived mental calculation. *Cognition*, **15**, 95–110.

Hebb, D. O. (1966). *A textbook of psychology*, 2nd edn. Philadelphia: Saunders.

Hishitani, S. (1980). Effects of the relationship between material-characteristics and coding difference on memory. *Japanese Journal of Educational Psychology*, **28**, 251–255.

Hishitani, S. (1983). Effects of meaningfulness of letter-string on formation and access of letter-string image. Paper presented to the 25th Annual Meeting of the Japanese Association of Educational Psychology, Kumamoto University.

Hishitani, S. (1987). Imagery processing of expert abacus operators. *Studies in Childhood Education* (Seinan Gakuin University), **12**, 27–41.

Hishitani, S. (1988). The usefulness of studies of imagery experts in imagery research. Manuscript submitted for publication.

Hishitani, S. and Yumino, K. (1976). A device for multi-purpose latency measurement system. *Japanese Journal of Psychology*, **47**, 40–44.

Hunter, I. M. (1986). Exceptional memory skill. In A. Gellatly (Ed.), *The skillful mind*. Milton-Keynes: Open University Press, pp. 76–86.

Kristofferson, M. W. (1972). Effects of practice on character-classification performance. *Canadian Journal of Psychology*, **26**, 540–560.

Marks, D. F. (1983). In defense of imagery questionnaires. *Scandinavian Journal of Psychology*, **24**, 243–246.

Miyazaki, K. (1983). Imagery researches in cognitive psychology. In K. Mizusima and T. Uesugi (Eds), *Psychology of imagery*. Tokyo: Seishin Shobo, pp. 158–191.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin*, **80**, 1–24.

Pylyshyn, Z. W. (1979). The rate of 'mental rotation' of images: a test of a holistic analogue hypothesis. *Memory and Cognition*, **7**, 19–28.

Pylyshyn, Z. W. (1981). The imagery debate: analogue media versus tacit knowledge. *Psychological Review*, **88**, 16–45.

Seamon, J. G. (1972). Imagery codes and human information retrieval. *Journal of Experimental Psychology*, **96**. 468–470.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, **153**, 652–654.

Weber, R. J. and Harnish, R. (1974). Visual imagery for words: the Hebb test. *Journal of Experimental Psychology*, **102**, 409–414.

Yuille, J. C. (1983). The crisis in theories of mental imagery. In J. C. Yuille (Ed.), *Imagery, memory and cognition: essays in honor of Allan Paivio*. London: Lawrence Erlbaum Associates, pp. 263–284.