Iconic store and partial report

SIU L. CHOW
University of Wollongong, Wollongong, New South Wales, Australia[‡]

The iconic store has recently been challenged on the grounds that data in its favor may have resulted from some procedural artifacts. The display-instruction compatibility and perceptual grouping hypotheses were reexamined in two experiments with the partial-report paradigm. When care was taken to rectify some procedural problems found in Merikle's (1980) study, it was established that the iconic store (as a hypothetical mechanism) can still be validly entertained. This report demonstrates one important procedural point in studying the iconic store with the partial-report task, namely, that subjects must be given more than token training on the partial-report task.

An example is useful in illustrating the meaning of "iconic store." A subject is given a five-letter array for 50 msec and is asked to recall as many letters as possible. In the absence of any masking (forward or backward), the subject can do the task quite easily. Of interest is the interval between the onset of the stimulus array and the time at which the subject finishes responding. The typical estimate is about 1.5 sec. Because the stimulus is present for only .05 sec, the subject is responding to something that is physically absent much of the time. This phenomenon suggests that information about the stimulus array must have been preserved long enough for the subject to respond. The crucial question is how the mechanism underlying such a feat should be characterized.

There are currently two schools of thought. Some investigators (e.g., Holding, 1970, 1971, 1973, 1975; Merikle, 1980), although they accept the phenomenon that a briefly shown stimulus may appear to persist longer than the actual stimulus duration, suggest that the inferred preservation of information can easily be accounted for in terms of the physical properties of the visual system. Moreover, such a hypothetical mechanism is deemed theoretically irrelevant because it does not have any ecological validity (Haber, 1983). The term "visual persistence" is used to stand for the underlying mechanism (Merikle, 1980). Relevant to this position is the finding that the locus of visual persistence may be at the retinal level (Haber, 1983; Sakitt, 1975, 1976a, 1976b; Sakitt & Appelman, 1978; Sakitt & Long, 1978, 1979). The term "visual persistence" seems to be compatible with either "neural persistence" or "visible persistence" as recognized by Coltheart (1980).

Other investigators (notably Coltheart, 1975, 1980, 1983; Coltheart, Lea, & Thompson, 1974; Haber, 1971; Neisser, 1967) are willing to ascribe to the underlying mechanism some properties not definable in terms of the physical properties of the visual system qua

a physical system (see particularly Coltheart, 1980). The mechanism is called "iconic memory" (Neisser, 1967) or "iconic store." This view corresponds to Coltheart's (1980) notion of "information persistence."

Much of the dispute can be attributed to some fundamental disagreements with regard to the procedures used to establish the notion of "iconic store" (Coltheart, 1980) as well as the converging operations (Garner, Hake, & Eriksen, 1956) used to substantiate some of the hypothetical properties of the iconic store. The properties in question are its relatively (with respect to the immediate memory span; Miller, 1956) large storage capacity and the allegedly "sensory" (in the sense of being precategorical à la Crowder & Morton, 1969) nature of information representation at the iconic level.

Evidence in support of the iconic store comes generally from the facts that a subject's partial-report performance is superior to his/her whole-report performance and that such a superiority declines as the delay of the probe tone in the partial-report task increases (Sperling, 1960). These observations will jointly be called the "large-capacity finding" subsequently.

Apart from the dimension of spatial position, other effective partial-report selection criteria have been identified. They are color (Clark, 1969; Turvey & Kravetz, 1970; von Wright, 1968), achromatic color (von Wright, 1968), size (von Wright, 1968), and shape (Turvey & Kravetz, 1970; von Wright, 1968). Category membership (i.e., letters or digits) is not an effective partial-report selection criterion (Sperling, 1960; von Wright, 1968). In other words, the large-capacity finding is obtained only when the partial-report selection criterion belongs to a physical ("sensory") dimension. This observation is called the "basic sensory finding" here.

The first objective of this report is to consider an interesting alternative interpretation of the "large-capacity
finding" suggested recently by Merikle (1980). The second objective is to reexamine the basic sensory finding.

EXPERIMENT 1

In order to appreciate Merikle's (1980) argument, it helps to look at Figure 1, in which is shown a "row display" (top), a "column display" (middle), and a "square display" (bottom). In terms of the Gestalt principle of spatial proximity, three rows or three columns may readily be seen when a row or a column display is shown. A square display is so called because the perceptual gestalt is a square.

Following Kahneman (1973), Merikle (1980) argued that a good gestalt would draw a perceiver's attention more readily. Consequently, an observer would favor rows and columns when given row and column displays, respectively. The same observer should

favor neither the rows nor the columns when given a square display. This component of Merikle's argument will be called the "perceptual grouping" hypothesis.

Whenever the large-capacity finding is obtained with Sperling's (1960) partial-report task, the partial-report selection criterion is exclusively spatial location. Merikle (1980) suggested that the perceptual grouping of the items (i.e., a row gestalt) is thus confounded with the recall (by row) instruction. Hence, partial recall by row is favored. This component of Merikle's argument will be called the "display-instruction compatibility" hypothesis.

```
B G S

W X D

N C P


X   M   G
D   W   H
L   N   R


F W N
C X M
T S Z
```
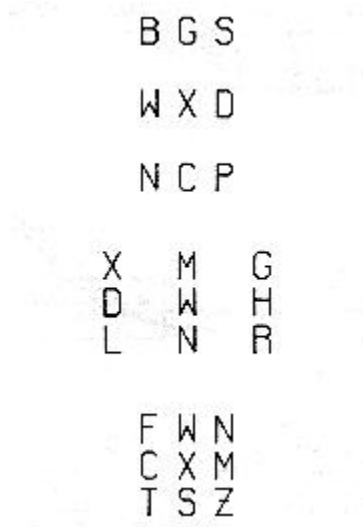
Figure 1. Three types of stimulus cards used in Experiment 1. "Row," "column," and "square" cards are represented in the top, middle, and bottom panels, respectively.

Using row and column displays, Merikle (1980) instructed different groups of subjects to recall either by row or by column, when given the partial-report task. The "recall by row" group was effectively given compatible and incompatible display-instruction relationships when shown row and column displays, respectively. By the same token, the "recall by column" group had incompatible and compatible tasks when given row and column displays, respectively.

Merikle (1980) found that partial-report performance was superior to whole-report performance at both levels of compatibility. However, the superiority was more pronounced when the display and instruction were compatible. These observations were obtained when perceptual grouping was achieved by spatial proximity (Merikle, 1980, Experiment 1) or by similarity in terms of brightness (Merikle, 1980, Experiment 2). It was concluded that the display-instruction compatibility definitely contributed to the

large-capacity finding. This conclusion is counter to the assumption that partial-report performance is superior because there are more items available in the iconic store than in the short-term store.

Despite the attractiveness of the display-instruction compatibility hypothesis, the empirical study in its support is not convincing because the partial-report cue was presented either *before* or *simultaneously with* the partial-report display in Merikle's (1980) study. This procedural feature violates the "experimental criterion for iconic memory" recognized by Coltheart (1980, p. 221), who insisted that the partial-report cue must be presented *at* or *after* display offset.

The methodological assumption underlying the partial-report task is that the subjects first select the relevant subset of items on the basis of a certain dimension before any further processing is carried out, a process called "select-then-process" (or "selective readout" by Coltheart, 1980, p. 187). However, the subjects seldom, if ever, have to engage in such selective readout under so fast a presentation condition. They naturally would treat the partial-report task as the attention-span task plus some editing before responding when they are inexperienced in the partial-report task, a strategy called "nonselective readout" by Coltheart (1980). Seen in this light, the fact that Merikle (1980) has given his subjects only 12 practice trials on the partial-report task is very unsatisfactory.

Experiment 1 had two objectives: (1) to test Merikle's (1980) interesting display-instruction compatibility hypothesis when the methodological assumption of the partial-report task is observed, and (2) to assess the importance of practice on the partial-report task to the largecapacity finding.

Method

Subjects. Twelve subjects were used in the experiment. All subjects were undergraduate students at the University of Wollongong. An honorarium was given to the subjects for participating in the experiment.
Apparatus. A Gerbrands four-channel tachistoscope with a 61159 tachistoscope interface was used to present the visual materials to the subjects. The tachistoscope was also connected to a Gerbrands 300 Series timer, an Exact Model 506 sweep function tone generator, and two automatic card changers. A pair of earphones was used to present the probe tones.

Table 1

Arrangement of Experimental Sessions in Experiment 1

Group	Subject	1	2	3	4	5	6	7	8	9
G1[a],G2[b] C	1	M[c]	M	C[d]	IC[e]	IR[f]	WR[g]	IR	IC	C
G1 ,G2 IR	2	M	M	IC	IR	C	WR	IC	C	IR
G1 ,G2 IC	3	M	M	IR	C	IC	WR	C	IR	IC

[a]Gl = Group l; given instruction to recall by row. [b] G2 = Group 2; given instruction to recall by column. [c]M = trained with "mixed" cards. [d]C = compatible display instruction (i. e., to recall by row and by column when given row and column cards, respectively). [e]IC = incompatible display instruction (i.e., to recall by row and by column when given column and row cards, respectively). [f]IR = irrelevant display instruction (i.e., to recall by row or by column when given square cams). [g]WR = whole report.

Materials. The stimulus ensemble consisted of the following letters of the alphabet in uppercase: B, C, D, F, G, H, J, K, L, M, N, P, R, S, T, V, W, X, and Z. Nine of them were selected randomly without replacement, for use on any one trial.

Eight sets of 100 cards were prepared with a plotter driven by a UNIVAC 1160 computer. Nine letters of the ensemble were plotted on each card. The eight sets of cards consisted of two "row" sets, two "column" sets, two "square" sets (see Figure 1 for some examples), and two "mixed" sets. A set of "mixed" cards was made up of 32 "row," 32 "column," and 32 "square" cards.

A "row" card subtended a visual angle of 1 ° 59' in width and 3 ° 33' in height. A "column" card subtended a visual angle of 3° 33' in width and 1° 59' in height. A square display subtended a visual angle of 1 ° 59' in width and in height.

Three tones were used in the experiment. They were 5000 Hz (high tone), 1000 Hz (medium tone), and 300 Hz (low tone).

Design. A 2 x 3 x 4 factorial design with repeated measures on the latter two factors was used in Experiment 1. The first factor was recall instruction (i.e., recall by row or by column), and it was a between-groups factor. The second factor was display-instruction compatibility. Its three levels were compatible, incompatible, and irrelevant. The third factor was the interstimulus interval (ISI), that is, the interval between the offset of the stimulus and the onset of the probe tone. The four levels of ISI were 0, 150, 500, and 1,000 msec. Display-instruction compatibility and ISI interval were within-subject factors.

Procedure. The 12 subjects were divided randomly into two groups (Group 1 and Group 2) of 6 each. The subjects in Group 1 were instructed to always recall by row and the subjects in Group 2 were to recall by column when given the partial-report task. They were tested (across nine sessions) in the order (from left to right) described in Rows 1, 2, and 3, respectively, in Table 1. All subjects were given the whole-report task in Session 6 and the partial-report task in all other sessions. This testing sequence was adopted in order to minimize any effect due to the order of testing.

The sequence of events involved in a trial under the whole-report condition is depicted in the upper panel of Figure 2. Upon hearing "Ready" from the experimenter, the subject was shown a "+" (which served as the fixation point) for 1,500 msec. After .75 sec, a nine-letter display was shown for 50 msec. The subject was instructed to recall (as soon as possible) as many letters as possible.
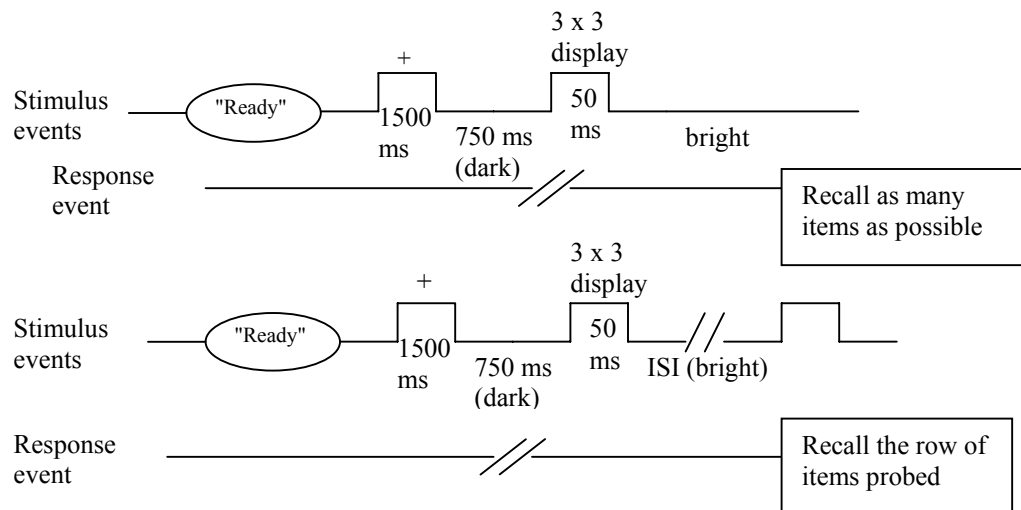

Figure 2. Sequence of events in the whole-report and the partial-report tasks.

The event sequence constituting a partial-report trial is depicted in the lower panel of Figure 2. Again, "Ready" from the experimenter indicated the onset of a trial. A "+" was first shown for 1,500 msec. After an interval of 750 msec, a nine-letter display was shown for 50 msec. One of three equally probable tones followed the nine-letter display at various delays. The subject was instructed to recall the top row (or the leftmost column) if a high tone was presented, the middle row (or the central column) if a medium tone was presented, and the bottom row (or the rightmost column) if the tone was a low tone.

There were 100 trials in every session for both the partial-report and the whole-report tasks. Data from the first 4 trials in every session were discarded. The 96 experimental trials were divided into four blocks of 24 trials each.

For the partial-report task, the ISI value was held constant within a block. The ISI values used in the first block were also used in the four warm-up trials. The order in which the four ISI values were tested in a session was determined randomly for individual subjects.

Within a block of 24 experimental trials, the three probe tones were used equally often. The choice of probe tones for the four warm-up trials was made at random and did not have any bearing on the experimental trials. The specific order in which the three rows of the partial-report display was probed was determined randomly within a block.

The sequence of events (as depicted in Figure 2) was thoroughly explained to the subjects. They were given a thorough training on discriminating among the three probe tones. The partial-report task did not start until the subjects could identify 30 tones (10 high, 10 medium, and 10 low, randomly arranged) without any error. Furthermore, the subjects were always tested on tone discrimination before any subsequent sessions. Whenever necessary, further practice on tone discrimination was provided before any session began.

The subjects were told that the purpose of the experiment was to establish how fast they could process visual information. The subjects were tested individually and were instructed to recall the items in the positions in which they had been presented.

Results

The subjects' performance on the partial-report and the whole-report tasks were tabulated according to the free-recall and the position-correct criteria. The two scoring criteria produced the same pattern of results. However, the free-recall procedure gave a higher absolute level than the position-correct criterion. For this report, only data scored with the free-recall criterion were considered.

Because Sessions 1 and 2 were training sessions (on the partial-report task), data collected in those sessions were not used in the analysis related to the display-instruction compatibility hypothesis. Since partial report was tested in two sessions at every compatibility level, data from the two occasions were averaged. Subsequently, the subjects' partial-report performance was subjected to a 2 (instruction group) x 3 (display-instruction compatibility) x 4 (ISI) ANOVA with repeated measures on the latter two factors. Data are shown in Figure 3.

Figure 3. Subjects' partial-report performance as a function of recall instruction, display-instruction compatibility, and ISI in Experiment 1.
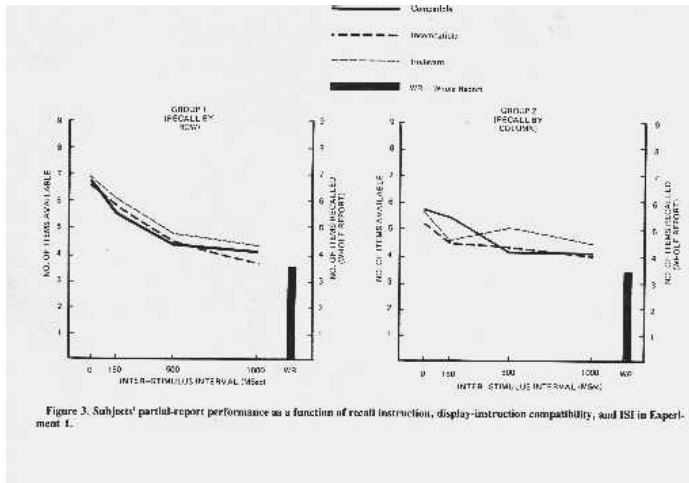
Figure 3. Subjects' partial-report performance as a function of recall instruction, display-instruction compatibility, and ISI in Experiment 1.

The ISI factor was significant [$F(3,30) = 29.00$, $p < .05$], as was the ISI x instruction group interaction [$F(3,30) = 3.68$, $p < .05$]. None of the other factors or interactions was significant.

To ascertain whether the basic large-capacity finding was obtained, the subjects' partial-report performance at the various levels of ISI was compared with their whole-report performance. For these comparisons, the subjects' whole-report performance was treated as the control ISI level. Data were subjected to an overall 2 (instruction group) x 3 (display-instruction compatibility) x 5 (four ISI levels and whole report) factorial ANOVA with repeated measures on the latter two factors. (For this analysis, only the ISI factor and all interactions involving ISI are relevant.)

The ISI factor was significant [$F(4,40) = 45.67$, $p < .05$]. Also significant was the ISI x instruction group interaction [$F(4,40) = 2.71$, $p < .05$]. None of the other interactions involving ISI was significant. Subsequently, the Newman-Keuls multiple comparison test was used to compare each of the four ISI levels with the whole-report condition separately for the two groups of subjects. As can be seen from Table 2, partial-report performance was superior to whole-report performance only at the 0- and 150-msec ISIs. This is true for both the "row recall" and the "column recall" groups.

As can be seen from Table 1, every subject was tested twice under the practice, compatible, incompatible, and irrelevant display-instruction conditions. The subjects' partial-report performance at the 0-msec ISI in the two practice and the two test sessions for each of the three compatibility conditions is shown in Figure 4. The obvious feature of Figure 4 is that the large-capacity finding is obtained only after the subjects have 96 practice trials on the partial-report task.

To validate this interpretation, data from the subjects' partial-report performance and their whole-report performance were subjected to a 2 (instruction group) x 3 (compatibility) x

5 (sessions) ANOVA with repeated measures on the latter two factors. (For this analysis, only the session factor and all interactions involving session are meaningful.) As can be seen from Table 3, partial-report superiority was not observed in Session 1. It was only in subsequent sessions that partial-report superiority was found.

Table 2

The Newman-Keuls Multiple Comparison Test Between Subject' Partial-Report Performance of Various Levels of ISI and Whole Report (Experiment 1)

		WR^a	1000^b	500^b	150^b	0^b
		Recall by Row Group				
		3.58	4.01	4.65	5.78	6.78
WR	3.58		0.43	1.07	2.20^*	3.20^*
1000	4.01			0.64	1.77^*	2.77^*
500	4.65				1.13^*	2.12^*
150	5.78					
		Recall by Column Group				
		3.51	3.98	4.21	5.11	5.47
WR	3.51		0.47	0.70	1.60^*	1.96^*
1000	3.98			0.23	1.13	1.49^*
500	4.21				0.99	1.26
150	5.11					0.36

[a]WR=Whole report; [b]1000=ISI value

[*]Significant at the 0.5 level by the Newman-Keuls multiple comparision test. The critical values for two-, three- and five-step diffeences are 1.09, 1.32, 1.45, and 1.55, respectively.
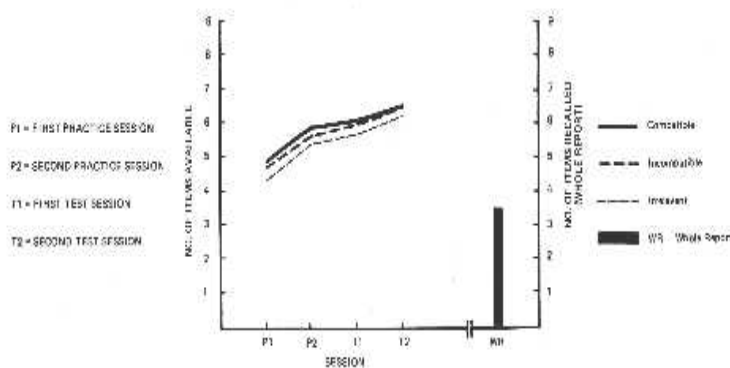

Figure 4. Subjects' partial-report performance at the 0-msec ISI as a function of session of testing in Experiment 1.

Figure 4. Subject's partial-report performance at the 0-msec ISI as a function of session of testing in Experiment 1.

Table 3

The Newman-Keuls multiple comparison for subjects' partial-report performance at the 0-msec ISI in various sessions and their whole-report performance (Experiment 1)

		Session				
		WR[a]	P1[b]	P2[c]	T1[d]	T2[e]
		3.52	4.64	5.62	5.85	6.38
WR	3.52		1.12	2.10*	2.33*	2.86*
P1	4.64			0.98	1.21	1.74
P2	5.62				0.23	0.76
T1	5.86					0.53

[a]WR = whole report; [b]P1 = first practice session; [c]P2 = second practice session [d]T1 = first test session; [e]T2 = second test session; *Significant at 0.05 level by the Newman-Keuls multiple comparison test. The critical values for two-, three-, four-, and five-step differences are 1.41, 1.70, 1.88 and 1.99, respectively.

Discussion

Results from this experiment differ from those of Merikle's (1980) study as a result of an important procedural difference. When the partial-report cue was presented *before* or *simultaneously* with stimulus *onset,* Merikle obtained data consistent with the display-instruction compatibility hypothesis. When the partial-report cue was presented (in accordance with the methodological assumption of the partial-report task; see Coltheart, 1980) *at* or *after* stimulus *offset,* the compatibility hypothesis was not supported. It can be suggested that the display-instruction compatibility hypothesis is incorrect.

It is observed here that the basic large-capacity finding is obtained only when subjects have 96 trials of practice on the partial-report task. This procedural feature should no longer be ignored in future applications of the partial-report task.

EXPERIMENT 2

Why is the basic sensory fording observed by some investigators (e.g., Sperling, 1960; Turvey & Kravetz, 1970; von Wright, 1968), but not by others (e.g., Dick, 1969, 1971, 1974; Merikle, 1980)? Although we have rejected the display-instruction compatibility hypothesis as an explanation of the large-capacity finding, it remains to be seen whether the perceptual grouping hypothesis can account for the basic sensory finding.

Merikle (1980) gave his subjects a circular array of eight alphanumeric items for 50 msec. Four of the eight items were letters, and four were digits. On half of the occasions,

items belonging to a particular category also shared the same brightness, a condition called "correlated physical dimension present" (Merikle, 1980, p. 287). The correlated physical dimension was absent on the other half of the occasions.

Subjects were instructed to recall letters only, digits only, or as many items as possible (regardless of category membership). The finding of interest is the observation that partial report was superior to whole report, despite the fact that the selection criterion for partial report was category information, both in the presence and in the absence of the correlated physical dimension. The conclusion was drawn that selection by category under the partial-report situation was possible. Consequently, "iconic memory might not be precategorical in nature" (Merikle, 1980, p. 288).

The validity of Merikle's (1980) conclusion can be questioned because the partial-report cue was given to the subjects before stimulus presentation. Experiment 2 was designed to test the perceptual grouping explanation of the basic sensory findings.

Method

Subjects. Six of the 12 subjects in Experiment I were given an honorarium to participate in Experiment 2.

Apparatus. The apparatus was the same as that used in Experiment 1.

Materials. There were three kinds of stimulus items: the ensemble of letters of the alphabet used in Experiment 1, the digits 2 through 9, and some common symbols. The ensemble of symbols was: $, ?, #, Q, %, and *. Three items were selected randomly, without replacement, from each of the three kinds of stimuli on every trial.

Three types of stimulus displays were prepared for Experiment 2. Examples of "square-good gestalt," "square-poor gestalt," and "column" cards can be found in Figure 5. A "square-good gestalt" card is one in which the three items from a particular category (e.g., letters) form a triangle in the display, whereas the rest of the six items from the other two categories are distributed randomly in the display. A "square-good gestalt" card subtends a visual angle of 2° 14' in width as well as in height
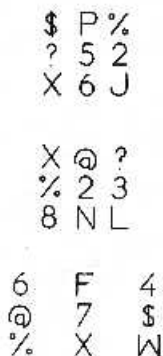

Figure 5. Three types of stimuli used in Experiment 2. A "square-good gestalt," a "square-poor gestalt," and a column card are shown in the top, middle, and bottom panels, respectively.

A "square-poor gestalt" card is one in which the nine items (three items from each of the three categories) are distributed randomly such that no good gestalt is discernible in terms of items from any one category. It subtends a visual angle of 2° 14' in width as well as in-height. This category of cards served, in fact, two purposes. They were used as stimulus cards without any good perceptual gestalt when the partial-report selection criterion was category information. These cards also served as good spatial cards when the partial-report selection criterion was spatial information.

A "column" card is one in which the nine items selected are cast in an arrangement that readily suggests three columns of three items each. Items from any category do not form any good discernible gestalt. It subtends a visual angle of 3° 57' in width and 2° 14' in height.

There were two sets of 96 cards each for every one of the three kinds of displays. From these six basic stimulus ensembles, six sets of experimental cards were formed. One such set was used in training subjects, one for testing subjects on the whole-report task, two for partial reports when category information was the selection criterion, and two for partial reports when spatial information was the selection criterion. Details of the six experimental sets of cards are as follows:

(1) Practice cards-- his deck of cards was made up of 24 square-good gestalt cards, 48 square-poor gestalt cards, and 24 column cards. Among the 24 square-good gestalt cards, 8 were chosen from letters, 8 from digits, and 8 from symbols.

(2) *Whole-report cards-- This* deck of cards was made up in the same way as the deck used in training the subjects.

(3) Partial-report by category cards-- Each of the two decks of cards consisted of 48 square-good gestalt and 48 square-poor gestalt cards. Among the former, there were equal numbers of letter, digit, and symbol cards.

(4) Partial-report *by* spatial location cards-- Each of the two decks of cards was made up of 48 square-poor gestalt and 48 column cards.

Other aspects of the materials used were the same as those in Experiment 1.

Design. A 2 (recall by row or category) x 2 (good or bad perceptual grouping) x 4 (ISI) factorial design with repeated measures on all factors was used. Whereas both the perceptual and the ISI factors were within-session factors, the recall-instruction factor was tested across sessions.

Procedure. In order to minimize any effect of the order of testing, the six subjects were tested in the order (from left to right) depicted in Table 4. The six subjects were assigned randomly to the three rows of Table 4. All subjects were given the whole-report task in Session 4. Because the subjects were well practiced on partial report with a spatial selection criterion, Session 1 was used for training the subjects on partial report with a category selection criterion. Two sessions were devoted to each of the two kinds of selection criteria. When the selection criterion was category information, the high, medium, and low tones signified letters, digits, and symbols, respectively. Other aspects of the procedure were identical to those employed in Experiment 1.

Table 4
Arrangement of Experimental Sessions in Experiment 2

Number of Subjects	Session					
	1	2	3	4	5	6
2	PR[a]	S[b]	C[c]	WR[d]	C	S
2	PR	C	S	WR	S	C
2	PR	S	C	WR	S	C

[a]PR = practice session. [b]S = spatial *location as partial-report cue.* [c]C =category *information as partial-report cue.* [d]WR = whole report.

Results

Data from the practice session were excluded when subjects' partial-report performance was examined to see whether the perceptual grouping hypothesis could account for the basic sensory fording. The analysis was a 2 (selection criterion) x 2 (perceptual grouping) x 4 (ISI) factorial ANOVA with repeated measures on all of the factors.

The overall ANOVA revealed a significant ISI effect [$F(3,15) = 36.21$, $p < .05$], a significant ISI x selection criterion interaction [$F(3,15) = 7.97$, $p < .05$], and a significant ISI x perceptual grouping interaction [$F(3,15) = 3.88$, $p < .05$]. However, the ISI x instruction x perceptual grouping three-way interaction (which was necessary for the perceptual grouping hypothesis) was not significant (see Figure 6).
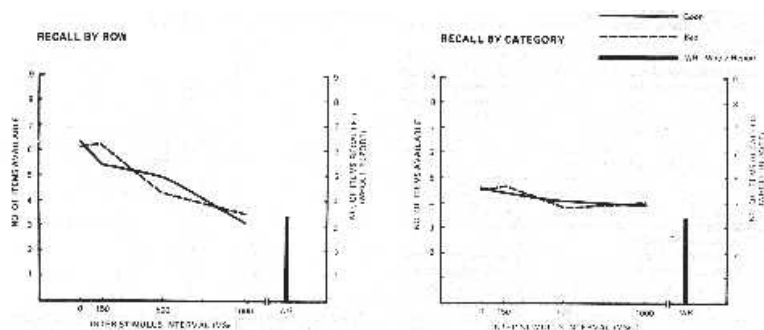

Figure 6. Subjects' partial-report data as a function of selection criterion, perceptual grouping, and ISI in Experiment 2.

Figure 6. Subjects' partial-report data as a function of selection criterion, perceptual grouping, and ISI in Experiment 2.

In order to assess whether the basic sensory finding had been replicated, it was necessary to ascertain whether the large-capacity finding was obtained separately for the selection-by-row and the selection-by-category conditions. For this purpose, the subjects' whole-report performance was treated as the control ISI level. Their partial-report data in the practice session were excluded. The subjects' performance on both tasks was then subjected to a 2 (recall instruction) x 2 (perceptual grouping) x 5 (ISI and whole report) factorial ANOVA with repeated measures on all factors.

The mean square error term for the ISI x recall instruction interaction was used to derive the standard error of the difference for the two series of planned t-statistic comparison tabulated in Table 5 (Wiper, 1962). As can be seen from Table 5, partial report is superior to whole report at delays shorter than 1 sec when selection is by row. When selection is by category, partial report is no better than whole report even when ISI is 0 msec.

Table 5
Planned Comparisons Between Partial Report and Whole Report
at Various ISI Levels (in Milliseconds) for the Spatial
Selection (Left Panel) and the Category Selection
(Right Panel) Conditions in Experiment 2

	Selection by Row					Selection by Column				
	0	150	500	1000		0	150	500	1000	WR[a]
PR[b]	6.29	5.93	5.07	3.35		4.55	4.49	4.32	3.90	3.46
>WR[c]	2.83	2.47	1.61	-0.11		1.09	1.04	0.44	0.44	
t	5.55[*]	4.84[*]	3.16[*]	-0.22		2.14	2.04	0.86	0.86	

[a]WR = whole *report (mean number of items recalled).* [b] *PR = partial report (mean number of items available).* [c] *> WR = magnitude of the partial-report superiority (partial report minus whole report).* [*]*Significant at the .05 level for the entire set of comparisons. Critical t(with k = 5; df for MSe = 20; n = 6) = 2.30.*

Discussion

A comment on the top panel of Figure 5 is helpful here. It is obvious that it may not be easy to perceive the triangle gestalt formed by the three letters. However, to appreciate Experiment 2, it is necessary to distinguish between what a theoretical position prescribes one should be able to do and what one actually can do.

The perceptual grouping hypothesis, together with the assumption that category information is available as a result of a brief stimulus exposure, prescribes that our perceptual system is capable of picking up the triangle in the top panel of Figure 5 so readily that a partial-report superiority by category should be observed. Whether the

triangle gestalt can readily be seen or not is a separate issue; hence, it is irrelevant to the validity of Experiment 2.

Can the perceptual grouping hypothesis account for the basic sensory fording? The perceptual grouping hypothesis prescribes that when the items of the probed category form a good gestalt (i.e., a triangle in Experiment 2), partial report by category should be better than when no such good gestalt is available. This theoretical expectation of the perceptual grouping hypothesis is not met. Moreover, the large-capacity finding is obtained only when selection is by row, not by category. In other words, results from Experiment 2 cast doubts on the suggestion that good perceptual organization within a stimulus may inflate subjects' partial-report performance in a typical partial-report experiment.

GENERAL DISCUSSION

Merikle (1980) proposed two alternative explanations for the large-capacity and the basic sensory findings, namely, the display-instruction compatility and the perceptual grouping hypotheses. Both of these hypotheses are primarily accounts of some procedural artifacts that might be operative in Sperling's (1960) partial-report paradigm.

The general thrust of Merikle's (1980) argument is that, if the artifacts are removed, both the large-capacity and the basic sensory findings should not be obtained. Consequently, the iconic store, as a hypothetical mechanism, would have to be treated very differently. Empirical data obtained from a series of experiments devised by Merikle to substantiate this argument led to the rejection of an underlying mechanism responsible for what Coltheart (1980) called "information persistence" in favor of a multichannel view of neural, or visible, persistence.

However, a necessary procedural requirement of the partial-report task was not observed in Merikle's (1980) study. There is no empirical support for either the display-instruction compatibility hypothesis or the perceptual grouping hypothesis when the partial-report cue is presented at or after stimulus offset.

Merikle (1980) also suggested that the absence of any partial-report superiority when category information is the selection dimension may be attributable to the fact that a wrong whole-report baseline was used. Merikle noted that, when whole report was made uncertain in Dick's (1969) study, category selection criterion is effective in bringing about partial-report superiority. It is suggested that the proper whole-report baseline should be the subject's whole-report performance when there is uncertainty about when to carry out the whole report (Merikle, 1980).

The partial-report task is a very unusual and difficult task. It is sometimes necessary to spend the first quarter of an hour of the first session in assuring the subjects that the partial-report task is not an impossible one. Partial-report data collected from subjects

without any reasonable amount of training are empirically suspect. Yet, Dick (1969) did not give his subjects any training on the partial-report task. This was particularly serious in Dick's study in which subjects were given a small number of practice trials on the whole-report task before being tested on the partial-report task. This methodological feature in Dick's study gives credence to Coltheart's (1980) suggestion that Dick's subjects might have been recalling from a more durable storage mechanism. At the same time, it means that Dick's study cannot be used to support the hypothesis that a wrong baseline was used in assessing the selection of category information after a brief stimulus exposure. Moreover, there is also a conceptual problem with such a suggestion. Why does partial report by category membership require a different baseline from partial report by spatial location?

Admittedly, the iconic store is assumed to be a mechanism found in everyday information-processing activities when the visual modality is involved. Yet, it has been emphasized repeatedly that the partial-report task used in studying the iconic store is an artificial one. A metatheoretical issue arises. How valid is it to study a behavioral phenomenon with an artificial method? A related issue is: How relevant is our understanding of the iconic store when such an understanding is achieved with ecologically irrelevant methods? These are the metatheoretical questions recently raised by Haber (1983) in his objection to the iconic store.

It must be noted that some investigators have found the partial-report task to be ecologically relevant (e.g., Coltheart, 1983). Even if the question about ecological validity is put aside, the iconic store can still be defended on metatheoretical grounds against Haber's (1983) metatheoretical objection. To begin with, it is not true that, when we study a phenomenon, we must use a task that mimics the phenomenon we are studying (Chow, in press; Manicas & Secord, 1983). Moreover, Haber's point of view "runs counter to scientific wisdom and practice developed over the past few milennia" (Loftus, 1983, p. 28). For example, much of what we know about gravity is gained by studying objects falling in near vacuums. Hence, it is important to realize that an obvious use of some phenomenon in the real world does not traditionally constitute a necessary condition for studying that phenomenon in the scientific laboratory. (Loftus, 1983, p. 28)

To conclude, the display-instruction compatibility and the perceptual grouping hypotheses fail to withstand the falsification attempts provided by the present two experiments. The methodological lesson to be learned is that, because the partial-report task is an unusual one, subjects must be given more than token opportunity to adopt the "select-then-process" strategy.

REFERENCES

CLARK, S. E. (1969). Retrieval of color information from preperceptual memory. *Journal of Experimental Psychology, 82,* 263-266.

COLTHEART, M. (1975). Iconic memory: A reply to Professor Holding. *Memory & Cognition, 3,* 42-48.

COLTHEART, M. (1980). Iconic memory and visible persistence. *Perception & Psychophysics, 27,* 183-228.

COLTHEART, M. (1983). Ecological necessity of iconic memory. *Behavioral and Brain Sciences, 6,* 17-18.

COLTHEART, M., LEA, C. D., & THOMPSON, K. (1974). In defense of iconic memory. *Quarterly Journal of Experimental Psychology, 26,* 633-641.

CROWDER, R. G., & MORTON, J. (1969). Precategorical acoustic storage (PAS). *Perception & Psychophysics, 5,* 365-373.

DICK, A. O. (1969). Relations between the sensory register and shortterm storage in tachistoscopic recognition. *Journal of Experimental Psychology, 82,* 279-284.

DICK, A. O. (1971). On the problem of selection in short-term visual (iconic) memory. *Canadian Journal of Psychology, 25,* 250-263.

DICK, A. O. (1974). Iconic memory and its relation to perceptual processing and other memory mechanisms. *Perception & Psychophysics, 16,* 575-596.

GARNER, W. R., HAKE, H. W., & ERIKSEN, C. W. (1956). Operationism and the concept of perception. *Psychological Review, 63,* 149-159.

HABER, R. N. (1971). Where are the visions in visual perception? In S. J. Segal (Ed.), *Imagery: Current cognitive approaches.* New York: Academic Press.

HABER, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behavioral and Brain Sciences, 6, 1-11.*

HOLDING, D. H. (1970). Guessing behaviour and the Sperling store. *Quarterly Journal of Experimental Psychology, 22,* 248-256.

HOLDING, D. H. (1971). The amount seen in brief exposures. *Quarterly Journal of Experimental Psychology, 23.* 72-81.

HOLDING, D. H. (1973). Recognition tests of visual information storage. *British Journal of Psychology, 64,* 9-16.

HOLDING, D. H. (1975). Sensory store reconsidered. *Memory & Cognition, 3,* 31-41.

KAHNEMAN, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice-Hall.

LOFTUS, G. R. (1983). The continuing persistence of the icon. *Behavioral and Brain Sciences, 6,* 28.

MANICAS, P. T., & SECORD, P. F. (1983). Implications for psychology the new philosophy of science. *American Psychologist, 38* 399-413.

MERIKLE, P. M. (1980). Selection from visual persistence by perceptual groups and category membership. *Journal of Experimental Psychology: General, 3,* 279-295.

MILLER, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81-97.

NEISSER, U. (1967). *Cognitive psychology.* New York: AppletonCentury-Crofts.

SAKITT, B. (1975). Locus of short-term visual storage. *Science, 190,* 1318-1319.

SAKITT, B. (1976a). Iconic memory. *Psychological Review, 83,* 257-276.

SAKITT, B. (19766). Psychophysical correlates of photoreceptor activity. *Vision Research, 16,* 129-140.

SAKITT, B., & APPELMAN, J. B. (1978). The effects of memory load and the contrast of the rod signal on partial report superiority in a Sperling task. *Memory & Cognition, 6,* 562-567.

SAKITT, B., & LONG, G. M. (1978). Relative rod and cone contributions to iconic storage. *Perception & Psychophysics, 23,* 527-536.

SAKITT, B., & LONG, G. M. (1979). Spare the rod and spoil the icon. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 19-30.

SPERLING, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74(11,* Whole No. 498).

TURVEY, M. T., & KRAVETZ, S. (1970). Retrieval from iconic memory with shape as the selection criterion. *Perception & Psychophysics,* 8, 171-172.

VON WRIGHT, J. M. (1968). Selection in visual immediate memory. *Quarterly Journal of Experimental Psychology, 20,* 62-68.

WINER, B. J. (1962). *Statistical principles in experimental design.* New York: McGraw-Hill.

[‡] Now at the Department of Psychology, University of Regina, Regina, Saskatchewan, Canada S4S 0A2.