COGNITIVE PSYCHOLOGY **35,** 135–199 (1998)
ARTICLE NO. CG980678

On the Relations among Different Measures of Visible and Informational Persistence

Geoffrey R. Loftus

University of Washington

and

David E. Irwin

University of Illinois

We report research designed to accomplish two goals. We first consider the question, raised by Coltheart (1980) and others, of whether three measures of visible and informational persistence—performance in temporally integrating two successively presented stimuli, subjective rating of the degree to which two successively presented stimuli appear to constitute a single or a dual temporal event, and partial-report performance—all measure the same underlying mental entity. We answer this question using a superset of dissociation logic called *state-trace analysis* (Bamber, 1979), and within the context of a systematic empirical foundation consisting of seven closely related experiments. Our second goal is to extend and apply a theory to data acquired from our seven experiments and also to data reported by other investigators. This theory, which has been confirmed in a variety of paradigms (see Busey & Loftus, 1994) assumes that (1) the initial stages of the visual system act as a low-pass linear filter which operates on a stimulus temporal waveform to produce a *sensory response;* (2) instantaneous rate of acquiring information from the stimulus is jointly proportional to sensory-response magnitude and proportion of as-yet-to-be-acquired stimulus information; (3) partial-report performance is determined by total amount of acquired information; (4) the probability that two events are perceived as contemporaneous is determined by the temporal correlation of their respective information-acquisition rate functions (which is similar to a suggestion

We extend our heartfelt gratitude to Vince Di Lollo whose suggestions initiated this research, and whose support and advice throughout (particularly on the topic of brightness compensation) have been invaluable. We also thank Pete Dixon, John Duncan, George Sperling, Mark Reinitz, George Wolford, Steve Yantis, and an anonymous reviewer for very valuable suggestions on early drafts. Finally, we thank Jennifer Larys, Steve Messe, Catherine Shantz, David Van Dyk, and Steve Weede for their assistance in data collection. The research was supported by NIMH Grant MH41637 to Loftus and NSF Grants BNS 89-08699 and SBR93-09564 to Irwin.

Address reprint requests to Geoffrey R. Loftus at gloftus@u.washington.edu or David E. Irwin at dirwin@psych.uiuc.edu.

0010-0285/98 $25.00

by Dixon & Di Lollo, 1994); and (5) temporal integration is successful to the degree that the two temporal events are perceived as contemporaneous. This theory was highly successful in accounting for our and other investigators' temporal-integration and completeness-rating data, and was moderately successful in accounting for partial-report data. We discuss the degree to which our three persistence measures can be united within the context of our theory; we comment on the distinction between objective and subjective measures of visible persistence; and we address the decades-old question: ''What is persistence good for?''

It has been known for centuries that the perceptual experience of a briefly presented visual stimulus persists beyond the physical stimulus itself; indeed, the writings of Aristotle (384–322 B.C.) reportedly contain the first known reference to such persistence (Allen, 1926). This phenomenon has been studied intensively for more than three decades, and has been endowed with various names, including ''visual information store'' (Sperling, 1960), ''short-term visual storage'' (Haber, 1969), ''iconic memory'' (Neisser, 1967), and ''visible persistence'' (Coltheart, 1980). Consequently, most textbook models of visual information processing have assumed the existence of a very short-term visual memory that stores the contents of a visual scene for some period of time after its offset.

Over the years, there has been a good deal of debate about the perceptual and cognitive processes that are entailed in this general phenomenon[1] and what role, if any, it plays in perception and cognition.[2] Two empirical facts are clear, however: First, something that *looks like* the physical stimulus continues to be present for a brief time following stimulus offset. Second, *information can be acquired from the stimulus* for a brief period following stimulus offset in much the same way as it can be acquired while the stimulus is physically present. Following Coltheart (1980), we refer to the first phenomenon as *visible persistence,* and to the second as *informational persistence.*

Visible and informational persistence have been investigated using a variety of experimental tasks, some of which are enumerated and briefly described in Table 1.[3] Although generally measured by separate tasks, both visible and informational persistence were originally assumed to be two measures of the same internal entity: a single, visible, precategorical, high-capacity, quickly decaying memory that holds incoming visual stimulation for further processing by later components of the information processing system (e.g., Coltheart, Lea, & Thompson, 1974; Dick, 1974; Neisser, 1967;

[1] For reviews, see Coltheart (1980); Di Lollo and Dixon (1988); Dixon and Di Lollo (1994); Irwin and Yeomans (1986a); Irwin and Brown (1987); Long (1980); Mewhort, Campbell, Marchetti, and Campbell (1981); and Nisly and Wasserman (1989).

[2] See, for example, Di Lollo and Dixon (1992b); Haber (1983, 1985); Loftus, Johnson, and Shimamura (1985); Massaro and Loftus (1995); and Turvey (1977).

[3] The list of tasks and studies in Table 1 is not exhaustive. Furthermore, others (e.g., Long, 1985; Nisly & Wasserman, 1989) have proposed alternative organizational schemes.

TABLE 1
Different Tasks Used to Investigate Visible and Informational Persistence

	Visible persistence	
Task name	Description	Examples from literature
Synchrony judgment	Stimulus presented followed by a variable interval, followed by a synchrony signal (e.g, a tone). Observer adjusts signal such that signal corresponds with phenomenological disappearance of stimulus	Efron (1970a, b); Bowen Pola, & Matin (1974)
Temporal integration	A stimulus (e.g, a CVC trigram) is broken up into two spatial halves. Halves are presented in temporal succession; observer's task requires being able to see whole stimulus, i.e., both halves superimposed (e.g., naming the trigram)	Eriksen & Collins (1966); Di Lollo (1980)
Completeness ratings	Like a temporal integration task. Instead of performing an objective task, observer provides a rating of how temporally complete the conjunction of the two halves seemed to be.	Loftus & Hanna (1989)
	Informational Persistence	
Procedure name	Description	Examples from literature
Partial report	An array of items is presented. A probe presented at some point, generally following stimulus offset, indicates that some part of the array should be reported.	Sperling (1960); Averbach & Coriell (1961)
Mask/No mask comparison	Stimulus is presented either followed or not followed by a mask. Difference between masked and unmasked performance provides an indication of information acquired from the iconic image	Intraub (1980); Loftus, Johnson, & Shimamura (1985)

von Wright, 1972). Three examples illustrate this view. First, Eriksen and Collins (1967) introduced the temporal-integration technique (see Table 1), saying about it, ‘‘This task . . . permits the study of a possible perceptual memory as suggested by Averbach and Coriell (1961) and Sperling (1963). . . . The nonsense syllable would seem to be capable of being perceived only if the two halves are perceived as psychologically simultaneous or the perceptual trace of the first half is still present when the second stimulus half occurs'' (p. 477). Second, Julesz and Chiarucci (1973) noted that, ''In the studies of Eriksen et al., the integration of successively presented dots into a form is a special case of the short-term memory integration extensively studied by Sperling.'' (p. 251). Finally, Haber & Standing (1970), describing a synchrony-judgment task, noted that, ''[An] indicator of the visual persistence of a flash . . . yields data of the same high order of reliability and magnitude found by the far more laborious and indirect procedures of Sperling (1960), Averbach and Coriell (1961) and others.''

This simple view, appealing though it was to early investigators, has been called into question. In an oft-cited and exhaustive article, Coltheart (1980) challenged this view based on dissociation logic.[4] In particular, Coltheart noted that two major independent variables—stimulus duration and stimulus luminance—appear to have negative effects on visible persistence, but little if any effect on informational persistence. Recently, however, the validity of this observation has been questioned by a number of investigators. For example, Long and Beaton (1982) reported a positive relation between stimulus intensity and partial-report performance, while Di Lollo and Dixon (1988; see also Di Lollo & Dixon, 1992a; Dixon & Di Lollo, 1994) discovered that under certain circumstances there is a negative relation between stimulus duration and partial-report performance. Nisly and Wasserman (1989) argued that stimulus luminance could have a positive, a negative, or no effect on either visible persistence or informational persistence depending on the luminances used and the instructions given to the subjects.[5]

Coltheart's conclusions, whether accepted or not, are based on a foundation that is seriously flawed in two respects. First, at the time of Coltheart's writing there had been little *systematic comparison* of different tasks. That is, there was little investigation wherein task alone was varied while all other relevant variables were held constant. To illustrate, consider a typical comparison made in the literature of a partial-report task in which duration is varied (e.g., Yeomans & Irwin, 1985) with a synchrony-judgment task in which duration is varied (e.g., Efron, 1970a, b). While both tasks do indeed involve investigation of stimulus duration effects on performance of some sort, the two experiments involved different stimuli, observers, duration ranges, luminances, contrasts, and probably other differences as well. These confoundings bring about two related consequences. First, the difference in

[4] Coltheart did not actually use the term ''dissociation logic'' as dissociation logic had not yet come into vogue in 1980.

[5] Based on such findings, Long (1979, 1985) and Sakitt and Long (1979) have argued that there are two kinds of visible persistence, rather than just one as Coltheart (1980) proposed. According to Long (1985), Type I persistence is affected negatively by increasing stimulus luminance and duration, which Type II persistence is affected positively by increasing stimulus luminance and duration. This proposal has been criticized for a number of reasons by several investigators (e.g., Bowling & Lovegrove, 1982; Coltheart, 1980; Di Lollo, 1983, 1984; Irwin & Yeomans, 1986a), however, so in this paper we adopt Coltheart's (1980) position.

the reported duration effects (null effects in partial report vs. inverse-duration effects in synchrony judgment) cannot be unambiguously attributed to any one factor. Second, any theory purporting to account for both informational and visible persistence effects must incorporate otherwise irrelevant subtheories of all differences between the two experiments in order to make explicit predictions.

A second, and more fundamental difficulty with Coltheart's conclusion is that he provided little formal consideration of the logic by which multiple persistence measures may or may not be inferred to reflect the same underlying thing. In a subsequent section, following a brief description of the specific tasks with which we are concerned, we will undertake such a consideration in some detail.

In sum, despite substantial discussion of the issue by Coltheart (1980) and others, it remains unclear whether visible and informational persistence reflect the same perceptual entity, or different perceptual entities. If the latter, it is of theoretical significance to determine the relations between these two types of persistence, in order to develop more accurate and more sophisticated theories of visual information processing.

TASKS

In this article, we compare three tasks (temporal integration, subjective completeness ratings, and partial report) that measure some aspect of visible and/or informational persistence under circumstances that are as similar as possible. Exposure duration and interstimulus interval are varied to determine whether these manipulations affect the three tasks in the same way. We briefly describe these tasks here because we will refer to them in our upcoming analysis of the logic by which different tasks may be inferred to measure the same thing or different things.

Experiments 1 and 2 were designed to investigate the relation between two of the tasks sketched in Table 1: an objective missing-dot temporal-integration task (Hogben & Di Lollo, 1974; Di Lollo, 1980) and a subjective completeness-rating task (Loftus & Hanna, 1989). In these first two experiments, 24 dots filling all but one of the 25 cells of an imaginary 5×5 matrix were presented as two temporally distinct 12-dot halves: each half was of some duration, and the two halves were separated by some interstimulus interval (ISI). In Experiment 1, half-1 duration (HID) and ISI were varied while half-2 duration (H2D) was always 20 ms (as in Di Lollo & Dixon, 1988). In Experiment 2, H1D and H2D were varied while ISI was always 0 ms (as in Dixon & Di Lollo, 1992; 1994). The objective task was to report the missing dot's position, while the subjective task was to rate whether the complete (both-half) combination appeared to entail one or two distinct temporal events.

Experiments 3–7 were designed to investigate the relation between subjective completeness and partial report. In these experiments, letter arrays were presented followed by a visual probe signaling the observer which letter was to be reported. The observer then either reported the signaled letter (as in Averbach & Coriell, 1961) or provided a subjective rating of the degree to which the letter array-probe combination appeared to entail one or two distinct temporal events. The range of stimulus durations and stimulus-probe ISIs was the same as in Experiment 1, thereby allowing direct comparisons among the experiments.

WHEN DO DIFFERENT MEASUREMENT TECHNIQUES MEASURE THE SAME THING?

Any science must be fundamentally concerned with the questions: What does a measurement technique measure, and when do physically different measurement techniques measure the same thing? Consider two examples.

Suppose first that one wishes to measure heights of trees. Two possible measurement techniques are these. First one could climb to the top of the tree and drop a stone; here the tree's height would be defined as the observed time required by the stone to reach the ground. Second, one could place a sextant on the ground at a specified distance, d ft., from the base of the tree and then aim the sextant at the tree's top; here, the tree's height would be defined as the sextant's observed angle, α. Intuitively, it seems that both of these techniques would measure the same thing: the intrinsic height possessed by the tree. Formally, this intuition translates into what we refer to as the *monotonicity prediction* which, in this instance is: If by one technique one tree, Tree A is shorter than another tree, B, then A will be shorter than B by the other technique as well. Likewise, two trees that are equally tall by one technique will also be equally tall by the other technique.[6]

As a second example, suppose one wishes to measure the two-dimensional size of cardboard rectangles. Again consider two techniques. First one could wrap a string around the to-be-measured piece's perimeter, noting the amount of string required to exactly circumnavigate the piece: In this ''perimeter technique,'' the piece's size would be defined as the string's observed length. Second, one could *weigh* the two-be-measured piece; here, the piece's size would be defined as the observed weight. Intuitively these techniques would not be measuring quite the same thing. This intuition would translate into failure of monotonicity. For instance one might observe two rectangles (e.g., a 4×4 rectangle and a 1×8 rectangle) whose size ordering is reversed for the two measurement techniques; likewise, one might observe two rectangles (e.g., a 4×4 rectangle and a 2×8 rectangle, or a 4×4 rectangle

[6] Specifically, in this case, the relation between t, the time in seconds measured by the dropping technique and α, the angle measured by the sextant technique, is $\alpha = \tan^{-1}(16t^2/2)$.

and a 2 × 6 rectangle) that are equal by one technique, but unequal by the other.

We now return to the measurement task at hand. In psychology, one often measures the magnitude of some assumed (but typically only loosely defined) internal construct, such as ‘‘magnitude of persistence’’ in terms of performance on some task. What does it mean in such a situation for different tasks to measure the same thing? The logic by which this question can be addressed is described in detail by Bamber (1979). With respect to the specific question at hand, the question can be addressed by reference to the hypotheses shown in Fig. 1.

Three Hypotheses

Figures 1A–1C presuppose a situation in which two stimuli are sequentially presented, and the observer must perform some task that, in one way or another, requires integrating the information in the two stimuli (for example, a first stimulus-second stimulus temporal-integration task, or a stimulus-probe partial-report task). The independent variables shown at the left (duration and ISI) refer to the duration of the first stimulus and to the ISI between the first stimulus's offset and the second stimulus's onset.

Figure 1A represents the hypothesis that different tasks measure the same thing, while Figs. 1B and 1C represent two alternative hypotheses. Note that the Fig. 1A hypothesis is exclusive of both the Fig. 1B and the Fig. 1C hypotheses, although the latter two are not exclusive of one another.

In each of Figs. 1A–1C, the following conventions have been used. Shaded rounded rectangles represent independent variables, and shaded ovals represent dependent variables. Open rectangles represent unobservable internal constructs. The unshaded rounded rectangle in Fig. 1C represents an uncontrolled internal variable.

Predictions

The hypothesis embodied in Fig. 1A is the conjunction of two assumptions. The first is that the independent variables whose combination defines the conditions in some experiment (duration and ISI in this example) jointly influence the formation of some internal representation that is *unidimensional*. By this we mean that whatever quality of the representation determines performance in the tasks of interest can be represented by a single number, R, on a unidimensional scale (for example, R might be ‘‘magnitude of perceived visibility.’’) The second assumption of the hypothesis is that this representation is the sole determinant of performance in each of two tasks (e.g., temporal integration and partial report); in other words, performance in Tasks 1 and 2 are monotonic functions of R, m_1(R), and m_2(R).

An example of this hypothesis is provided by Dixon and Di Lollo (1994). In their admirably well articulated theory, Dixon and Di Lollo assume that

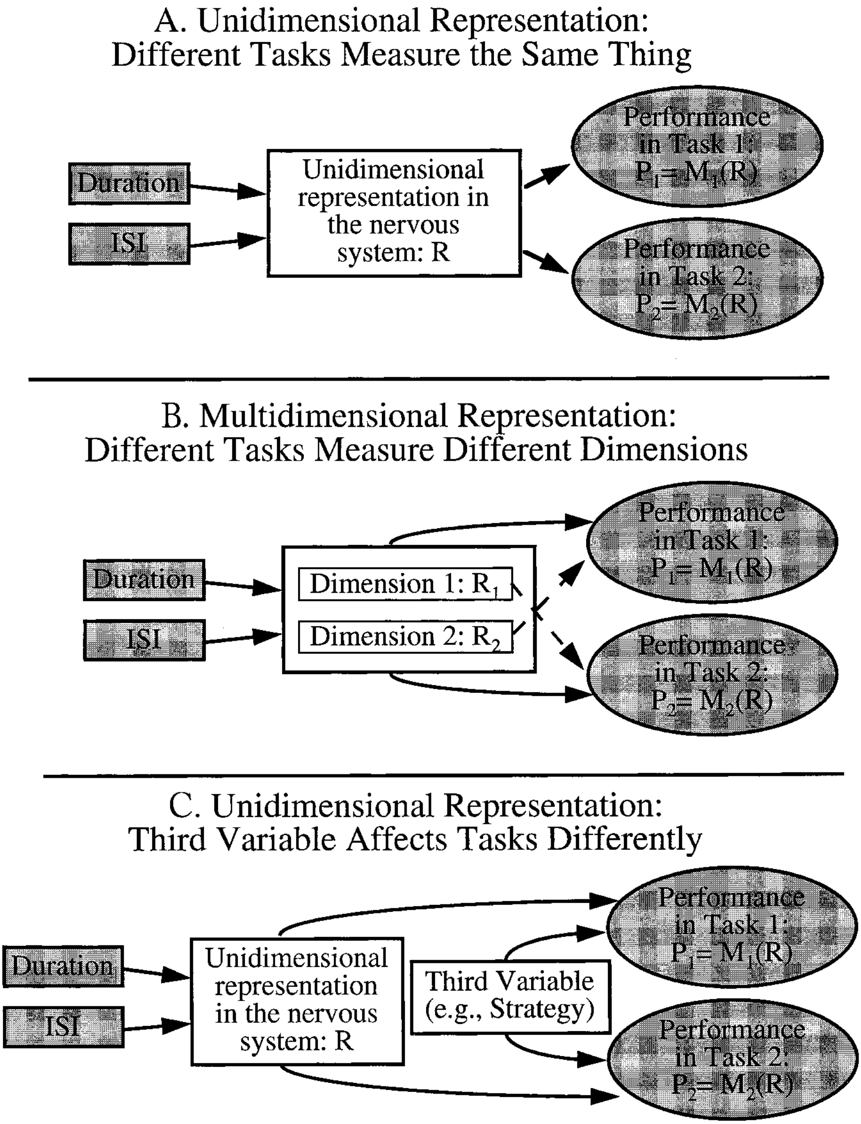

FIG. 1. Three hypotheses about the relation of different persistence tasks. Top panel: Two tasks measure the same thing. Bottom two panels: Two tasks measure somewhat different things.

two sequentially presented stimuli—either two spatial halves of a complete object in the case of temporal integration, or a stimulus array followed by a probe in the case of partial report—give rise to a temporally weighted temporal *correlation* between two internal temporal waveforms engendered by the stimuli. Correlation magnitude is thus the relevant unidimensional scale in Dixon and Di Lollo's theory; temporal-integration and partial-report performance are then assumed to be monotonic functions of correlation magnitude.

One parenthetical note is in order here. We wish to emphasize that the assumed unidimensionality of R in Fig. 1A does *not* require that the entire memory representation be unidimensional: only that facet of it that determines performance in the tasks under consideration is, by the hypothesis, unidimensional. The hypothesis allows other dimensions that may influence performance on other tasks.

The Monotonicity Prediction

As we suggested earlier, the Fig. 1A hypothesis implies the monotonicity prediction, which can be expressed simply as: the rank order correlation over experimental conditions between Task-1 performance and Task-2 performance is 1.0. The reasoning behind this prediction is straightforward. Consider any two experimental conditions, i and j. Designate Task-1 performance in the two conditions $P_1(i)$ and $P_1(j)$; likewise, designate Task-2 performance in the two conditions $P_2(i)$ and $P_2(j)$. A finding that $P_1(i) > P_1(j)$ implies that $R(i) > R(j)$, where $R(i)$ and $R(j)$ are the memory representations issuing from Conditions i and j, which in turn implies that $P_2(i) > P_2(j)$. This is the definition of an over-condition monotonic relation between P_1 and P_2.

Possible Inferences Given Failure of the Monotonicity Prediction

Confirmation of the monotonicity prediction would confirm the hypothesis shown in Fig. 1A. Suppose, however, that the monotonicity prediction fails. Such an outcome would disconfirm the general proposition that Task 1 and Task 2 measure the same underlying thing, and would be consistent with a variety of alternative possibilities, two of which are represented by Fig. 1B and 1C.

Multidimensional systems. In Fig. 1B, the prediction fails because of failure of the unidimensionality assumption. Here, as illustrated, there are two dimensions in memory: The value along Dimension 1 determines (or primarily determines) performance in Task 1, while the value along Dimension 2 determines (or primarily determines) performance in Task 2. As an example of such a multidimensional system, Di Lollo and Dixon (1988) proposed a model to account for both partial-report and temporal-integration data. By this model, Dimension 1 is ''visible persistence'', which primarily determines temporal-integration performance, while Dimension 2 is ‘‘schematic persistence’’ which primarily determines partial-report performance.

Strategy effects. In Figure 1C, the memory representation is again assumed to be unidimensional, but the prediction fails because the memory representation is not the sole determinant of performance in the task. Instead, there is at least one additional variable (‘‘strategy’’ in our example), that is assumed to affect the two tasks differently. Complex and sometimes idiosyncratic strategies have often been found to play an important role in partial-report tasks (e.g., Sperling, 1960; Gegenfurtner & Sperling, 1993) thus one might suspect a priori that Hypothesis C might well characterize the relation between partial report and other persistence related tasks.

Confirmation of the monotonicity prediction—which constitutes, in turn, confirmation of the Fig. 1A hypothesis—across tasks that measure visible persistence and tasks that measure informational persistence would provide strong support for the view of iconic memory, exemplified by our earlier quotes of Erickson and others, which assumes temporal-integration tasks and partial-report tasks measure the same thing—i.e., that visible and informational persistence arise from the same source. In contrast, failure of the monotonicity prediction would imply that, one way or another, visible and informational persistence arise from different perceptual and cognitive processes.

We should emphasize that confirmation of the monotonicity prediction between temporal-integration and partial-report tasks seems unlikely on an apriori basis (this was the major point of Coltheart’s argument, although see Dixon and Di Lollo, 1994 for an opposing view). There are many ways in which the single-dimension hypothesis can fail. In Fig. 1, we provide two general such ways. As we have noted, Di Lollo and Dixon (1988) among others provide a model according to which more than two dimensions are necessary to account for both temporal-integration and partial-report data, and below we offer another such model.

State-Trace Logic and Dissociation Techniques

Over the past decade, the use of *dissociation techniques* has become quite popular as a means of testing whether or not two tasks do or do not measure the same underlying entity. We have already noted that dissociation techniques were central to Coltheart’s (1980) logic, although Coltheart did not explicitly use the term. In recent research, dissociation techniques have been used, for example, to conclude that because declarative memory task performance is affected by attention, while implicit memory task performance is not, these two techniques measure two different kinds of memory (e.g., Gardiner & Parkin, 1990).

It is worthwhile to note that the kind of state-trace techniques that we have described, and in particular the monotonicity prediction *subsume* dissociation logic. By this we mean the following. By dissociation logic, one assumes two tasks to be measuring the same thing if some independent variable has an effect on performance using one technique, but no effect (or the opposite effect) on performance using the other technique. If this situation holds, then the monotonicity prediction must fail. However, a dissociative outcome, while sufficient, is not *necessary* for the monotonicity prediction to fail. Let us reconsider, for instance, our two ‘‘rectangle size’’ measurement techniques—the perimeter technique and the weighing technique which, as we have noted, do *not* measure the same thing. Two independent variables might be the length and width of the rectangles. As either variable increases, size would increase by either technique; thus, there would be no dissociation with respect to either variable. Yet, as demonstrated earlier, it is quite easy to select examples by which the monotonicity prediction is disconfirmed.

Theory

As noted earlier, our second goal in this article is to account for our data with a theory that has been presented elsewhere by Loftus and his colleagues (e.g., Loftus, Busey, & Senders, 1993; Loftus & Ruthruff, 1994; Busey & Loftus, 1994; 1998). Application of this theory to visible persistence has its roots in models described by Loftus and Hogden (1988), Loftus and Hanna (1989), Loftus and Busey (1992), and Dixon and Di Lollo (1994).

In a later section, we describe the theory in detail. Briefly, it is as follows. A stimulus, characterized as a function relating intensity to time, t, since stimulus onset, engenders a *sensory response* in the nervous system. The sensory-response function relating sensory-response magnitude to t is temporally blurred relative to, and lags behind, the physical stimulus. The sensory response serves as a basis for acquiring information from the stimulus; accordingly, we can derive a function relating information-acquisition rate to t. Stimulus visibility at any given time is identified with the magnitude of the information-acquisition rate at that time.

When two stimulus halves are separated in time, as in Experiments 1 and 2, there are *two* information-acquisition rate functions, one corresponding to each stimulus half. Both temporal-integration and subjective-rating performance are assumed to depend on the degree to which these two acquisition-rate functions are temporally correlated with one another (cf. Dixon & Di Lollo, 1994).

In a partial-report task, there are also assumed to be two information-acquisition rate functions: one corresponding to the stimulus array, and the other corresponding to the probe. As in the dot-matrix task, subjective completeness is determined by the temporal correlation of the two functions. However, partial-report performance is determined by the amount of information acquired from the stimulus array which, logically, is equal to the integral over time of the information-acquisition rate.

EXPERIMENTS 1 AND 2: COMPARISON OF TEMPORAL INTEGRATION AND SUBJECTIVE COMPLETENESS

The temporal-integration task, first reported by Eriksen and Collins (1967), has long been used to study effects of certain variables on visible-persistence duration. In this particular temporal-integration task—the missing-dot task, introduced by Hogben & Di Lollo (1974)—an array of dots (typically a 5×5 array) is randomly divided into two spatial halves of 12 dots per half. Note of course that two 12-dot halves in a 5×5 matrix leave one dot missing. If the two halves are presented contemporaneously, it is exceedingly easy to detect and correctly report the cell that doesn't contain a dot. If, however, the two halves are separated in time, then under some conditions temporal-integration-report performance declines precipitously. Such a decline takes place, in particular, with increases in (1) the ISI separating half 1 offset from half-2 onset (Hogben & Di Lollo, 1974), (2) H1D (Di Lollo, 1980), and (3) H2D (Dixon & Di Lollo, 1992a; 1994).

A general hypothesis to explain these findings is that performance depends on the degree to which the two halves are perceived as a single or a dual temporal event. Subjective experience lends this hypothesis a good deal of appeal. If the ISI is long, the two halves are, of course, seen as two temporally distinct displays. Although not nearly as obvious, lengthening the duration of either the first or second half leads to the same impression. If, for instance, the first half is lengthened, it appears subjectively to be ''processed and over with'' by the time the second half appears. The second half then seems to be processed as a separate entity. Because the two halves are seen as temporally disjoint events, it isn't possible to perceive the missing dot's location. In contrast, when H1D, ISI, and H2D are all short (e.g., 20 ms or less) then the two halves are perceived as a single temporal event, consisting of 24 dots that leave one obviously empty cell whose location is easy to report.

Loftus and Hanna (1989) investigated the proposition that temporal integration is related to the subjective experience of simultaneity. They presented dot matrices in temporally successive halves, varying H1D and ISI. Instead of asking observers to detect a missing dot, however, they asked observers to *rate* how temporally unified the entire display appeared to be on a scale from 4 (indistinguishable from all dots displayed simultaneously) to 1 (two halves appear to be completely separate temporal events). The rating data mimicked temporal-integration data: mean rating declined in an orderly and dramatic fashion with both H1D and ISI.

While consistent with the proposition that temporal-integration performance is intimately associated with subjective completeness, Loftus and Hanna could not assess the degree of intimacy very strongly because they did not test their observers in an actual temporal-integration task. The present Experiments 1 and 2 were designed to collect both temporal-integration and subjective-completeness judgments, thereby allowing a direct comparison between the two tasks. In Experiment 1, 5 values of H1D were factorially combined with 5 values of ISI (with H2D held constant at 20 ms), while in Experiment 2, the same 5 values of H1D were factorially combined with 5 values of H2D (with ISI held constant at 0 ms). The major purpose of Experiments 1 and 2 was to examine whether the phenomenological experience of stimulus completeness and the ability to accurately detect the location of the missing dot measure the same underlying perceptual quantity, as illustrated in Fig. 1A, wherein ''Task 1'' is a measure of subjective completeness and ''Task 2'' is a measure of temporal integration. The major question we address is: to what degree is performance in the two tasks correlated over conditions? As discussed earlier, a perfect rank-order correlation would represent the strongest possible confirmation of the Fig. 1A hypothesis. A less-than-perfect rank-order correlation would confirm the Fig. 1B and/or the Fig. 1C hypotheses.

General Method

Each observer in Experiments 1 and 2 completed a series of trials in both the temporal-integration and subjective-completeness tasks. In each task the stimuli consisted of 5×5 dot matrices that had been randomly divided into two halves of 12 or 13 dots each, which were presented in rapid succession. In the (objective) temporal-integration task, 24 of the 25 dots in the matrix were presented, with 12 dots per half, and the observer's task was to report the location of the matrix dot that had not been presented. In the (subjective) completeness-rating task one half contained 12 dots and the other contained 13 dots. The observer's task was to rate how complete the 5×5 matrix appeared to be following presentation of the two halves.

Experiment 1

Experiment 1 consisted of a 5 (H1D) $\times$ 5 (ISI) factorial design. H1D ranged from 20 to 100 ms in 20-ms steps, and ISI ranged from -20 to 60 ms in 20-ms steps. Two performance measures were collected: temporal-integration performance and subjective-completeness ratings.

Method

Observers. Fourteen observers from the Michigan State University community, including both undergraduates and graduate students, served as observers. Observers had little or no knowledge of the experimental hypotheses. All observers reported normal or corrected-to-normal vision, and each was paid $30 for participating in 6 sessions.

Stimuli. The stimulus patterns consisted of 5×5 dot arrays that had been randomly divided into two halves, with 12 or 13 dots in each half. Twenty-five separate random assignments of 12 or 13 dots to matrix locations were created. We refer to these as the A-patterns. The complements of the A-patterns (i.e., the other 12 or 13 of the 25 dots) are referred to as the A_c (c for ''complement'') patterns. The 25 12-dot/12-dot patterns were such that the missing dot occurred once in each of the 25 cells. The same A and A_c patterns were used repeatedly throughout the experiment.

Apparatus. Stimuli were displayed on a Hewlett Packard 1340A x-y oscilloscope (P31 phosphor) driven by a Digital Equipment Corporation Micro-11/23+ computer through digital to analog converters. The computer also recorded responses entered by the observer into the terminal keyboard. Observers were seated 36 cm from the display and used a chin rest to keep their heads steady. At this viewing distance the oscilloscope subtended 20 deg of visual angle horizontally and 15 deg vertically. The 5×5 dot matrices subtended 3.2 deg horizontally and vertically; each dot subtended 0.04 deg, and the spaces between dots subtended 0.75 deg.

The experimental chamber was dimly illuminated during the experiment, so a red filter and a blue filter were lowered over the face of the display scope to reduce phosphor persistence visibility. The luminance of this background was 2 cd/m^2; stimulus displays were presented with an effective luminance of 21 cd/m^2. Shutter tests similar to those described by Irwin, Jonides, and Yantis (1983) confirmed that no phosphor persistence was visible 5 ms after stimulus offset.

Procedure. The sequence of events during each trial was very similar in the two tasks. To begin a trial, the observer pressed the carriage return key on the terminal keyboard. This caused a fixation point (two small dots centered above and below the middle of the display) to be presented for 500 ms. This was followed 500 ms later by the first half of dots, which was presented for a H1D of either 20, 40, 60, 80, or 100 ms. A blank ISI of either −20, 0, 20, 40, or 60 ms then elapsed before the second half of dots was presented for a H2D of 20 ms. An ISI of −20 ms signifies that the second half of dots overlapped in time with the last 20 ms presentation of the first half of dots (note that in the 20 H1D/−20 ISI condition, the entire array was presented for 20 ms). In the temporal-integration task, each half contained 12 dots; in the subjective-completeness task, one half contained 12 dots and the other contained 13 dots. The half containing the extra dot was counterbalanced across trials.

Following presentation of the two dot halves, the observer entered his or her response into the keyboard. In the temporal-integration task, a response consisted of the row and column coordinates corresponding to the location at which no dot had been presented. In the subjective-completeness task, the observer provided a rating ranging from 1 to 4 of how temporally integrated the 5×5 dot matrix appeared to be. Observers were instructed that a rating of ''4'' meant that one complete matrix appeared to have been presented, whereas a rating of ''1'' meant that two temporally separate displays appeared to have been presented. Ratings of ''2'' and ''3'' were used for intermediate perceptions. No feedback was provided.

Each observer completed 6 sessions, each session containing two blocks of trials. The observer performed the temporal-integration task in one of the two blocks, and the subjective-completeness task in the other. The order in which these tasks were performed alternated over sessions. Half of the observers began their initial block with the temporal-integration task and half began with the subjective-completeness task. The first session consisted of 100 practice trials in each task; these data were discarded. The remaining sessions contained 250 trials in each task. Thus, over the course of the experiment, each observer completed 1250 temporal-integration trials and 1250 subjective-completeness trials. These 1250 trials consisted of 50 replications of the 25 conditions formed by the factorial combination of H1D and ISI. The observers saw each of the twenty-five 5×5 dot patterns twice in each condition, once with the ''A'' pattern appearing first in the display sequence and once with the corresponding ''A_c'' pattern appearing first. H1D and ISI were sequenced randomly over trials, but observers saw the same patterns in the same order under the same timing conditions in the two experimental tasks.

Results

Figure 2, which shows the main data, is organized as follows. Each panel represents some measure as a function of ISI with different curves plotted

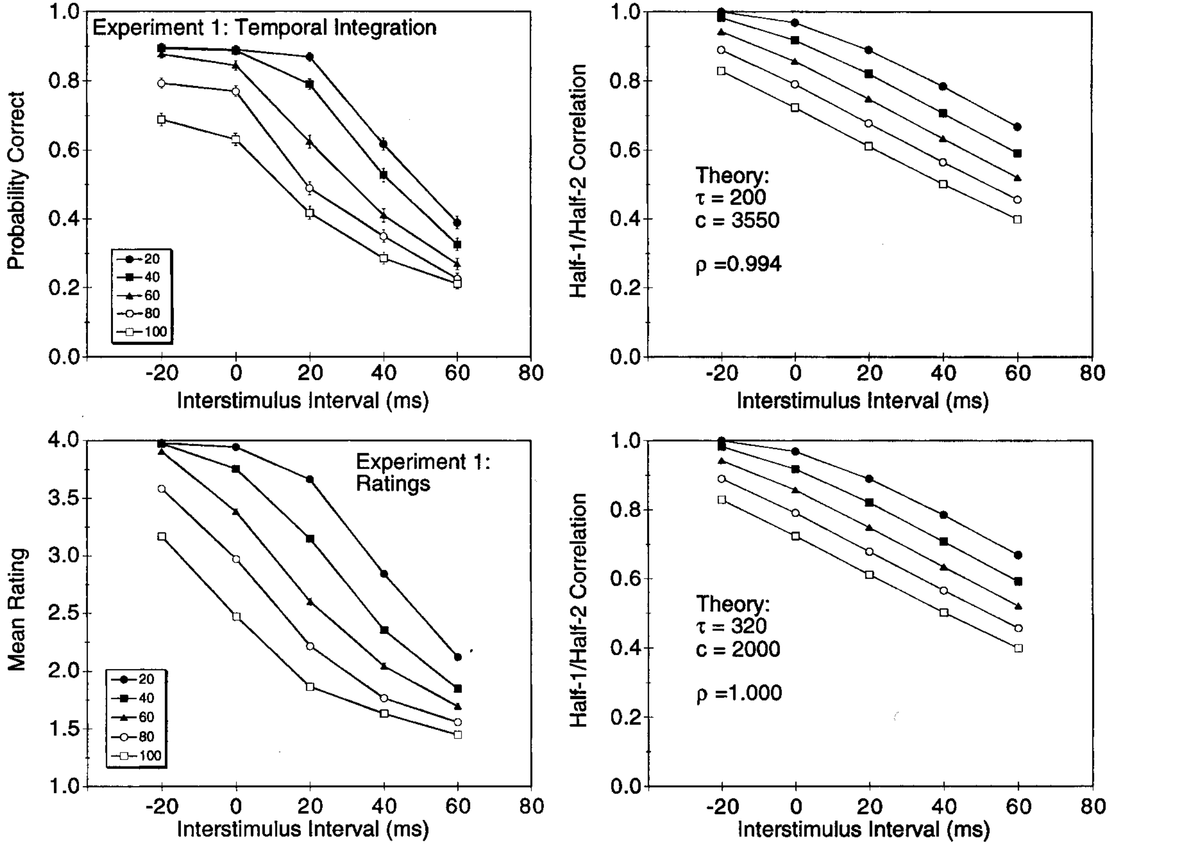

FIG. 2. Experiment 1: Main results. Top panels represent objective measure (temporal integration); bottom panels represent subjective data (completeness ratings). Left panels represent observed data; right panels represent theoretical predictions (to be described below). In all cases, performance, or the theoretical determinant of performance is plotted as a function of ISI. Different curves represent different H1D levels, as indicated by the figure legend. H2D was always 20 ms.

for the different H1D values as indicated in the figure legend. (Except where explicitly noted otherwise, this and subsequent figure legends *always* represent different H1D values, and the same curve symbols always correspond to the same H1Ds). The left panels show the data, while the right panels show the output of the theory that we shall describe below. The top panels depict the objective measure (temporal-integration performance) while the bottom panels depict the subjective measure (rating). In the (left-hand) data panels, the error bars represent one standard error.[7]

The Fig. 2 data panels indicate very systematic, and similar, effects of ISI and H1D on both the objective and subjective measures: each declines with increases in both variables. The subjective-completeness data provide essentially a perfect replication of analogous data reported by Loftus and Hanna (1989).

Is performance determined solely by SOA? Given a performance decline with both H1D and ISI, a simple hypothesis suggests itself: perhaps performance is determined entirely by the stimulus onset asynchrony (SOA) that intervenes between the onset of half-1 and the onset of half-2. To address this hypothesis, Fig. 3 shows performance as a function of ISI with different curves for the different SOA values (note plotting as a function of ISI is arbitrary: within each curve, increasing ISI corresponds to decreasing H1D). We can clearly reject the hypothesis that SOA alone determines performance. With the exception of three of the −20-ms ISI conditions, performance declines uniformly with ISI, even with SOA held constant. This effect is less dramatic for the subjective than for the objective task, but it holds consistently for both tasks nonetheless. This finding can be interpreted to mean that the cognitive determinant of performance in the tasks declines faster with increases in ISI than with increases in H1D.

The objective/subjective correlation. Our fundamental question, designed to address the hypotheses shown in Fig. 1, was: to what degree are the two performance measures correlated over the 25 experimental conditions? Table 2 provides these objective-subjective correlations (Spearman ρ's) for all seven experiments. For Experiment 1, the correlation was 0.977, which, while high, is not perfect. Fig. 4, top panel, shows the scatterplot corresponding to this correlation: here, objective temporal-integration performance is plotted against subjective completeness rating. Again the different symbols correspond to different H1D values. Points corresponding to a given H1D are fit with quadratic equations.[8] Of some interest is that the less-than-perfect

[7] In many instances there appear to be no error bars. This is because the error bars in these instances are smaller than the curve symbols.

[8] The selection of quadratic equations to fit these curves was somewhat arbitrary. Such equations fit these and other analogous curves quite well, and provide a visual basis for comparing the scatterplot points belonging to different H1D levels.

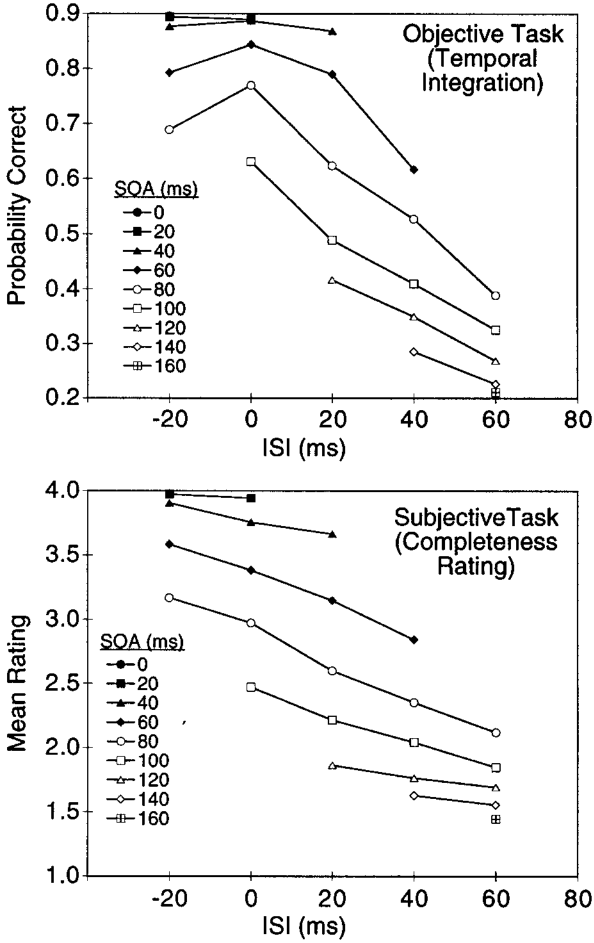

FIG. 3. Experiment 1: SOA effects. Performance is plotted as a function of ISI for different levels of SOA = (H1D + ISI). Objective and subjective measures are shown in top and bottom panels.

correlation between the two measures does not occur as a result of noise in the data. Instead, it is clear that the long-H1D curves in the scatterplot tend to be to the left of the short-H1D curves. This finding can be interpreted as follows. Consider a set of points from different H1Ds that lead to equal objective performance. As indicated by the narrow horizontal rectangle on the figure, such a set of points corresponds to a horizontal cross-section of the Fig. 4 scatterplot. Holding objective performance constant, the longer the H1D, the lower is subjective performance. This means that subjective (completeness-rating) performance declines more with increasing H1D than

TABLE 2
Correlations between the Objective Measure (Temporal Integration or Partial Report) and the Subjective Measure (Completeness Ratings) for Experiments 1–7

Experiment	Rank-order correlation (r)
1	0.977
2	0.979
3	0.628
4	0.744
5	0.808
6	0.667
7	0.774
4–7 (mean)	0.801

does objective (temporal-integration) performance.[9] This finding is consistent with a general rule that will be underscored by subsequent data: the more some dependent variable measures subjective experience, the greater the effect of H1D.

Discussion

The Experiment 1 results provide several sorts of information. First, both temporal-integration and subjective-completeness measures are affected similarly by ISI and H1D: both measures decline with increases in both independent variables. Second, the over-condition correlation between the two measures is high ($\rho = 0.977$) but not perfect. This lack of perfect correlation is not due to noise, as there are systematic differences between the two measures evident in Fig. 4: the subjective measure declines faster with H1D than does the objective measure.

The high intermeasure correlation indicates that to a large extent the objective and subjective tasks measure the same underlying thing. However, the small but still systematic differences between the two measures indicate that, contrary to the hypothesis depicted in Fig. 1A, the objective and subjective tasks do not measure *precisely* the same thing. In keeping with these observations, we next offer a hypothesis about the relation between our subjective rating measure and other objective measures.

This hypothesis can be expressed by the following two assumptions. First,

[9] To gain a feel for this reasoning, imagine that subjective performance was inversely affected by H1D, but that objective performance was not affected at all. In that case, the five curves constituting the scatterplot would all be horizontally aligned; that is, for any given ISI, there would be five H1D values falling along a horizontal line (corresponding to equal temporal-integration performance), going from the longest H1D (at a smaller rating value) to shortest H1D (at a larger rating value).

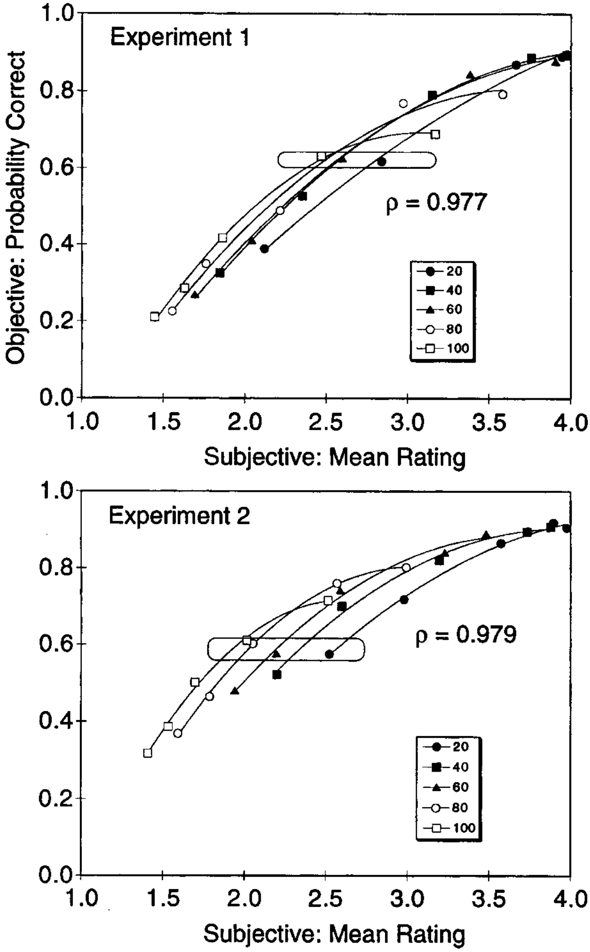

FIG. 4. Experiment 1 (top panel) and Experiment 2 (bottom panel): Correlations between objective measure (ordinate) and subjective measure (abscissa) over 25 experimental conditions. Different curve symbols represent different H1D levels. Data points within each H1D level are fit by quadratic functions.

two stimuli presented close in time can be perceived as either a single or a dual temporal event. The degree to which one or the other of these two perceptions occurs is accessible to consciousness and can be directly expressed as a subjective rating. The second assumption is that this subjective experience—or whatever underlies the subjective experience—heavily determines performance in an objective task, such as the temporal-integration task. However, as in any objective task, certain strategies, both general and idiosyncratic, also play a role in determining performance. In a temporal-integration task, for instance, encoding the location of a few half-1 dot locations in short-term memory might constitute one such strategy. Such a strategy might allow a subject to overcome to some extent the otherwise deleterious effects of increases in H1D and ISI on objective temporal-integration task performance: subjects might respond on the basis of information coded in short-term memory rather than on the phenomenological appearance of the stimulus. Strategies such as these would presumably be more useful in the objective task than in the subjective task, which requires only a judgment about the phenomenological appearance of the persisting stimulus. Thus, as depicted in Fig. 1C, the two tasks measure slightly different things and are less than perfectly correlated. In short, subjective performance appears to be a more sensitive measure of the quality of the subject's perceptual experience than is objective performance.

Experiment 2

In Experiment 1, H1D and ISI were varied to determine their effects on temporal-integration accuracy and subjective completeness; H2D was constant at 20 ms. In Experiment 2, H1D and H2D were varied while ISI was held constant at zero. Dixon and Di Lollo (1989; 1994) have demonstrated that increases in H2D lead to decreases in temporal-integration accuracy similar to those engendered by increases in H1D and ISI. In Experiment 2 we compared the effects of H1D and H2D in the temporal-integration task and in the subjective-completeness task in order to determine whether these manipulations affected performance in the two tasks in the same way.

Method

Observers. Eleven Experiment 1 veterans also served in Experiment 2. Each observer was paid $30 for participating in 6 sessions.

Stimuli and apparatus. The same stimulus patterns and apparatus used in Experiment 1 were used in Experiment 2.

Procedure and design. The procedure was very similar to that used in Experiment 1, except ISI was always 0 ms and H2D varied from 20 to 100 ms in 20 ms steps. Thus, the experiment consisted of a 5 (H1D = 20, 40, 60, 80, or 100 ms) $\times$ 5 (ISI = 20, 40, 60, 80, or 100 ms) factorial design, in which two performance measures—performance correct in a temporal-integration task, and subjective completeness ratings—were collected.

As in Experiment 1, each observer completed 6 sessions, each containing two blocks of trials: a temporal-integration block and a subjective-completeness block. Randomization and counterbalancing were the same as in Experiment 1.

Results

Figure 5 shows the main Experiment-2 results. In Fig. 5, unlike all other data figures, the abscissa is not ISI, but is H2D; note that the curve parameter remains H1D.

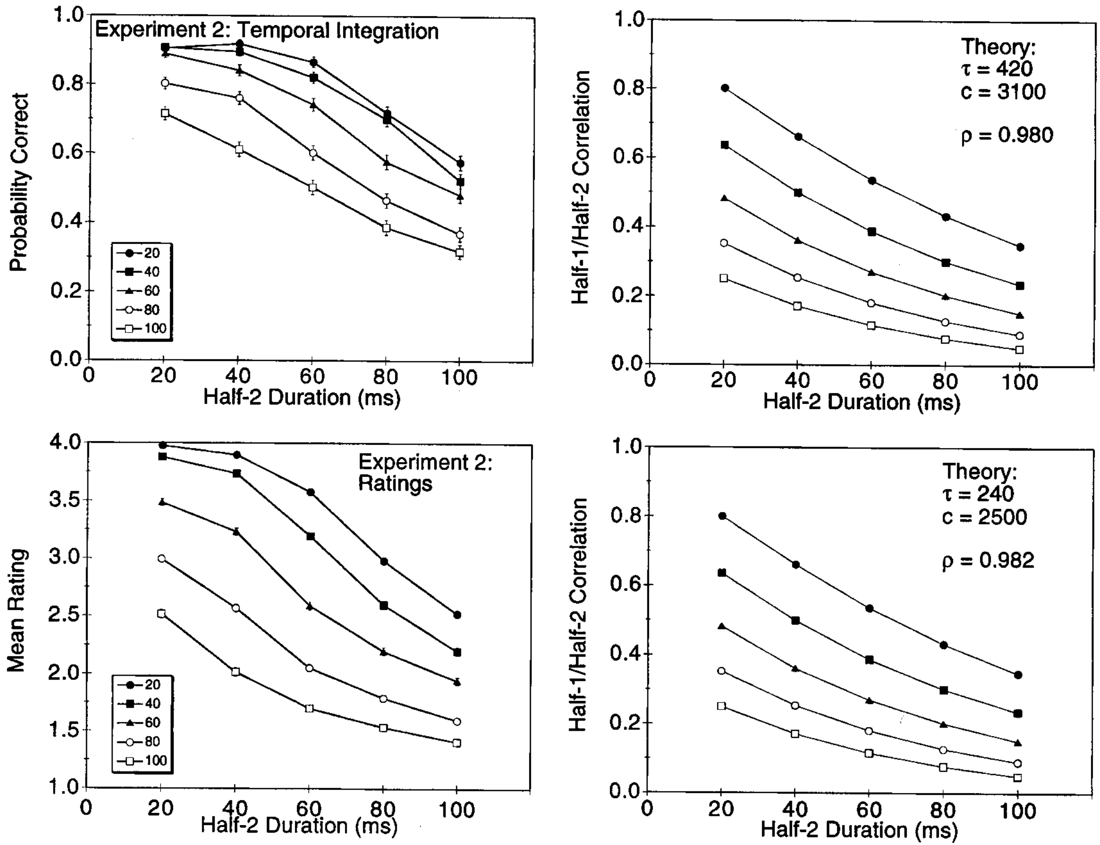

FIG. 5. Experiment 2: Main results. Figure 5 is organized like Fig. 2, except that performance, or the theoretical determinant of performance, is plotted as functions of H2D rather than ISI. ISI was always 0 ms.

It is clear that H2D closely mimics ISI in its effects on both performance measures. Comparison of Figs. 2 and 5 indicates that ISI and H2D effects are quite similar. The temporal-integration data provide essentially a perfect replication of analogous data reported by Dixon and Di Lollo (1992a; 1994).

Do H1D and H2D have symmetrical effects? Figure 5 appears to indicate that effects of H1D and H2D are very similar; both cause strong performance decreases. To determine the degree to which the effects are similar, the left panels of Fig. 6 show objective and subjective performance as functions of H2D, with different curves for different values of total duration, (i.e., of H1D + H2D). For the subjective task, these curves are essentially flat: for a given total duration, it makes no difference what portion of that duration comes from half 1 vs half 2. For the objective task, there is a small asymmetry between the two halves: in most cases, lengthening H2D has a marginally greater detrimental effect than lengthening H1D.

Dixon and Di Lollo (1994) report a temporal-integration experiment in which, as in the present Experiment 2, H1D and H2D were both varied. Their data are presented in the right panels of Fig. 6 which shows data for stimuli presented on a dim and on a bright background; again, different curves are presented for different total-time values. As is true with the present objective data, performance generally (but not always) declines modestly with increases in H2D.

Dixon and Di Lollo account for their data using a theory that predicts half-1 and half-2 duration effects to be symmetrical; hence their data, and our corresponding objective data, represent a minor disconfirmation of their theory. Their theory is confirmed, however, by our subjective data. We will have more to say on this topic in a later section.

The objective/subjective correlation. Recall that, as indicated in Fig. 1, the over-condition correlation between the two performance measures represents the degree to which the two tasks measure the same thing. All the logic depicted in Fig. 1 applies to Experiment 2; one need only substitute H2D for ISI. As in Experiment 1, the correlation was high, $\rho = 0.979$ (see Table 2), but was not perfect. The scatterplot, shown in Fig. 4 (bottom panel) is organized like that of Fig. 4, top panel, and shows a very similar pattern: again, subjective performance is affected more by H1D than is objective performance.

Discussion

The Experiment 2 data tell essentially the same story as do the Experiment 1 data. First, as reported by others (e.g., Dixon & Di Lollo, 1992a; 1994), H2D has a strong negative effect on temporal-integration performance, as does H1D and ISI. Second, there is a high, but not perfect, over-condition correlation between the objective and subjective tasks, indicating that they measure much, but not completely, the same thing. As in Experiment 1, the failure of perfect correlation is systematic: subjective performance declines faster with H1D than does objective performance, suggesting that subjective

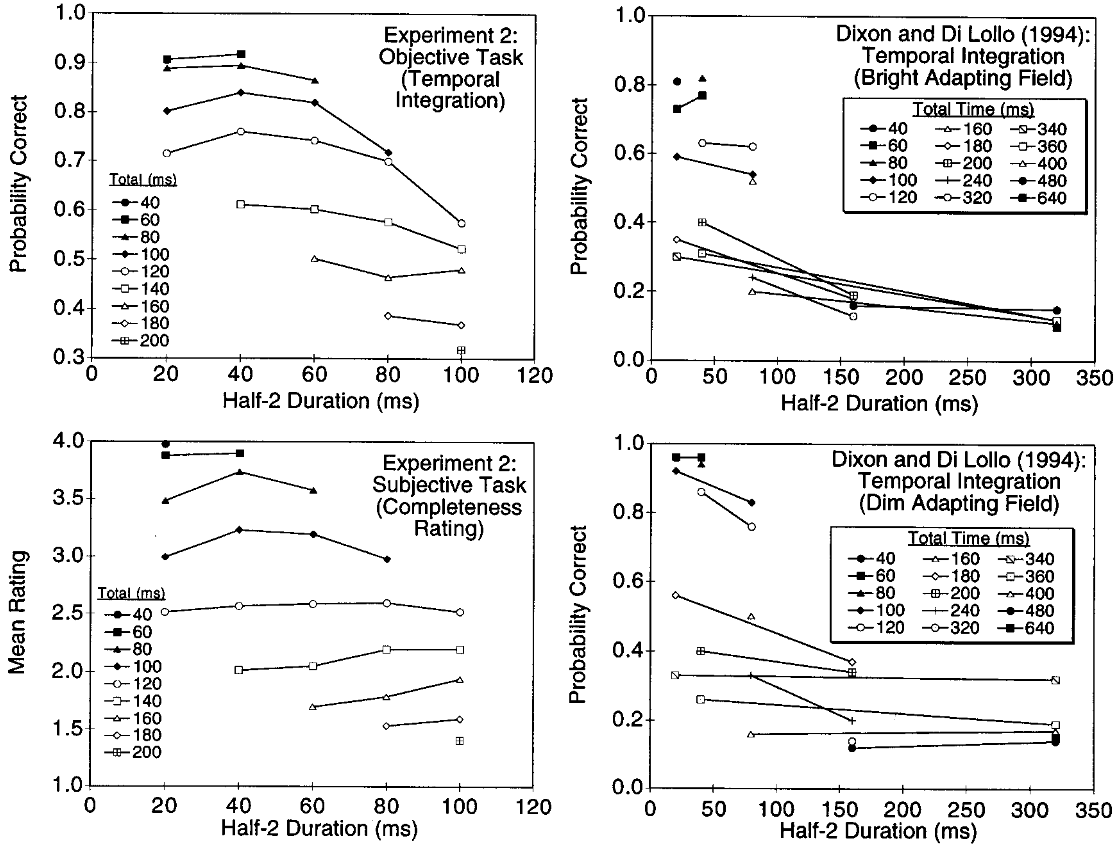

FIG. 6. Left panels: Experiment 2: Total-time effects. Performance is plotted as a function of H2D for different levels of Total Time = (H1D + H2D). Objective and subjective measures are shown in top and bottom panels. Right panels: Data replotted from Dixon and Di Lollo (1993). Top and bottom panels show data from dim and bright background conditions.

performance is a more sensitive measure of persistence. Finally, H1D and H2D have symmetrical effects on completeness-rating performance, but slightly asymmetrical effects on objective temporal-integration performance.

Theory: Temporal Integration and Completeness Rating

In this section, we describe a theory that has been applied to a variety of perceptual and cognitive phenomena, including iconic information acquisition in both picture recognition tasks (Loftus & Hogden, 1988), and digit recall tasks (Loftus, Duncan & Gehrig, 1992), synchrony-judgment performance (Loftus & Hogden, 1988), duration-intensity tradeoffs (Loftus & Ruthruff, 1994), and perceptual integration of temporally distributed information (Busey & Loftus, 1994; Loftus, Busey, & Senders, 1993) and binocular information acquisition (Busey & Loftus, 1998). The theory has also been applied to the present subjective-completeness task (Loftus & Hanna, 1989), and has qualitatively, but not quantitatively, accounted for both inverse-duration and ISI effects in partial report (Loftus & Busey, 1992).

The theory that we describe here is a refinement of that described by Loftus and Hanna (1989). This refinement borrows from a theory proposed by Dixon and Di Lollo (1994) in the sense of viewing perception of two successively presented stimuli as depending strongly on the *temporal correlation* between the internal perceptual functions assumed to result from presentation of the two stimuli. In what follows, we will develop the theory as it applies to the temporal-integration and completeness-rating data gathered in Experiments 1 and 2. In a later section, we apply the theory to the partial-report data of Experiments 3–7.

The Sensory-Response Function and Information Acquisition

As depicted in Fig. 7A, a stimulus is conceptualized as a function, $f(t)$, relating some stimulus attribute, such as intensity or contrast, to time, t, since stimulus onset. In Experiments 1 and 2, the stimulus was a square wave: as indicated, it abruptly appeared at time $t = 0$, remained on for a duration of d ms ($d = 20$ in Fig. 7), and abruptly disappeared.

The sensory-response function, a(t). The stimulus is assumed to engender a time-varying sensory response in the nervous system. The function relating sensory response magnitude to t, referred to as $a(t)$, is shown in Fig. 7B (the vertical lines in Fig. 7B and 7C represent the times of stimulus onset and offset). With others (e.g., Groner, Bischof & Di Lollo 1988; Dixon & Di Lollo, 1994; Loftus & Ruthruff, 1994; Sperling, 1964; Watson, 1986; Wolford, 1992) we assume that the sensory response results from a low-pass temporal filter applied to the stimulus input waveform, $f(t)$. As detailed in Appendix A, the equation for $a(t)$ is

$$a(t) = \begin{cases} \phi G(t) & \text{for } t < d \\ \phi[G(t) - G(t - d)] & \text{for } t \geq d, \end{cases} \qquad (1)$$

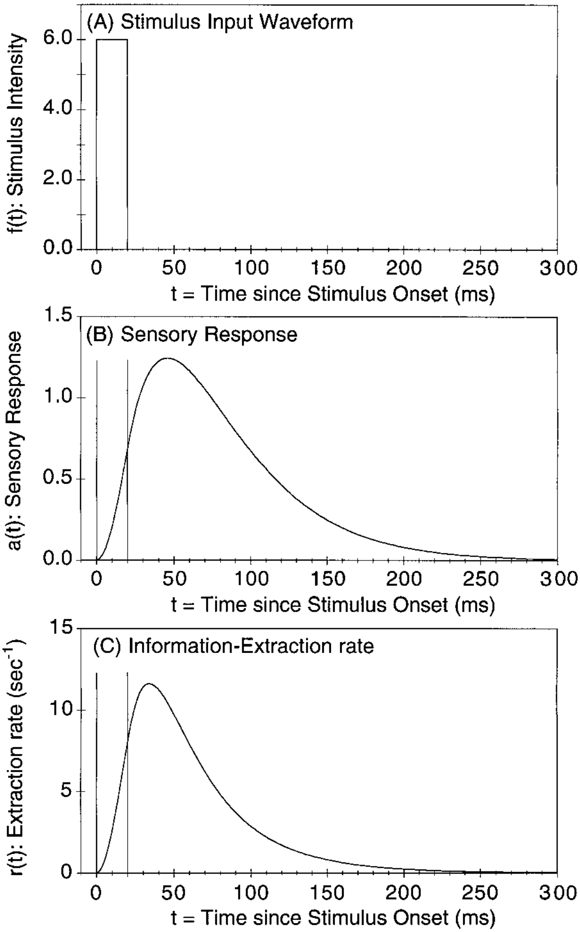

FIG. 7. Three major functions in the theory. All functions show some measure as a function of time since stimulus onset. Top panel: description of physical stimulus (intensity). Middle panel: sensory-response magnitude. Bottom panel: information-acquisition rate.

where ϕ is stimulus intensity and $G(x)$ is the integral from zero to x of the gamma function, $g(x)$, defined as,

$$g(x) = \frac{(x/\tau)^{n-1}\, e^{-x/\tau}}{\tau(n-1)!} \tag{2}$$

In Eq. 2, n and τ are free parameters: n is a positive integer, and τ is a positive real number.

The information-acquisition rate function, r(t). We next assume that some information-acquisition mechanism operates on the sensory response function to acquire stimulus information. At time t, information is acquired at an instantaneous rate, $r(t)$ that is computed by:

$$r(t) = a(t)h[I(t)] \tag{3}$$

Here, $I(t)$ is the proportion of stimulus information already acquired by time t, and h is a monotonically decreasing function, constrained such that $h[I(t)]$ approaches zero as $I(t)$ approaches 1.0. The intuitive interpretation of Eq. (3) is that information is acquired from the stimulus at a rate that is proportional to the sensory-response magnitude, but declines, in diminishing-returns fashion, with amount of already-acquired information (see Bundesen, 1990; Kowler & Sperling, 1980; Loftus & Kallman, 1978; Massaro, 1970; Rumelhart, 1971, Shibuya & Bundesen, 1988; Townsend, 1981, for similar information-acquisition models).

In past work, the theory has fit various sorts of data quite well by letting

$$h[I(t)] = [1.0 - I(t)]/c, \tag{4}$$

where c is a free parameter. Note that Eq. (4) follows from the presumption that information is acquired randomly and with replacement from the stimulus (see Loftus, Busey, & Senders, 1993; also, see below). Note also that c is in units of time (ms in our treatment). Intuitively, c represents the amount of time to sample some fixed amount of information from the stimulus and conversely, $1/c$ (‘‘per time’’) is the unit of processing rate: if c were very small, for instance, $r(t)$ would be initially high, and stimulus information would be acquired quickly.

Equations (3) and (4) imply that $r(t)$ is computed by

$$r(t) = a(t)[1.0 - I(t)]/c \tag{5}$$

Appendix A shows that, given Eqs. (3) and (4), the resulting equation for the information-acquisition rate, $r(t)$ is:

$$r(t) = [a(t)e^{-A(0,t)/c})]/c \tag{6}$$

where $A(0,t)$ is the area under the $a(t)$ function from time zero to time t. Figure 7C shows the r(t) function that results from the $f(t)$ and $a(t)$ functions of Fig. 7A and 7B.

A metaphor. Our model is a modified random-sampling-with replacement model. It is useful at this point to provide a (limited) physical metaphor for the information-extraction rate function embodied in Eqs. (5) and (6). Imagine the stimulus to be comprised of some large number of features (e.g., 1000 features). Stimulus onset initiates the gradual accumulation of these features in an urn, while stimulus offset initiates the gradual disappearance of the features from the urn. Thus the ‘‘number of features in the urn’’ function follows the time course depicted in Fig. 7B. Suppose further that the features assume a gas-like quality, in that they distribute themselves randomly within the urn’s volume; thus, the more features in the urn, the more densely packed they are. Finally, suppose that the features randomly redistribute themselves over time within the urn’s volume.

Features in the urn are *sampled* by a device that every c ms acquires all features within a given fixed volume of the urn. Thus, the number of features sampled per unit time is proportional to the number of features *in* the urn. Each sampled feature is determined to be a *new* feature (i.e., one that has never been sampled before) or an *old* feature (one that has been sampled at least once already). After such determination, all sampled features are returned to the pool (thus sampling is with replacement). Accordingly, the rate of sampling new features is also proportional to the number of new features in the urn; i.e., it is proportional to the product of total features in the urn (analogous to $a(t)$) and the proportion of those features that are new (analogous to $1.0 - I(t)$). It is this *rate of sampling new features* that is described by Eqs. (5) and (6).

The diminished status of visible ‘‘persistence’’ within the theory’s context. Before proceeding, a remark is in order about the general notion of ‘‘persistence’’. Traditionally, persistence has been associated with sensory and perceptual events that occur—and in some fashion, ‘‘decay’’—following the time of physical stimulus offset. Physical stimulus offset time is represented by the vertical lines in Fig. 7A and 7B, so ‘‘persistence’’ within the theory’s context, is represented by those portions of the functions to the right of the vertical lines.

The theory is conceptually at odds with traditional notions of ‘‘persistence’’ in at least two respects. First, neither $a(t)$ nor $r(t)$, necessarily decays immediately following stimulus offset. Instead $a(t)$ (and, in this example, $r(t)$ as well) continue to rise for some period following stimulus offset, before eventually falling.[10] Second, physical-stimulus offset time is not especially interesting or important. That is, the theory’s focus is on the *entire* $a(t)$ and $r(t)$ functions, not just those portions that follow stimulus offset. From this perspective, effects of independent variables like H1D or ISI on measures like ‘‘persistence duration’’ are concomitantly uninteresting; instead the theoretical focus is on the effects of such variables on the entire a(t) and r(t)

[10] It can be easily shown that, by the linear-filter model, the sensory-response function, $a(t)$, *always* rises for a brief time following stimulus onset (see Wolford, 1992). The information-extraction rate function $r(t)$ sometimes continues to rise (as in the Fig. 7 illustration) and other times does not, depending on the exact values of duration, contrast, and the theoretical parameters.

functions. Given this view, the traditional ''icon'' is just that part of the sensory response function that happens to be occurring after stimulus offset.

Information Acquisition as Phenomenological Appearance

To account for visible persistence data, including both temporal-integration and synchrony-judgment data, Loftus and his colleagues have assumed that *r(t) determines phenomenological appearance;* that is, a stimulus is phenomenologically present to the degree that $r(t)$ is high, and a stimulus phenomenologically disappears when $r(t)$ declines below some threshold.

Explanation of the inverse-duration effect. With this assumption, the inverse-duration effect found in both temporal-integration and synchrony-judgment paradigms is accounted for quite naturally. The key to the theory's account of these effects is that as duration increases, $r(t)$ becomes smaller at the time of stimulus offset and, accordingly, it takes less time for $r(t)$ to decline to any given criterion level.

Previous phenomenological-appearance hypotheses. Over the years, the default assumption has been that phenomenological appearance has been intimately tied to iconic decay which, it turn, has been assumed to be exponential (e.g., Di Lollo, 1984; Hawkins & Schulman, 1979). A small number of investigators, however, have explicitly set out to measure the shape of the phenomenological-appearance function. The most direct such measurements were reported by Weichselgartner and Sperling (1985) using a technique in which luminance of a test patch of light was matched to the brightness of a target stimulus at varying times following stimulus offset. The resulting observed functions were describable at a very general level (they both rose and fell gradually) but differed across individuals. Weichselgartner and Sperling concluded that ''[these observed functions] do not seem to derive from any generic function'' (p. 721). However, they do not appear to be inconsistent with the $r(t)$ functions that we have postulated.

Application to experiments 1 and 2. In Experiments 1 and 2, two separate stimuli (halves) were presented, each of some duration, separated by some ISI. The theory's depiction of ISI, H1D, and H2D effects on $r(t)$ are shown in Fig. 8. Consider first, Panel A, which depicts a ''base condition'' of 20 ms H1D, 0 ms ISI, and 20 ms H2D. Each half engenders its own $f(t)$, $a(t)$, and $r(t)$ function; it is the half-1 and half-2 r(t) functions that are shown in the figure (note that rectangles correspond to stimulus presentations, curves correspond to $r(t)$ functions, and solid and dashed lines represent half-1 and half-2 stimuli). We assume that the stimulus is seen as a unitary whole to the degree that the $r(t)$ functions corresponding to the two halves are *similar.* Such similarity can be quantitatively expressed in a variety of ways. The manner we chose was based on a technique proposed by Dixon and Di Lollo (1994): computation of the temporal correlation between the two $r(t)$ functions. In particular, we considered the $r(t)$ functions up to $t = 800$ ms (by which time they have essentially fallen to zero), divided that duration into

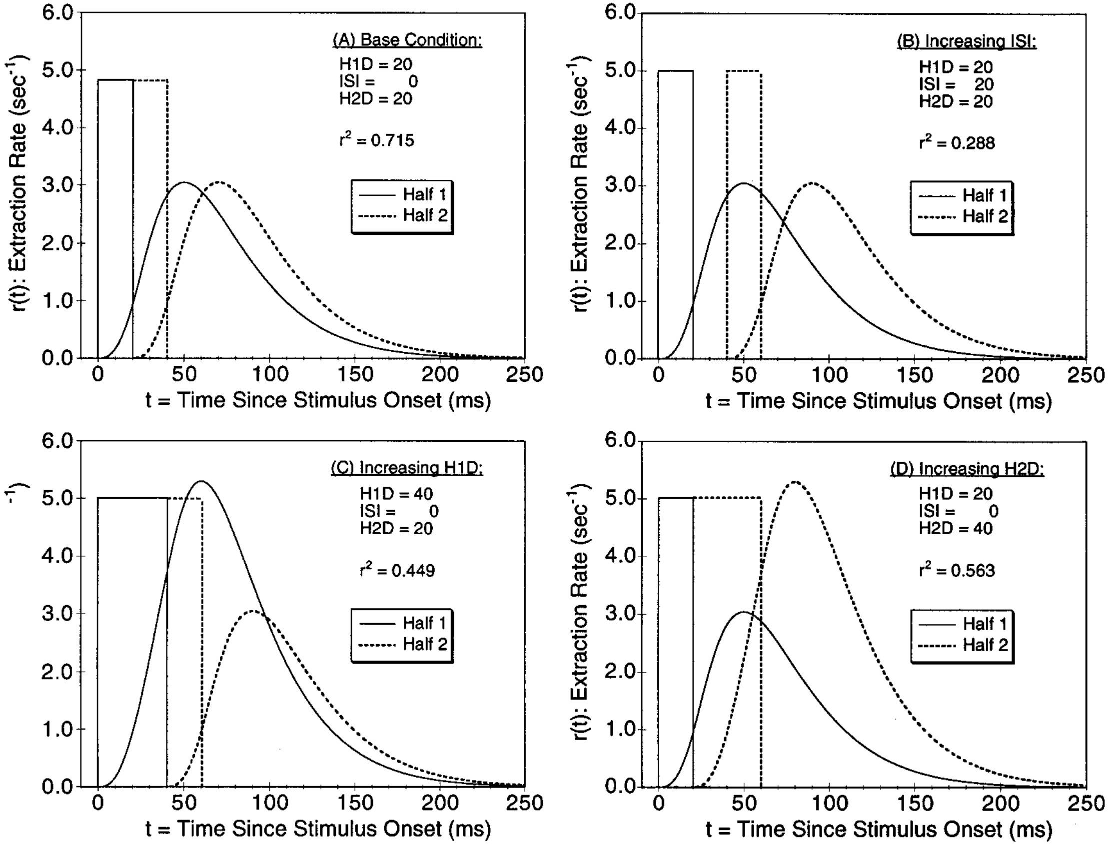

FIG. 8. Theoretical account of H1D, ISI, and H2D effects. Each panel represents $r(t)$, the determinant of phenomenological presence, to time since stimulus onset. Rectangles represent physical stimulus presentations. Solid lines corresponds to the first-half presentation, while dashed lines corresponds to the second-half presentation. Pearson r^2 values are correlations of the two $r(t)$ functions from 0 to 800 ms in 10-ms intervals.

80 10-ms intervals, computed the r(t) value for each half at each interval, and correlated the resulting value pairs. For the Fig. 8A condition, this correlation (Pearson r^2) is 0.715.

Performance—both temporal-integration and subjective-rating performance—was then assumed to be monotonically related to this Pearson r^2 measure. Figures 8B–8D show the effects on this correlation measure of the Experiment 1 and 2 manipulations: increasing ISI, H1D, and H2D, respectively. In all cases, the half-1/half-2 correlation decreases, as did both observed performance measures.

The parameter values, n, τ, and c, used to generate the curves in Fig. 8 (and Fig. 7) were the best-fitting values for the Experiments 1 and 2 data (described below). Two aspects of these theory predictions are of some interest. First, as shown in Fig. 3, the observed ISI effect is greater than the H1D effect (in the example, increasing ISI by 20 ms decreases performance more than increasing H1D by 20 ms); comparisons among Figs. 8A, 8B, and 8C indicate that this is mirrored by the theory. Second, as shown in Fig. 6, the observed H1D and H2D effects on performance are roughly equivalent; comparisons among Figs. 8C and 8D indicate that this is also mirrored by the theory. We note, however, that this rough equivalence of H1D and H2D effects come about as a result of the particular parameter values that we have chosen. Different parameter values would allow the theory to account for asymmetrical H1D/H2D effects such as those found by Dixon and Di Lollo (1994).

Application to Experiments 1 and 2

We fit the theory to Experiments 1 and 2 simultaneously. The goodness-of-fit measure was the over-condition rank-order correlation between the Pearson r^2 predicted by the theory, and the observed data measure (probability correct or mean rating).[11]

Rendering Experiment 1 and Experiment 2 data comparable. Although there are 50 conditions across the two experiments, five of the conditions (the five H1Ds corresponding to the Experiment-1 0-ms ISI level, and the five H1Ds corresponding to the Experiment-2, 20-ms H2D level) were common to the two experiments. In a perfect world, performance would be identical for the Experiment-1 and the Experiment-2 versions of these five conditions. However, there were in fact slight interexperiment differences for the five common conditions. Accordingly, we corrected the Experiment-2 data for each of the two tasks using the following algorithm. First, for each of

[11] It is easy to become confused by this proliferation of correlations. To forestall such confusion, we believe it is worthwhile to reiterate at this point that we are using correlational measures in two entirely different ways. First, as indicated in Fig. 8, the theory generates a Pearson r^2 as a measure of similarity between the two $r(t)$ functions engendered by the two stimulus halves: one such r^2 is generated for each experimental condition. The overall goodness of fit measure is then the *rank-order correlation,* over the 25 conditions, between the these Pearson r^2s and the observed data points.

the five common conditions (one condition corresponding to each of the 5 H1Ds), we computed the ratio of Experiment-1 to Experiment-2 performance. Second, we multiplied each of the five Experiment-2 data points for the H1D level corresponding to that condition by this ratio. Accordingly, we had 45 degrees of freedom across the 50 total Experiment-1 and Experiment-2 conditions.

Results. The theory has three parameters: n and τ, the two parameters entering into the $a(t)$ function (Equation 2); and c, the scaling parameter relating proportion acquired information to $r(t)$, the information acquisition rate (Equations 5 and 6). We carried out a gridsearch procedure on these three parameters using, as indicated above, the theory/data rank order correlation as the criterion fit measure. The best-fitting parameter values for all experiments are shown in Table 3. The overall best fit to Experiments 1 and 2 combined produced rank-order correlations of 0.993 and 0.999 for the objective (temporal-integration) and subjective (completeness-rating) measures. We also computed best fits for Experiments 1 and 2 individually, as shown in Rows 2 and 3 of Table 3. For the objective measure, the resulting correlations were 0.994 for Experiment 1 and 0.980 for Experiment 2. For the subjective measure, the individual correlations were higher: 1.000 for Experiment 1 and 0.982 for Experiment 2.

The theory's account of the subjective data is slightly better than its account of the objective data. As we suggested earlier, we believe that the subjective rating data, which are simple, direct, and not very prone to elaborate strategies, probably form the best measure of the observer's internal perception of the stimulus ensemble. The objective temporal-integration data, on the other hand, are almost certainly prone to idiosyncratic strategies (e.g., memorization strategies) that are not included in the theory. This is the sort of situation depicted in Fig. 1C. It is the use of such strategies that allows an observer to maintain high performance in the temporal-integration task even given an H1D or H2D at which perceptual integration has begun to deteriorate. Because the theory does not attempt to account for the use of these minor strategies, it does not fit the objective data quite as well as it fits the subjective data. Nonetheless, the theory fits both data sets very well, demonstrating that it accounts for the major determinants of performance in both tasks. In sum, the rating task and the temporal integration task appear to measure much the same thing, but subjects' strategies generate some minor performance differences between the two tasks.

Theory Application to Other Experiments

To demonstrate the generality of our theory's application to temporal integration, we apply it here to data reported by other investigators.

Dixon and Di Lollo (1994). Dixon and Di Lollo (1994) reported a temporal-integration task, similar to our Experiment 2, in which H1D and H2D were factorially combined. ISI was zero, and stimuli were presented against either a dim or a bright background. Of some interest is that Dixon and Di

TABLE 3
Best-Fitting Parameter Values for Experiments 1–7 along with Goodness of Fit (Theory-Data Rank-Order Correlation, ρ)

Experiment	Performance measure	Best-Fitting Parameter Values*						Model fit
		n	τ	c	c_p	Y	c/c'	(ρ)
1 and 2	**Temporal integration**	**2**	**210**	**3100**	—	—	—	**0.993**
(combined)	**Completeness ratings**	**2**	**120**	**2500**	—	—	—	**0.999**
1	**Temporal integration**	**2**	**200**	**3550**	—	—	—	**0.994**
(alone)	**Subjective ratings**	**2**	**320**	**2000**	—	—	—	**1.000**
2	**Temporal integration**	**2**	**420**	**3100**	—	—	—	**0.980**
(alone)	**Subjective ratings**	**2**	**240**	**2500**	—	—	—	**0.982**
3	Partial report	2	31	11.3	—	0.95	0.34	0.949
	Subjective ratings	**2**	**55**	**385**	**195**	—	—	**0.994**
4	Partial report	2	31	8.8	—	0.76	0.03	0.894
	Subjective ratings	**2**	**90**	**180**	**450**	—	—	**0.946**
5	Partial report	2	43	9.8	—	0.01	0.03	0.916
	Subjective ratings	**2**	**30**	**230**	**550**	—	—	**0.969**
6	Partial report	2	58	11.6	—	0.82	0.28	0.869
	Subjective ratings	**2**	**100**	**115**	**70**	—	—	**0.991**
7	Partial report	2	54	13.6	—	0.72	0.29	0.947
	Subjective ratings	**2**	**100**	**115**	**70**	—	—	**0.982**
Mean	Partial report	—	—	—	—	—	—	0.972
4–7	**Subjective ratings**	—	—	—	—	—	—	**0.974**

Note. Top number in each cell indicates value for objective test (temporal integration in Experiments 1–2; partial report in Experiments 3–7). Bottom number in each cell indicates value for subjective test (completeness ratings in all experiments). Temporal integration and subjective rating information is indicated in **boldface,** while partial report information is indicated in normal type.

[a]Parameters are n and τ: $a(t)$ parameters, c is overall speed of encoding, c_p is speed of encoding the probe in a partial-report task, Y is pre-probe encoding asymptote in a partial-report task, and c' is post-probe speed of encoding in a partial-report task.

Lollo's H1D and H2D values were generally higher than ours, ranging from 20–320 ms, in contrast to the 20–100 ms used in our Experiment 2. Is our theory sufficiently general to account for temporal integration performance with durations of this magnitude?

Figure 9, which is arranged in the same fashion as Fig. 5, shows Dixon and Di Lollo's data in the left panels. To fit our theory, we allowed four free parameters: n and τ, and two c values: one for each illumination value. The results are presented in the right panels of Fig. 9, and in Table 4. The fit is reasonable: $\rho = 0.951$. The ''Bright'' data contain some nonmonotonicities at short H2D values which, according to Dixon and Di Lollo were probably due to the brightness-matching technique that they used, and which are not captured by our theory. We do note that the c values that emerged from our fit (see Table 4) were 6200 and 4200 for the Bright and Dim conditions respectively. These values make sense in that the Dim condition would produce higher contrast and thus a faster information-acquisition rate—which is reflected by the smaller c value for the Dim condition.

Wolford (1992). Wolford (1992) reported a temporal-integration task in which H1D, ISI, and H2D were factorially varied. In addition, Wolford's observers viewed stimuli either monocularly or binocularly; accordingly there were four independent variables.

Wolford's data are shown in the left panels of Fig. 10. As expected, H1D, ISI, and H2D all had negative effects on temporal-integration performance. Less expected is that binocular performance is *worse* than monocular performance. To fit our theory, we assumed simple ocular additivity: that the binocular $a(t)$ function is the monocular $a(t)$ function multiplied by 2 (see Busey & Loftus, 1998, for a justification of this assumption). As indicated in the right panels of Fig. 10, the theory captures the main effects of all independent variables, including the monocular-viewing superiority. The best-fitting parameter values, along with the fit ($\rho = 0.950$) are in Table 4. Intuitively, the theory's account of the negative binocular effect is as follows. With greater magnitude binocular $a(t)$ functions, information is extracted at a faster rate and accordingly, the $r(t)$ functions peak earlier. This causes the $r(t)$ functions to be narrower and more distinct from one another, i.e., less highly correlated.

EXPERIMENTS 3–7: PARTIAL REPORT AND SUBJECTIVE-COMPLETENESS RATINGS

Experiments 3–7 were designed to compare subjective-completeness performance in the rating task with accuracy performance in a partial-report task. In these experiments, a stimulus array was presented for some duration, H1D, followed by a blank ISI, followed by a 20-ms visual probe indicating one of the array's letters. On some trials observers performed a standard partial-report task and attempted to report the probed letter. On other trials, they rated the subjective completeness of the ensemble formed by the stimu-

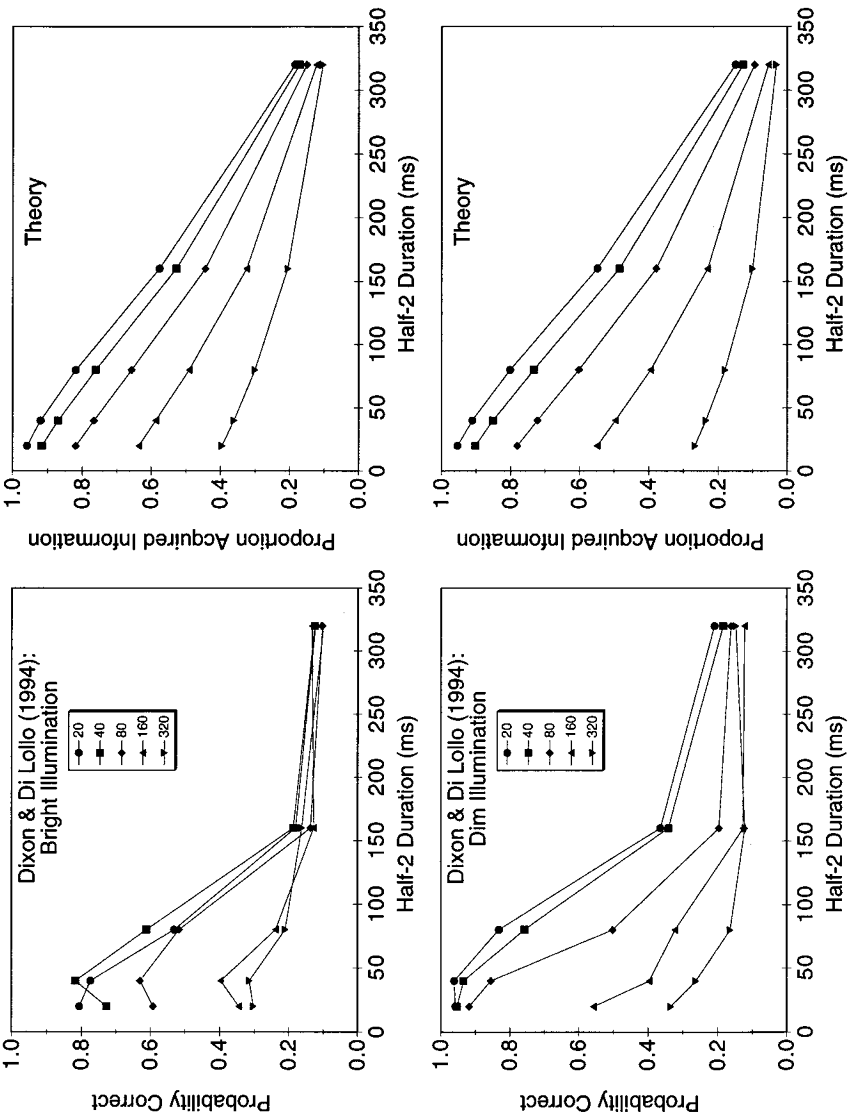

FIG. 9. Data from Dixon and Di Lollo (1994). Top and bottom panels show data from dim and bright background conditions. Left panels show data; right panels show theoretical output.

TABLE 4
Best-Fitting Parameters and Theory Fit for The Dixon and Di Lollo (1993) and Wolford (1992) Data

Dixon and Di Lollo (1993) Data				
n	τ	*c*(Bright)	*c*(Dim)	Model fit
2	245	6200	4200	$\rho = 0.951$
Wolford (1992) Data				
n	τ	*c*		Model Fit
2	130	202		$\rho = 0.940$

lus array and the probe. The H1D and ISI ranges were identical to those of Experiments 1 and 2.

Experiment 3

Experiment 3 differed from Experiments 4–7 in that brightness compensation—endowing shorter-duration stimuli with higher luminance to achieve equal subjective brightness—was not used in Experiment 3, but was used in Experiments 4–7. Accordingly, we first describe Experiment 3, and we then go on to describe Experiments 4–7 together.

Method

Experiment 3 (and Experiments 4–7 as well) used largely the same methodology as Experiment 1. The two main changes were the following. First, instead of half-1 and half-2 dot stimuli, there was a letter array and a bar probe. Second, the objective task was partial report rather than temporal integration. As in Experiments 1 and 2, the subjective task was a completeness rating of the degree to which the letter array and the probe constituted one or two distinct temporal events.

Observers. Ten undergraduate and graduate students at Michigan State University participated in Experiment 3. None had participated in Experiments 1 or 2, and none had any knowledge of the experimental hypotheses. All reported normal or corrected-to-normal vision, and each was paid $30 for participating in 6 sessions.

Stimuli. Fifty different letter arrays were used as stimuli. Each array contained 10 letters in a 2 rows $\times$ 5 columns format. The letters were drawn randomly from the set of all consonants, excluding Y. A short vertical line appearing above (top row) or below (bottom row) one of the array locations was used as the partial-report cue.

Apparatus. Stimuli were displayed on a Tektronix 608 x-y oscilloscope (P15 phosphor) driven by a Digital Equipment Corporation Micro-11/23+ computer through digital to analog converters. The computer also recorded observer responses entered into the terminal keyboard. Observers viewed the display from a distance of 40 cm, set by a chin rest. At this viewing distance the oscilloscope subtended 16.7 deg of visual angle horizontally and 13.4 deg vertically. The letter arrays subtended 3.2 deg horizontally and 2.0 deg vertically. Each letter subtended 0.32 deg horizontally and 0.5 deg vertically; the letters were separated horizontally by 0.4 deg and vertically by 1 deg. The vertical bar probe was 0.25 deg high and 0.04 deg wide; it was presented 0.25 deg above or below the location of the probed letter.

The experimental chamber was dimly illuminated during the experiment. The P15 phosphor

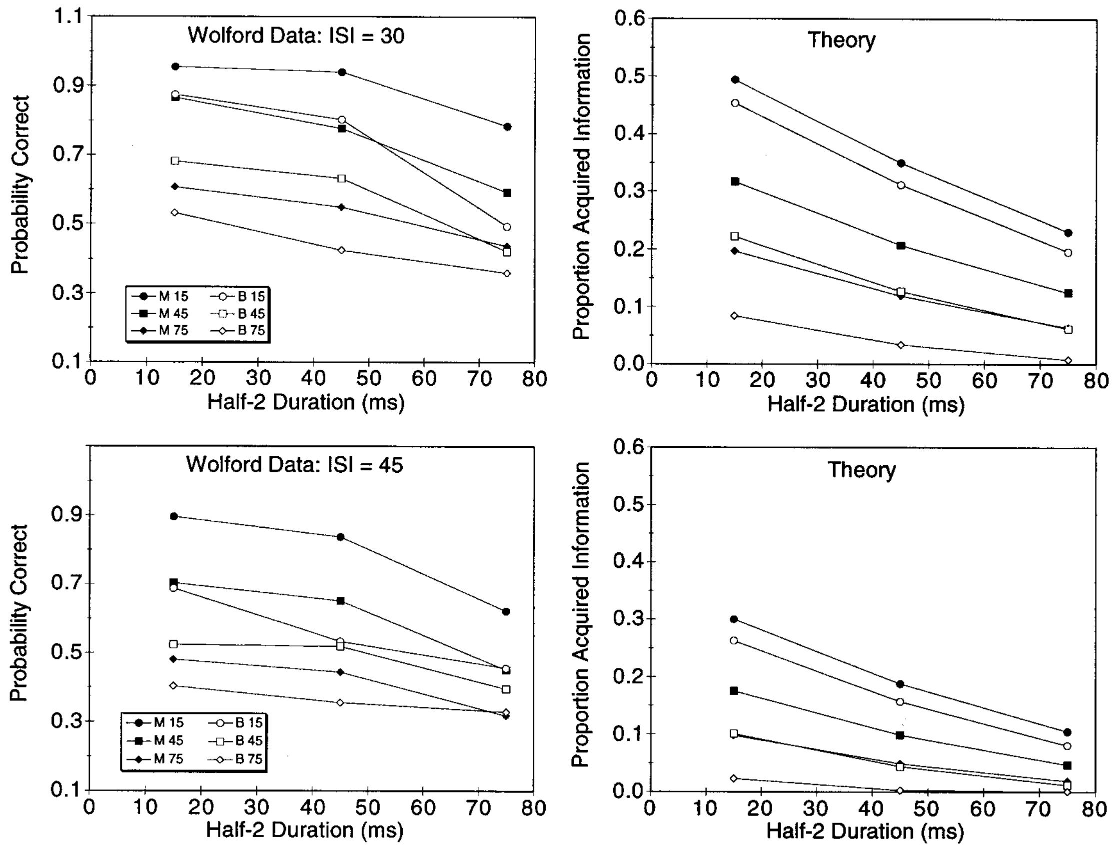

FIG. 10. Data from Wolford (1992). Top and bottom panels show data from ISI = 30 and ISI = 45 conditions. Left panels show data; right panels show theoretical output.

has no detectable persistence (Groner, Groner, Muller, Bischof, & Di Lollo, 1993), however, so no filters were placed over the display scope. The luminance of the display background was 2 cd/m^2; stimulus displays were presented with an effective luminance[12] of 21 cd/m^2. Shutter tests similar to those described by Irwin, Jonides, and Yantis (1983) confirmed that no phosphor persistence was visible 5 ms after stimulus offset.

Procedure and design. The sequence of events on each trial was as follows. After the observer pressed the return key on the terminal keyboard to begin the trial, a fixation point (a small plus sign centered in the middle of the display) was presented for 500 ms; this was followed 500 ms later by the 2 × 5 letter array.

Following the ISI, the bar probe was presented for 20 ms, centered above (top row) or below (bottom row) one of the letter locations. Following bar-probe presentation, the observer entered his or her response into the terminal keyboard. As in Experiments 1 and 2, observers carried out both an objective and a subjective task. In the objective (partial-report) task, the observer attempted to report the letter that had appeared in the probed position. In the subjective-completeness task, the observer provided a rating ranging from 1 to 4 of how temporally integrated the ensemble formed by the letters and the bar probe appeared to be. As in Experiments 1 and 2, observers were instructed that a rating of ''4'' meant that the letters and the bar probe appeared to be simultaneously present, whereas a rating of ''1'' meant that the letters and the bar probe appeared to be completely separated in time. Ratings of ''2'' and ''3'' were used for intermediate perceptions. No feedback was provided.

It was our intent to have the same set of H1Ds (20 − 100 ms) and ISIs (−20–60 ms) as used in Experiment 1. However, because of technical complications the actual values of H1D, ISI, and H2D were slightly different from the corresponding intended or nominal values. The details of these complications, as well as the actual H1D, ISI, and H2D values, are provided in Appendix B.

Each observer completed 6 sessions containing two blocks of trials; in one block the observer performed the partial-report task and in the other the observer performed the subjective-completeness task. The order in which these tasks were performed alternated over sessions. Half of the observers began their initial block with the partial-report task and half began with the subjective-completeness task. The first session consisted of 100 practice trials in each task; these data were discarded. The remaining sessions contained 250 trials in each task. Thus, over the course of the experiment, each observer completed 1250 partial-report trials and 1250 subjective-completeness trials. These 1250 trials consisted of 50 replications of the 25 conditions formed by the factorial combination of stimulus exposure duration and cue delay. The observers saw each of the fifty 10-letter arrays once in each condition. Stimulus exposure duration (H1D) and cue delay (ISI) were sequenced randomly over trials, but observers saw the same arrays in the same order under the same timing conditions in the two experimental tasks. The bar probe appeared equally often at each letter location in each of the two tasks over the course of the experiment.

Results

Figure 11, which is organized like Figs. 2 and 5, shows the main results of Experiment 3. The subjective (rating) data, shown in the bottom-left panel, look very much like the Experiment-1 rating data: there are strong effects of both H1D and ISI. The objective (partial-report) data are somewhat different. The usual negative ISI effect is evident, but there appears to be very

[12] Effective luminance was the luminance obtained from a stimulus consisting of a 13 × 13 pixel array displayed continuously. The screen luminance remained the same, at this effective luminance throughout Experiment 3. However, due to other constraints to be described below, actual luminance was reduced because the display was not present 100% of the time.

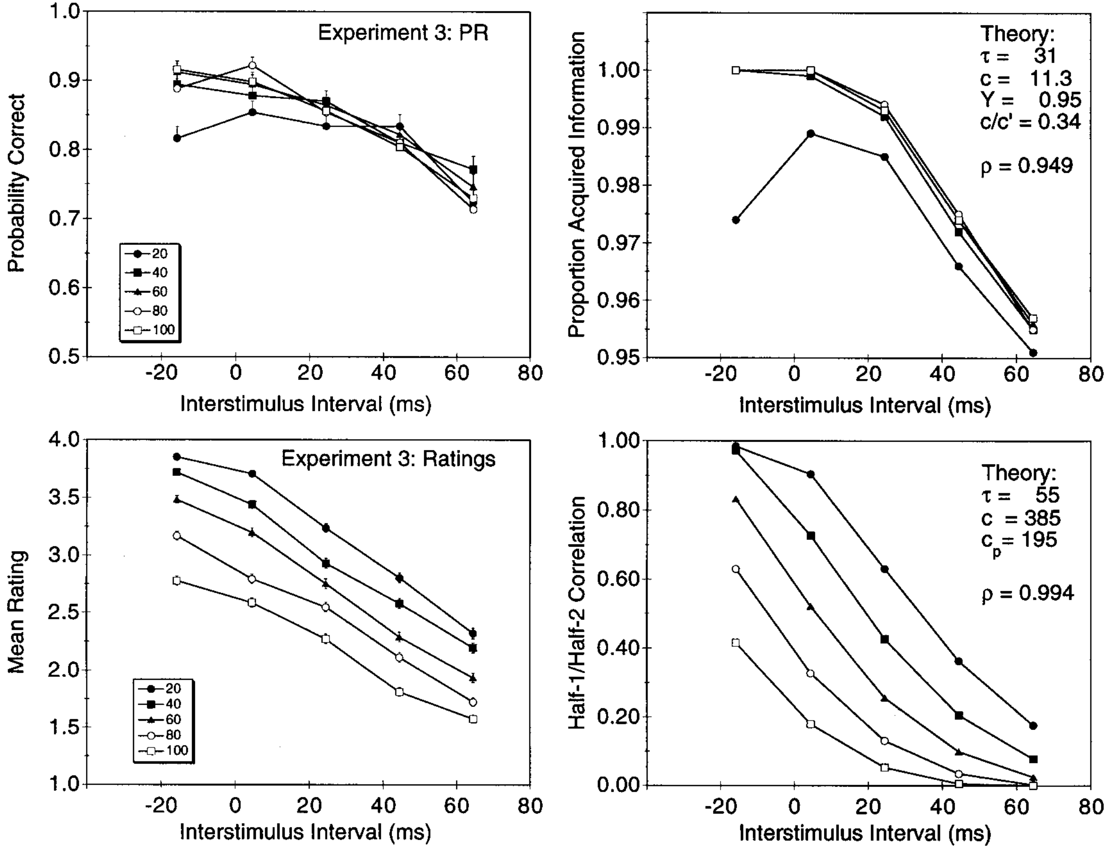

FIG. 11. Experiment 3: Main results. Figure 3 is organized like Fig. 2, except that objective measure is partial report rather than temporal integration.

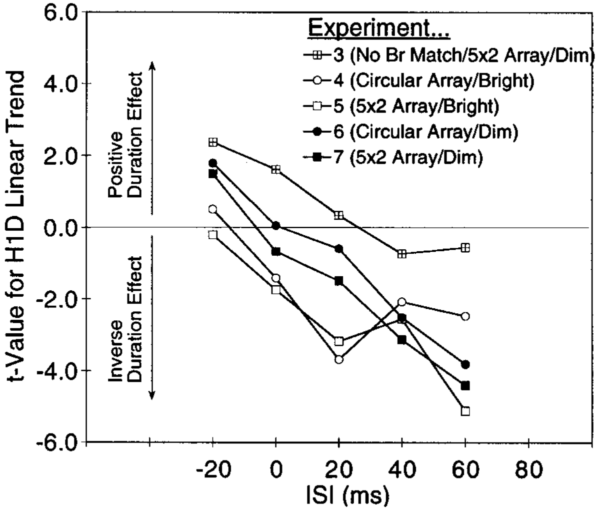

FIG. 12. Experiments 3–7: Measures of the inverse-duration effect. Each curve corresponds to a given experiment, with curve symbols carefully chosen to be visual mnemonics for the experiment's array-shape and luminance status. The measure is the *t* value corresponding to linear trend for partial-report performance across H1D at each ISI level. Positive and negative *t* values indicate positive and negative duration effects on performance.

little effect of H1D; indeed, if anything, there appears to be a positive duration effect, at least at the short ISIs, thereby replicating similar positive-duration effects found by Irwin and Yeomans (1986a) and Irwin and Brown (1987).

Thus, at first glance, it appears that partial report and completeness ratings behave quite differently. This will prove to be the case for subsequent glances as well. The extension of our theory, which generated the results shown in the right panels, will be described shortly.

The inverse-duration effect. To quantify the nature of the H1D effect for the objective measure (partial-report in Experiment 3), we computed a $t(36)$-value for H1D linear trend at each ISI value. This value is positive or negative corresponding to the sign of the H1D-performance relation.[13] The *t* value, plotted as a function of ISI, is shown as the top curve in Fig. 12. The H1D effect is positive for the first three ISI values, and becomes slightly negative for the final two ISI values. (The remaining curves in Fig. 12 are discussed at greater length in a later section.)

Di Lollo and Dixon (1988; 1992a) reported robust negative relationships

[13] Normally, the effect of linear trend would be expressed as an $F(1,df\mathrm{I})$ value where $df\mathrm{I} = 36$ is degrees of freedom in the subject-by-H1D interaction. However, the F value is indifferent to the sign of the linear trend. The *t* value, which is the square root of the F value, is positive or negative depending on the direction of the trend.

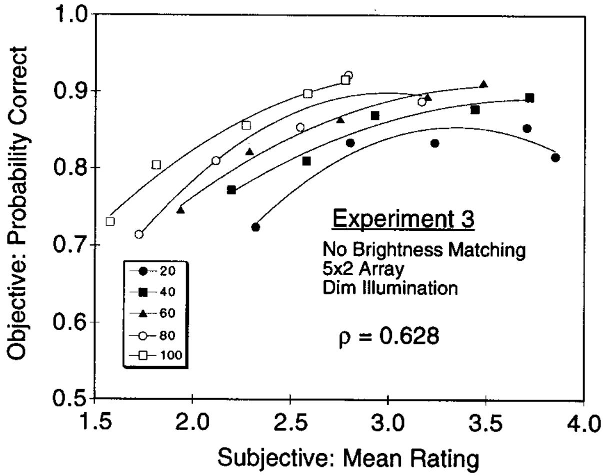

FIG. 13. Experiment 3: Correlations between objective measure (ordinate) and subjective measure (abscissa) over 25 experimental conditions. Figure 13 is organized like each panel of Fig. 4 except that objective measure is partial report rather than temporal integration.

between H1D and performance in a partial-report task (henceforth called an inverse duration effect, or IDE). We found no IDE in Experiment 3. There are several possible (non-exclusive) reasons for this apparent conflict. First, identifying a single letter in our 5×2 array may not have been spatially demanding; according to Di Lollo and Dixon (1988), high spatial demand may be a necessary condition for obtaining an IDE. Second, we did not use brightness compensation; i.e., stimulus luminance was identical for all H1D values. Finally, we used relatively short durations (20–100 ms) and relatively dim stimuli, while Di Lollo and Dixon used substantially longer times (on the order of 20–400 ms) and substantially brighter stimuli. In Experiments 4–7 we examine (to some degree) which of these variables is critical for obtaining an IDE in partial report.

Intermeasure correlation. Figure 13, which is organized like each panel of Fig. 4, shows the correlation between the objective and subjective performance measures. The correlation, $\rho = 0.628$, is quite low. In contrast to the corresponding correlations in Experiments 1 and 2, wherein the two performance measure ranks shared over 95% of the variance, the corresponding value in Experiment 3 is less than 40%. This low correlation does not come as a surprise: it was evident in Figure 11 that there was a large IDE associated with the subjective measure, but no IDE associated with the objective measure. The degree that the two measures are correlated reflects the usual inverse-ISI effect found in both.

Discussion

Experiment 3 is conceptually very similar to Experiment 1. Both experiments involved (essentially) the same ranges of H1D and ISI. In both, an objective and a subjective measure were obtained. The subjective measure—completeness rating—behaved very similarly across the two experiments. To capture this similarity, Table 5 shows the interexperiment rank-order correlations (across the 25 conditions) for Experiments 1 and 3–7. The top right entries are for the objective measures (temporal integration and partial report) while the bottom left entries are for the subjective measure (rated completeness).

The Experiment 1–Experiment 3 rank-order correlation for the rating data is $\rho = 0.994$ (Table 5, first column, second row), indicating that the subjective ratings measure much the same thing in the two experiments, even though the stimuli were entirely different. In contrast, the correlation for the objective data is much lower: $\rho = 0.683$ (second column, first row). The reason for this lowered interexperimental similarity is that in Experiment 3 the usual IDE effect vanished. It is clear that the IDE, which is powerful and robust for any visible-persistence measure—synchrony judgment, temporal integration, and completeness rating—is not robust at all with respect to partial report. This confirms Coltheart's (1980) assertion that partial report and temporal integration measure different perceptual entities, as depicted in Figs. 1B and 1C.

Theory

In this section, we return to the theory we described earlier and extend it to make predictions about the partial-report and subjective-completeness data from Experiments 3–7.

Application to Completeness Ratings

The theory accounts for the subjective rating data largely as described earlier, with one exception. Recall that the parameter c in the theory (Equation 3) is essentially a measure of how fast information is acquired from a given stimulus: the lower is c, the less time is required to acquire any given proportion of stimulus information. In the dot-matrix tasks, the two stimulus halves were qualitatively similar; accordingly we assumed that the c values were the same for both halves. Such is not necessarily the case for Experiments 3–7, however: the letter array and the probe are quite different stimuli, and we have no reason to assume them to be processed at the same rate. Accordingly, we assume a second parameter, c_p to characterize speed of probe processing. The $r(t)$ equation corresponding to the stimulus array is, then, the same as given in Eq. (3). However, the $r(t)$ equation corresponding to the probe is as given in Eq. (6) except that a new parameter, c_p, is substi-

TABLE 5

Across-Experiment Intercorrelations (Spearman ρ) for Experiments 1 and 2–7

	Objective measures (above the diagonal)						
Experiment	1	3	4	5	6	7	4–7 (mean)
1	—	0.683	0.875	0.930	0.748	0.836	0.911
3	0.994	—	0.690	0.672	0.822	0.864	0.795
4	0.902	0.929	—	0.893	0.772	0.811	—
5	0.932	0.947	0.976	—	0.700	0.810	—
6	0.958	0.975	0.980	0.986	—	0.906	—
7	0.980	0.990	0.958	0.972	0.988	—	—
4–7 (mean)	0.943	0.960	—	—	—	—	—

Note. Subjective measures (below the diagonal).

Note. Entries above the Minor Diagonal are for Objective Measures (Temporal Integration and Partial Report) while Entries below the Minor Diagonal are for Subjective Measures (Completeness Ratings).

tuted for *c*. Accordingly, the theory as applied to the Experiment 3–7 rating data has four free parameters: *n, t, c,* and c_p. As before, we assume rating performance to be monotonically related to the resulting half-1/half-2 correlation.

Application to Partial Report

The theory must be extended in a somewhat more complex fashion to be able to predict partial-report data. Such an extension has been sketched by Loftus and Busey (1992), and we describe it in detail here.

As noted earlier, information is acquired at a rate $r(t)$. When the stimulus is to be later remembered, (e.g., when it is a letter array), it is assumed that acquired information is placed into short-term memory where it can serve as a basis for subsequent recall. In a partial-report task, we must consider separately the information acquisition that occurs prior to probe presentation vs. the information acquisition that occurs subsequent to probe presentation.

Asymptotic pre-probe information acquisition. Prior to probe presentation, information acquisition can, at best, be carried out from random locations in the stimulus array. With typical arrays (e.g., the 10-item arrays used in Experiment 3), it is not possible to acquire all stimulus information because of the well-known short-term capacity limit of 4–5 letters (e.g., Sperling, 1960, 1963, 1967). Accordingly, we assume that, prior to probe presentation, information is acquired randomly from the array, but that the proportion of acquired information approaches some asymptote, $Y < 1.0$, which is a free parameter. The probability that the eventually probed letter is acquired prior to the probe, p_r, is equal to,

$$p_r = Y[1.0 - e^{-\mathrm{A(0,\ H1D+ISI)}/c}] \tag{6}$$

where A(0, H1D + ISI) is the area under the $a(t)$ function up to the point at which the probe occurs (i.e., the area between time $t = 0$ and time $t =$ (H1D + ISI).

Remark. The random-acquisition process we have just described is by no means a process that universally occurs. Various investigators, beginning with Sperling (1960), have shown that observers can and do use many different strategies in a partial-report task. For instance, an observer may sample randomly from the array prior to probe presentation (as assumed in the discussion above); alternatively, there may be systematic biases in which locations are sampled from, such as those near the fixation point (e.g., Gegenfurtner & Sperling, 1993). The strategy that is used in any given experiment is probably determined fairly idiosyncratically, but may also dramatically affect the pattern of results. For this reason, partial report is a difficult task to model in any simple and universally applicable fashion.

Post-probe processing. Once the probe has signaled the to-be-reported array letter, the observer can concentrate fully on it rather than acquiring information randomly from the array. Because full concentration is presumably on a single letter, two changes in processing occur. First, because a single letter is within the bounds of short-term memory, information acquisition from the target letter can now asymptote at 1.0. Second, the information can be acquired at a faster rate. The pre-probe rate parameter is c (see Equation 3 above); we assume the corresponding post-probe parameter to be c', another free parameter. The probability that the target letter is identified subsequent to the probe, p_s, is

$$p_s = 1.0 - e^{-A(\text{H1D} + \text{ISI}, \infty)/c'}, \tag{7}$$

where A(H1D + ISI, ∞) is the area under the $a(t)$ function from the time the probe occurs (H1D + ISI) to infinity.

Predicting partial-report performance. The probabilities p_r and p_s (Equations 5 and 6) are assumed to be independent; accordingly as also assumed by others (e.g., Averbach & Coriell, 1961; Averbach & Sperling, 1961; Di Lollo and Dixon, 1988), the overall probability of encoding the probed letter, p, is obtained by probability summation:

$$p = p_r + (1.0 - p_r)p_s \tag{8}$$

If we were studying only partial report, we would simply assume that p in Eq. (8) is the predicted partial-report performance value for a given condition, and fit the theory accordingly. However, to provide a more direct comparison with the theory's fit to the rating data, we make the weaker assumption that observed partial-report data is only monotonically related to p (i.e., we make the same assumption about theory-data correspondence for both measures). Accordingly, our goodness-of-fit measure for both rating and partial report is, as in Experiments 1 and 2, the over-condition rank-order correlation, ρ, between predicted and observed measures.

A remark is in order about the pre-probe/post-probe independence implication embodied in Equation 8. Although, as we have noted, it has been assumed by some investigators, it has been explicitly disconfirmed by partial-report data reported by Gegenfurtner and Sperling (1993). We incorporate independence because it is directly implied by the random-sampling assumptions of our theory. As we will see shortly, our theory does not fit partial-report data especially well. Part of its problem in this domain may lie in the independence assumption.

Experiment 3: Theory Fit

To predict the rating data, the theory has four free parameters: n, τ, c, and c_p. To predict the partial-report data, the theory has five free parameters: n,

τ, c, Y, and c'. For clarity of exposition, we report the ratio, c/c' rather than simply c'. Because a smaller c (or c') corresponds to faster information processing, a c/c' value greater than 1.0 would correspond to an increase in post-probe processing rate, while a value less than 1.0 would correspond to a decrease in post-probe processing rate.

The right panels of Fig. 11 show the best-fitting predicted parameter values. The partial-report fit was reasonable ($\rho = 0.949$) whereas the rating fit was somewhat better ($\rho = 0.994$). The best-fitting parameter values are shown in Table 3. Several remarks about these fits apply to all fits for Experiments 3–7. First, for both measures, the parameter space was quite flat, producing many ties in the criterion measure, ρ. In particular, while there is probably no parameter value set that produces a *higher* ρ value than the best fitting ones that we report, there were many sets that produced equal values, or values that differ only in the 3rd decimal place. In this sense, the reported best-fitting parameter set is somewhat arbitrary. Second, the best-fitting value of n was 2 throughout.[14] Third, c was estimated to be much lower for the partial-report than for the rating data. Ideally, this difference should permit the conclusion that processing is much slower when the observer is trying to make a rating vs. encoding the letters in the array. However, because the theory application is so different for the two measures, such a conclusion (or any conclusion based on quantitative comparisons among the parameters for ratings vs. report) are quite weak. Finally, the c/c' ratio is less than 1.0. This replicates a finding reported by Loftus and Busey (1992), who applied this same theory to several data sets reported by Di Lollo and Dixon (1988, 1992b). Assuming the theory's validity, the counterintuitive conclusion is that processing rate decreases (i.e., c increases) when processing switches from the entire array to a single letter.

Theory: Discussion

Even though the subjective-completeness task in Experiment 3 was quite different from that of Experiments 1 and 2, the same theory fit the resulting data quite well, thereby increasing confidence in the theory's generality in accounting for the subjective experience of temporal completeness.

In contrast, while the theory can account for the qualitative pattern of partial-report results (see also Loftus & Busey, 1992), the theory's quantitative fit is less good. The rank-order correlation between theory and data is somewhat lower, and the estimation of a less-than-1.0 c/c' ratio is entirely inexplicable; one would expect at the very least that an adequate partial-report theory would prophesize an increased processing rate on a single target letter relative to a letter array whose target letter is unspecified.

[14] Although the best-fitting value of n was 3 for the fit of Experiments 1 and 2, the best fit constraining n to be 2 produced only a marginally lower value of ρ.

TABLE 6
Organization of Experiments 4–7

Luminance Level	Array Type: Circular (15 characters)	Array Type: 5 × 2 (10 characters)
Bright (Effective Luminance = 171.0)	Experiment 4	Experiment 5
Dim (Effective Luminance = 21.0)	Experiment 6	Experiment 7

Note. In each experiment, the letter array is followed by a bar probe indicating a single to-be-reported letter.

Experiments 4–7

Experiments 4–7 were all very similar to Experiment 3 and to one another. Table 6 provides their organization. Experiments 4–7 were run primarily to generalize the empirical and theoretical findings that emerged from Experiment 3. As indicated in Table 6, the four experiments can be conceptualized as occupying the four cells of the 2 × 2 array defined by two levels of stimulus luminance (referred to as ''Bright'' and ''Dim'') and two array types (a circular 15-letter array and a rectangular, 5 × 2, 10-letter array).

Methods

Experiments 4–7 were essentially identical except for the manipulations shown in Table 6.

Observers. Six undergraduate and graduate students at Michigan State University participated in Experiments 4–7. None had participated in any of Experiments 1–3, and none had any knowledge of the experimental hypotheses. All reported normal or corrected-to-normal vision, and each was paid $120 for participating in 6 sessions of each experiment.

Stimuli and apparatus. The same stimulus patterns and apparatus used in Experiment 3 were used in Experiments 4–7.

Procedure and design. The procedure was very similar to that used in Experiment 3, the principal difference being that brightness compensation was used; i.e., effective luminance was greater for shorter exposure durations. This manipulation was accomplished by varying the stimulus refresh timing. As discussed earlier, the manipulation produced functional durations that were somewhat different from the nominal durations as described in Appendix B. All other procedures, including counterbalancing procedures, were as in Experiment 3.

Results

There were only minor differences among the results of Experiments 4–7; accordingly we discuss the four experiments in concert. The main results are shown in Figures 14–17, which are organized like Figures 2, 5, and 11: objective measure on the top and subjective measure on the bottom; data at the left and theory on the right. As in Experiments 1–3, the rating data are very clean and show rather large inverse-duration effects as well as inverse-

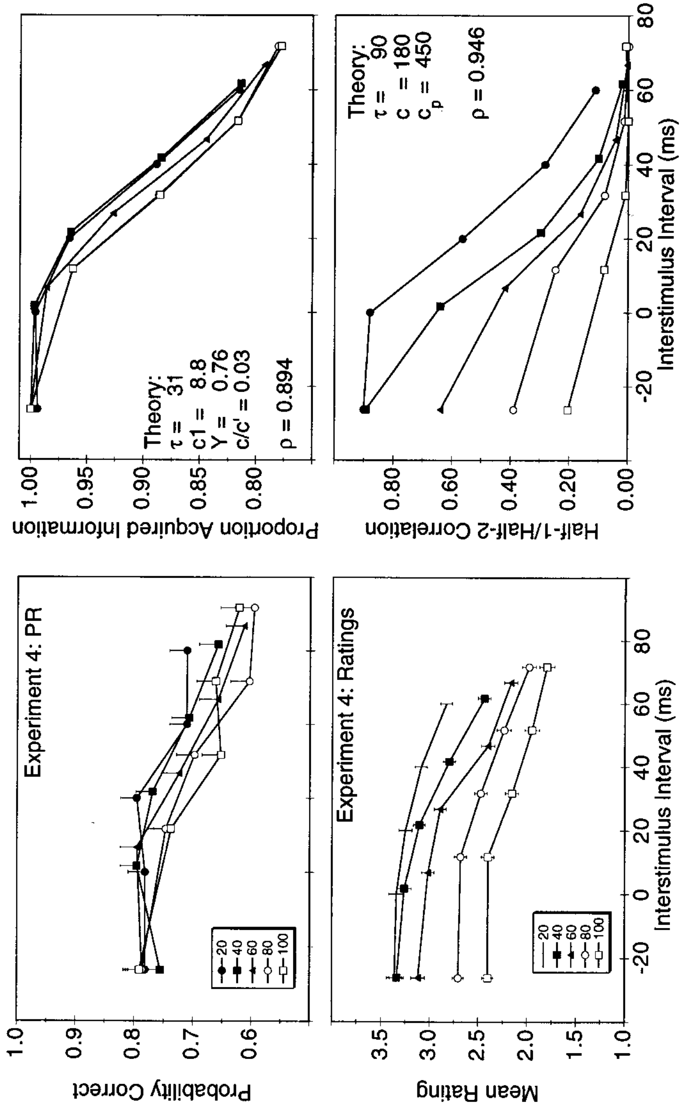

FIG. 14. Experiment 4: Main results.

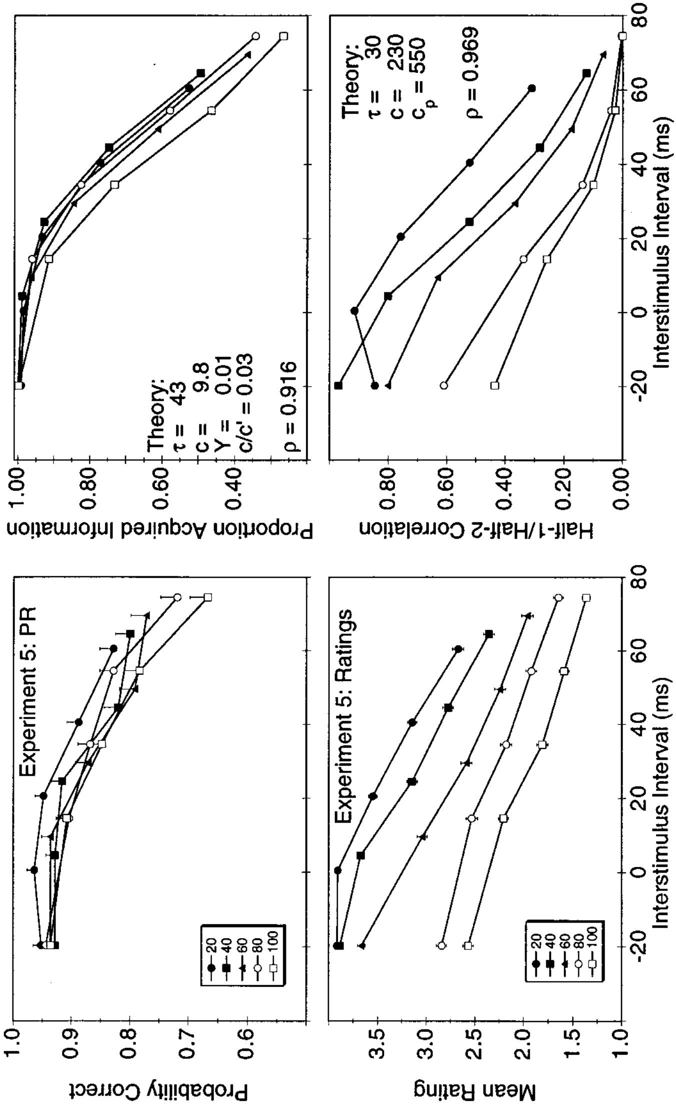

FIG. 15. Experiment 5: Main results.

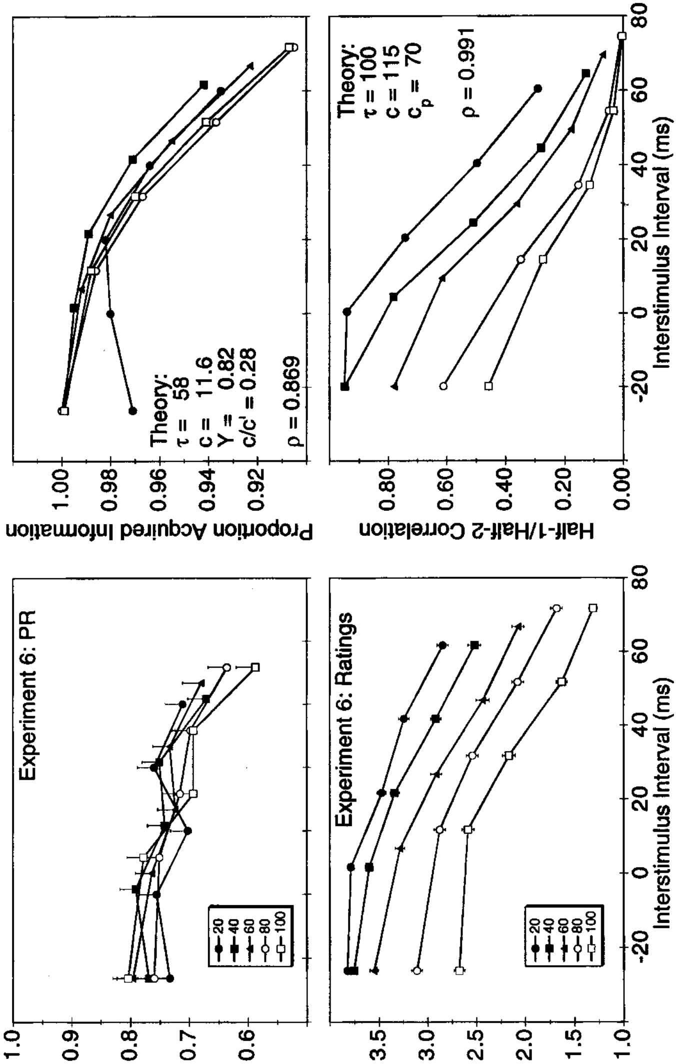

FIG. 16. Experiment 6: Main results.

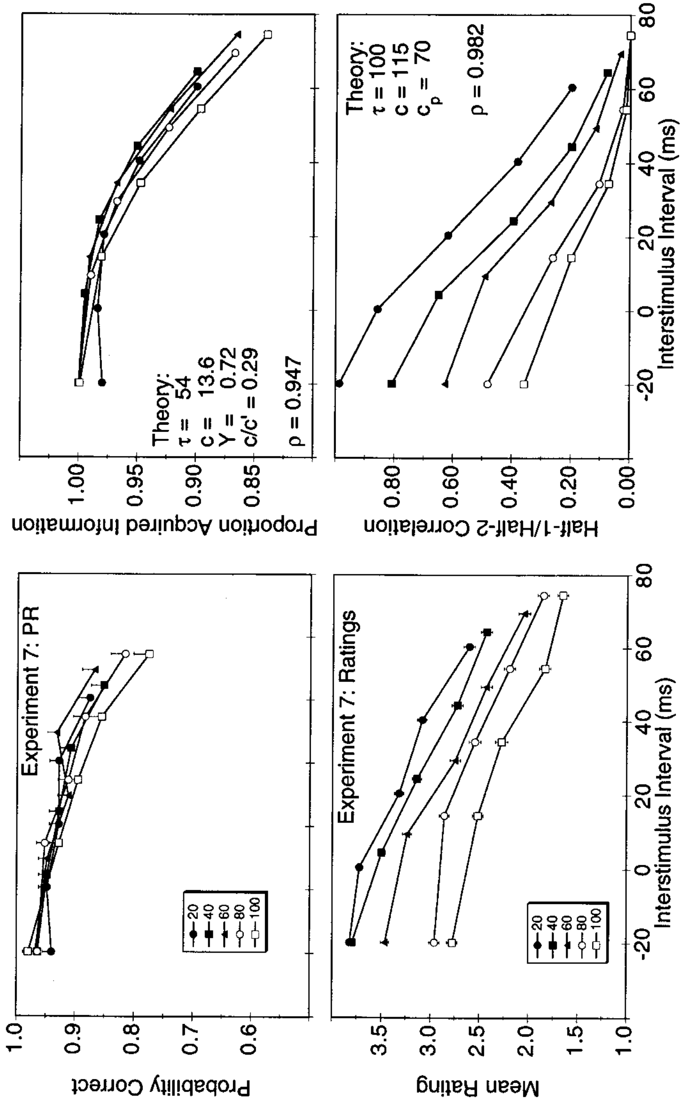

FIG. 17. Experiment 7: Main results.

ISI effects. The partial-report data are somewhat noisier. There is the usual inverse-ISI effect and, in addition, an inverse-duration effect emerges to some degree

The inverse-duration effect for objective measures. Figure 12 shows a measure of the IDE—the *t* value for a linear duration effect at each ISI—for Experiment 1 temporal-integration performance (bottom curve) and Experiments 3–4 partial-report performance (top five curves). The degree to which a given experiment produces an IDE can be identified with the overall height of the curve: the lower the curve, the greater the magnitude of the IDE. Note that there is a dramatic separation between the temporal-integration curve (which indicates a large IDE at all ISIs) and the partial-report curves (which indicate generally smaller IDEs or positive-duration effects). The five partial-report curves are somewhat variable. As the top curve indicates, there was a positive-duration effect in Experiment 3 at short ISIs, and essentially no duration effect at all at longer ISIs. The same general pattern obtains for Experiments 4–7, in which the duration effect shifts from more positive to more negative across ISI. As is evident in Fig. 12, there is no strong pattern across Experiments 4–7. Generally speaking, however, bright displays appear to lead to a greater IDE than dim displays, and the 5×2 (10-letter) displays appear to produce marginally stronger IDEs than circular (15-letter) displays.

A comparison of the Figure 11, Experiment 3 curves with the corresponding curves from Experiments 4–7 supports the proposition that the brightness compensation used in Experiments 4–7 was partially responsible for the IDE effect. This agrees with similar conclusions that issue from similar findings reported by Di Lollo and Dixon (1992a).

Objective/subjective comparisons. Figure 18, which is analogous to Figs. 4 and 13, shows objective/subjective scatterplots (and ρ values) for Experiments 4–7. As would be expected, given the observed IDE in partial report, these correlations are higher than the corresponding Experiment 3 correlation (of $\rho = 0.628$). However, they are considerably below the corresponding Experiment 1 and 2 correlations (of 0.977 and 0.979; see Fig. 4 and Table 2). It is clear, in short, that while completeness rating and partial report overlap in what they measure ($\rho >> 0.0$), they also differ systematically ($\rho << 1.0$).

Theory application to Experiments 4–7. The theory was applied to Experiments 4–7 in the same way as it was applied to Experiment 3. The right panels of Figures 14–17 show the theory's predictions for both the objective and subjective tasks, along with the rank-order goodness-of-fit measures. As in Experiment 3, the theory accounts for the subjective rating data quite well (ρ's ranging from 0.946–0.991), but for the objective partial-report data less well (ρ's ranging from 0.869–0.947). The parameter values are included in Table 3. As noted earlier, the parameter space was quite flat for all experiments, and accordingly, the best-fitting parameter values cannot be taken

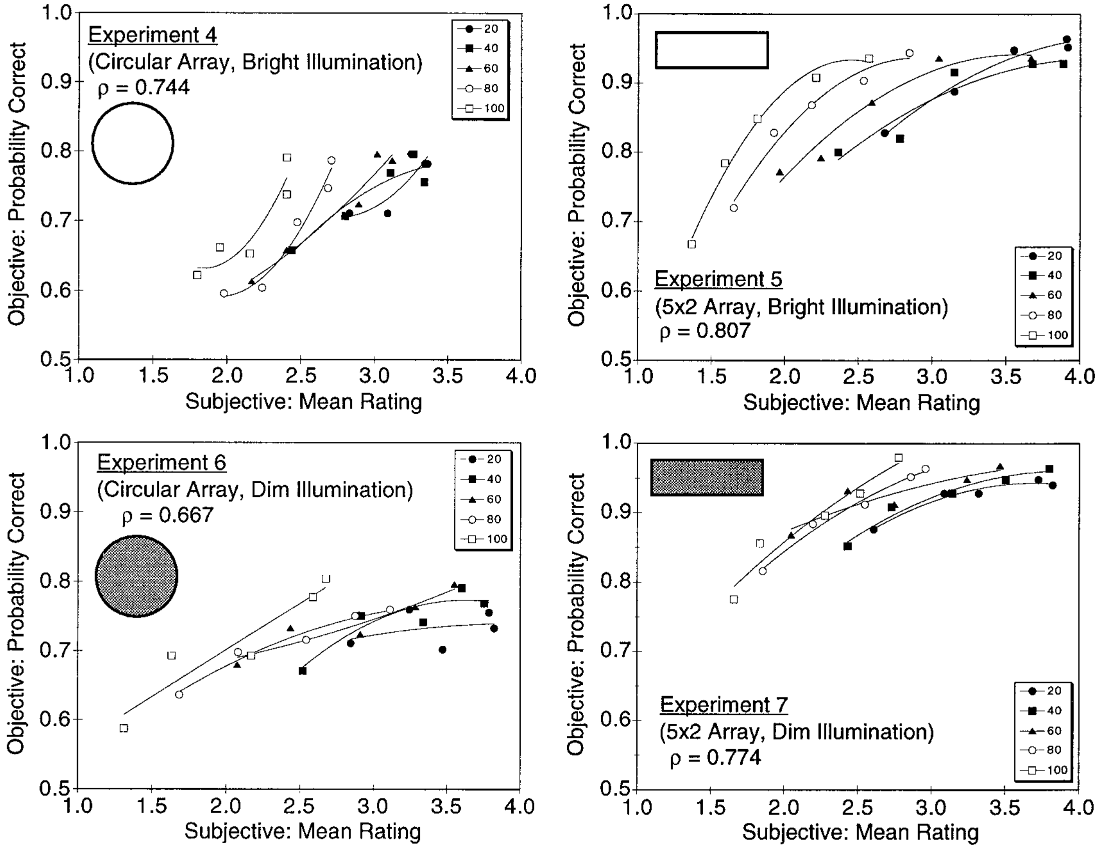

FIG. 18. Experiments 4–7: Correlations between objective measure (ordinate) and subjective measure (abscissa) over 25 experimental conditions. Each panel of Figure 18 is organized like Figure 13. The large symbol in each panel is a mnemonic for stimulus configuration (circular array vs rectangular array) and stimulus brightness (white = bright; gray = dim).

very seriously (especially in the case of the somewhat noisy partial-report data).

Discussion

The data from Experiments 4–7 allow conclusions first about the nature of the objective vs. the subjective task, and second about the nature of the IDE.

Objective-subjective task differences. A fundamental question in this article is the degree to which objective and subjective measures of visual persistence measure the same thing. In Experiments 4–7 we have confirmed what we first observed in Experiment 3, which is that the subjective-rating data and the partial-report data differ systematically in several respects. These differences are encapsulated in Fig. 18 and result in large part because the IDE is more dramatic and consistent in the subjective than in the objective measure. The objective-subjective differences are also exhibited in two other guises.

First, the completeness-rating data are statistically clean (see bottom-left panels of Fig. 14–17) and quite consistent across the four experiments (see Table 5, lower left intercorrelations among Experiments 4–7). In contrast, the objective partial-report data are statistically messy (see top-left panels of Figures 14–17) and much less consistent across the four experiments (see Table 5, upper right, intercorrelations). It is noteworthy that the intercorrelations of Experiments 4–7 with Experiments 1 and 3 are greater for the subjective measure (Table 5, bottom row) than for the objective measure (Table 5, rightmost column).

Second, the completeness ratings are better fit by our theory than are the objective partial-report scores (see Figs. 14–17, right panels).

On the inverse-duration effect. The IDE, while pervasive in visible-persistence tasks such as temporal integration and synchrony-judgment, has been elusive in partial report. It has been reported by Di Lollo and Dixon (1988; 1992a), who intimated that rather special conditions were necessary to obtain it. We believe that additional research is still necessary to completely identify the reasons for, and concomitantly the necessary and sufficient conditions for obtaining an IDE in partial report. We do note that with the relatively short exposure durations that we used, the following is true (see Fig. 12). First, brightness compensation is necessary for obtaining an IDE. Second, stronger IDEs were found (1) with higher-luminance displays and (2) with 5×2 10-letter displays than with circular 15-letter displays.

The finding that brightness compensation is necessary (and, in our experiments, sufficient) to obtain an IDE in partial report permits a fairly simple account of the effect in our procedure: it has long been known that information is acquired faster from high-intensity than from low-intensity stimuli (e.g., Loftus, 1985). Thus, the IDE could arise at least in part as a result of increased luminance for the shorter-duration stimuli. We note though that

Di Lollo and Dixon (1992a) have shown an IDE in the absence of brightness matching for relatively long-duration stimuli; accordingly, such an explanation would not suffice as a general account of the IDE.

GENERAL DISCUSSION

We first summarize our major findings, and then discuss two issues relevant to the general topic of visible persistence: the status of subjective vs. objective persistence measures, and the role of persistence in perception.

Summary

As noted at the outset, we had two major goals. The first was to determine whether three different tasks alleged to measure persistence-temporal integration, subjective completeness, and partial report-measure the same or different internal events. The second was to apply a theory to the objective and subjective data from our seven experiments. We discuss these goals in turn.

Do Different Tasks Measure Different Internal Events?

To determine the degree to which our different tasks measure the same or different underlying mental event, we calculated out a variety of over-condition, interexperimental correlations that are presented in Tables 2 and 5. Several conclusions are readily apparent.

1. As indicated in Table 2, the Experiments 1–2 correlations between completeness-rating and temporal-integration performance are very high (0.977 and 0.979), indicating that, as suggested by Loftus and Hanna (1989), these two tasks measure something very close to the same thing.

2. As also indicated in Table 2, the corresponding Experiments 3–7 correlations between completeness-rating and partial-report performance are rather low (ranging from 0.628 to 0.808). An examination of Figs. 11 and 14–17 illuminate these correlations. Both subjective rating and partial report decline quite robustly with ISI which leads to a positive relation between the two measures; however, rating declines much more robustly with H1D than does partial-report performance, which prevents the correlation from being very high.

3. As indicated in Table 5, completeness-rating performance is highly correlated over conditions for Experiment 1 (which incorporated two successive dot stimuli) on the one hand and Experiments 3–7 (which incorporated a letter array followed by a bar probe) on the other hand. The magnitude of these correlations—which range from 0.902 for the Experiment 1/Experiment 4 comparison to 0.994 for the Experiment 1/Experiment 3 comparison—is remarkable given that the actual stimuli were so different for Experiment 1 vs. Experiments 3–7. It is also notable that the lower correlations of this group—those of Experiment 1 with Experiments 4, 5, 6, and 7—in-

volved one experiment in which there was no brightness matching (Experiment 1) with another experiment in which brightness matching was implemented (Experiment 4, 5, 6, or 7). All of this indicates that the experience of temporal unity is determined much more by the timing characteristics of to-be-integrated stimuli than by the spatial configuration or physical characteristics of the stimuli.

4. In contrast to subjective completeness performance, *objective* (temporal integration or partial-report) performance is *not* highly correlated over conditions for Experiments 1 (temporal integration) on the one hand and Experiments 3–7 (partial report) on the other hand. These correlations range from 0.683 for the Experiment 1/Experiment 3 comparison to 0.930 for the Experiment 1/Experiment 5 comparison. The reason that Experiment-1 objective performance correlates higher with Experiments 4, 5, 6, and 7 than with Experiment 3 is, of course, that the brightness-matching procedures of Experiments 4–7 produced the inverse-duration effect that was so prominent in Experiment 1; there was no inverse-duration effect in the non-brightness matched Experiment 3.

In short, temporal integration and completeness rating measure much the same thing, while partial report is the odd measure out. The traditional view of the icon (e.g., Neisser, 1967), which assumed that all these various persistence tasks measure the same thing, is clearly incorrect. Rather, as Coltheart (1980) and many others have since argued, visible persistence (the lingering visible trace of a stimulus after its offset) differs from informational persistence (knowledge about the properties of a recently extinguished stimulus). One virtue of the present investigation is that differences in task performance were established under conditions in which task alone was varied while all other relevant variables were held constant; as we argued earlier, this eliminates alternative explanations that might be proposed based on other, uncontrolled differences between experiments.

Theory Application

We consider a virtue of our approach to be the development of a general theory designed to account for performance in all tasks. We applied this theory to all performance measures from the seven experiments. Theory application was somewhat different for temporal integration and completeness-rating performance on the one hand, and partial-report performance on the other.

Temporal integration and subjective completeness. To account for temporal integration and completeness ratings the theory incorporates two general assumptions, plus two specific assumptions. The general assumptions (used in all applications of the theory) are as follows. First, a linear filter operates on the input stimulus temporal waveform, $f(t)$, to generate a sensory response function, $a(t)$. Second, the rate of acquiring information, $r(t)$, is a positive function of $a(t)$ but an inverse function of already-acquired information $I(t)$. The specific assumptions, required to apply the theory to these particular tasks, are as follows. First, conscious awareness of any stimulus is determined by the level of $r(t)$. Second, the degree to which two temporally separated events are perceived as contemporaneous is determined by the degree to which the $r(t)$ functions are correlated over time.

The theory fits were quite good: the theory-data rank-order correlations for temporal integration and all subjective tasks range from 0.946 to 1.000. It should be noted that these fits are to 42 free data points (50 conditions minus 5 ''corrected conditions'' minus 3 parameters) in the combined Experiments 1 and 2, and to 21 free data points (25 conditions minus 4 free parameters) in each of Experiments 3–7.

Other linear-filter theories of temporal integration have recently been put forth by Dixon and Di Lollo (1994) and by Wolford (1992). By these, as by the present theory, a linear filter operates on the stimulus to generate a corresponding sensory-response function. The theories differ principally in the assumed relation between the sensory-response function and the various performance measures. In Dixon and Di Lollo's theory, temporal integration performance is assumed to depend on the temporal correlation between the $a(t)$ functions corresponding to the two stimulus halves; in particular, at any time, t, the system computes a ''traveling correlation window'' that correlates the $a(t)$ functions backward in time with progressively older function values receiving progressively less weight in the correlation. In Wolford's theory, functions corresponding to the two halves are simply summed to produce a single function. For some temporal configurations (e.g., H1D, ISI, and H2D all short) this summed function is single-peaked, while for others (e.g., long ISI) the function is twin-peaked. Temporal integration performance is assumed to depend on the degree that the function is single- rather than twin-peaked. Unlike the present theory, neither Dixon and Di Lollo's nor Wolford's theory incorporates the notion of ''information acquisition.'' While providing more parsimony than the present theory in their applications to visible persistence, lack of an assumed information-acquisition process does not allow simple application of these theories to memory tasks in which acquired information is the prime determinant of performance.

Partial report. To apply our theory to partial report, we added two specific assumptions to the two general assumptions sketched above. First, information is acquired randomly from the stimulus array prior to probe presentation, but exclusively from the probed position following probe presentation. Second, the correct response can be made based on either pre-probe or post-probe information independently.

The theory fits to the partial-report data are generally not as good as the theory fits to the temporal integration and completeness rating data: the theory-data rank-order correlations ranged from 0.869 to 0.949. One possible reason for the poorer fits is that partial-report performance is not nearly as statistically clean as is temporal-integration and completeness-rating data; a lower rank-order correlation would be expected on that basis alone. Support for this conjecture is provided in the bottom row of Table 3 where information is presented for the mean of Experiments 4–7 (all brightness-matched experiments). While interpretation of results obtained from averaging over experiments is certainly problematical, it is nonetheless provocative that the theory fit is substantially higher ($\rho = 0.972$) than for any of the individual partial-report data sets.

There are, in any event, several aspects of the partial-report paradigm that mitigate against its being well fit by the present theory. First, as noted earlier, partial report is particularly prone to subject strategies—particularly strategies involving the manner in which the stimulus array is processed prior to probe presentation. One recent, relatively successful partial-report theory (Gegenfurtner & Sperling, 1993) explicitly assumes quite specific attentional strategies. Although such strategies can be experimentally induced (Holding, 1970), one cannot assume that subjects will always use the strategies specified by the theory. Accordingly, no particular theory is likely to be successful in accounting for all partial-report data.

The second problem is that our theory assumes partial-report performance to be entirely determined by something called ''acquired information,'' which, in turn, is assumed to be a unitary substance acquirable either before or after probe presentation. Various investigators, notably Irwin (e.g., Irwin & Yeomans, 1986a; Irwin & Brown, 1987; see also Di Lollo & Dixon, 1988) have provided evidence that at least two kinds of information contribute to partial-report performance. Irwin characterized these two information types as ''visible persistence'' (which begins to decay following stimulus onset) and ''visual analog information'' (which begins to decay following stimulus offset). Such models can account for the inverse duration effect in partial report, while the present theory accounts for the inverse duration effect only to the degree that it occurs as a result of brightness matching.

The final problem is elucidated by a finding originally reported by Townsend (1973), that has been recently confirmed and generalized by Dixon and Di Lollo (1994): an increase in *probe duration* (from 40 to 900 ms in the Townsend study and from 20 to 320 ms in the Di Lollo and Dixon study) causes a relatively robust decline in partial-report performance. This finding, which is utterly unpredicted by the present theory's application to partial report, indicates that, at least when a visual probe is used, partial-report performance, like temporal integration performance, depends in part on whether the entire array-probe configuration is perceived to be a unitary or a dual temporal event.

Theoretical Unification of Performance Measures

We have argued that (apart from minor strategy differences) temporal-integration and subjective-completeness tasks measure the same thing which, in turn, is somewhat different from what is measured by partial report. These conclusions have been based on the pattern of correlations among the various measures. That is, both within and across experiments, subjective-completeness performance measures are highly correlated, with one another and with temporal-integration measures. In contrast, partial-report performance measures are *not* highly correlated with subjective-completeness or temporal-integration measures either within or across experiments.

Our theory allows us, however, to view all of these performance measures in a somewhat more unified fashion, in the sense that all are determined, in one way or another, by the theory's information-extraction rate ($r(t)$) function: In particular, phenomenological appearance is (essentially) assumed to be determined by the *shape* of the $r(t)$ function,[15] while memory performance is (essentially) assumed to be determined by the *area under* the $r(t)$ function.

The $r(t)$ function is affected in various ways by manipulating such variables as H1D, ISI, and H2D, and ensuing performance differences are, by the theory, concomitantly affected. However, (roughly speaking) some variable's effect on $r(t)$'s *shape* is not necessarily the same as its effect on *area* which means (strictly speaking) that the variable's effect on appearance-related measures—such as temporal integration and subjective completeness—is not necessarily monotonically related to its effect on memory-related measures such as partial report.

On the Relation between Objective and Subjective Measures

In earlier sections, we have argued that subjective ratings offer the most direct way of measuring what we believe to be a fundamental perceptual event: the perception of whether two temporally distinct stimuli are seen contemporaneous, or as two discrete temporal events. We argued further that objective tasks, such as temporal integration and partial report, measure this event imperfectly, because various strategies and other sorts of information intervene between perception and performance.

We believe that the subjective rating data, which are simple, direct, and not very prone to elaborate strategies, probably form the best measure of the observer's internal perception of the stimulus ensemble. The objective temporal-integration data, on the other hand, while largely consistent with the rating data, are almost certainly prone to idiosyncratic strategies (e.g., memorization strategies that help overcome the deleterious effects of H1D on visibility) that are not included in the theory. This is the sort of situation

[15] Note, of course, that the mapping of shape to the performance measure depends on task. In a synchrony-judgment task, for instance, a stimulus is assumed to be phenomenologically present to the degree that $r(t)$ is high (and is reported to have disappeared when $r(t)$ drops below some threshold). In a temporal-integration or subjective-completeness task (as we have described in detail) two stimuli are perceived as contemporaneous to the degree that their $r(t)$ functions are positively correlated over time.

depicted in Fig. 1C. Finally, performance in a partial-report task seems to depend not only on a phenomenal, lingering trace of an extinguished stimulus, but on information about letter position and letter identity stored in other memory systems as well, as several investigators have recently suggested (e.g., Irwin & Yeomans, 1986a; Irwin & Brown, 1987; Mewhort, Campbell, Marchetti, & Campbell, 1981). H1D has little effect on partial report performance because subjects can respond based on information acquired during stimulus presentation (informational persistence), rather than on their internal perception of the stimulus array (visible persistence). Thus, completeness rating and temporal integration seem to tap largely into the same underlying representation, but partial report accesses a different source of information as well; this is the sort of multidimensional representational situation depicted in Fig. 1B. Our theory is not designed to take strategies (temporal integration data) and multidimensional stimulus representations (partial report data) into account. It is probably for this reason that the theory's account of the subjective data is better than its account of the objective data.

What Is Persistence Good for?

Back in the days when the traditional view—that visible and informational persistence tasks measured the same thing—was widely accepted, a great deal of discussion centered on the questions: Why did evolution design persistence? What role does it play in perceptual functioning? Although, as we've discussed above, the traditional definition of persistence is incorrect, the same question is still relevant: Why do we experience a lingering trace of a stimulus after its offset? Some investigators have denied that persistence plays any role at all in ''normal'' perception (e.g., Haber, 1983). In contrast, others have suggested that some form of persistence may be useful for integrating information across eye fixations (e.g., Breitmeyer, Kropfl, & Julesz, 1982; Jonides, Irwin, & Yantis, 1982; McConkie & Rayner, 1976). Over the past decade, however, a number of studies has ruled out this latter possibility: fixation durations are typically too long for retinotopic visible persistence to survive a saccadic eye movement (Irwin, Brown, & Sun, 1988), and spatiotopic visible persistence appears not to exist (e.g., Irwin, 1991; Irwin, Yantis, & Jonides, 1983; Irwin, Zacks, & Brown, 1990; Jonides, Irwin, & Yantis, 1983; Rayner & Pollatsek, 1983; see Irwin, 1992 for a comprehensive review). Thus, the question of the usefulness of persistence is still extant. We would like to offer two answers to this question.

Persistence as a Consequence of Imperfect Temporal Resolution

Any system that lacks temporal resolution will have something analogous to persistence—a residue of the input that outlasts the input itself. From this perspective, it would be enormously surprising if any stage of the visual system did *not* have some neural residue that outlasted input from the previous stage. In short, persistence is an entirely natural and expected part of the visual system.

This assertion seems almost self-evident in the context of recently proposed theories—of which linear-filter theories are a prominent class—wherein the visual system, like many physical systems, has an output that in some fashion lags behind and is temporally blurred relative to the input. Given this kind of theory persistence, as noted earlier, becomes relatively uninteresting: it is simply that part of the system's response that occurs later in physical time than stimulus offset (cf., Eriksen & Schultz, 1978).

In contrast, the ''persistence'' about whose purpose earlier investigators puzzled was conceptually different: it was a residual activity (or set of activities) stuck onto the end of the ''standard'' perceptual processing—where ''standard perceptual processing'' was the processing that was assumed to take place when it *ought* to take place, namely during stimulus presence. Such persistence was indeed a unique, but mysterious entity whose existence demanded explanation. From our theory's perspective, there is nothing mysterious about persistence, nor is persistence itself intrinsically interesting; it exists, and we study it, because it is part of the larger story of how the visual system works.

Temporal Integration vs Temporal Separation

It is natural to think that a visual system should be designed to perceive the world as veridically as possible; for example it should be designed such that it perceives as temporally distinct those events that are indeed physically separated in time. Recently, however, Dixon and Di Lollo (1994) and Wolford (1992) have pointed out that the visual system actually has two conflicting goals. The first, to be sure, is to *distinguish* closely spaced temporal events that somehow belong apart. The second, however, is to *integrate* those temporally separate events that somehow belong together (for instance, successive views of a predator moving behind bushes, or successive frames of a motion picture). To accomplish both these goals, the system needs to be designed such that it has what is in some sense an optimal amount of temporal irresolution (where amount of irresolution could, for instance, be expressed by choice of a temporal linear filter's parameters). The investigation of persistence is, in fact, an investigation of what it is that is being optimized, and how such optimization is accomplished by the visual system.

APPENDIX A

Theory Equations for Predicting Temporal Integration

1. Generation of $a(t)$. The characteristics of a low-pass temporal filter are entirely determined by the system's response to an *impulse* input, where an impulse is defined to be an infinitesimally short stimulus of infinite intensity

and unit area. A commonly assumed such *impulse-response function* is a gamma function, $g(t)$, as indicated in text Eq. (2).

The system's response function, $a(t)$ to an arbitrary stimulus waveform, $f(t)$ is obtained by assuming the stimulus to be composed of a series of impulses, scaled by intensity, whose individual responses sum. Accordingly, the response function is the convolution of $f(t)$ and $g(t)$. When the stimulus is a square-wave function, as used in the present experiments, the convolution is provided by text Eq. (1).

2. Generation of $r(t)$. The information-extraction rate, $r(t)$ is the derivative of acquired information, I(t) with respect to time, dI/dt. This, with text Eq. (3) and (4), thus implies:

$$r(t) = dI(t)/dt = a(t)[1.0 - I(t)]/c \tag{A1}$$

or,

$$dI(t)/[1.0 - \mathrm{I}(t)] = ca(t)dt \tag{A2}$$

Integrating both sides of Eq. (A2), with initial conditions of $I(t) = 0$ when $t = 0$ yields,

$$-\mathrm{l}n[1.0 - \mathrm{I}(t)] = c\mathrm{A}(t) \tag{A3}$$

where A(t) is the integral from 0 to t of $a(t)dt$. Algebraic manipulation of Eq. (A3) provides

$$I(t) = 1.0 - e^{-\mathrm{A}(0,\, t)/c)} \tag{A4}$$

Finally, differentiating both sides of Eq. (A4) yields text Eq. (6).

APPENDIX B

Actual ISI, H1D and Luminance Values for Experiments 3–7

Because stimulus luminance was not under software control, luminance had to be manipulated on line via the stimulus refresh rate: to achieve the desired luminance, the stimulus was painted over a 5.5-ms interval once every 10 ms. However, the bar probe was always presented at exactly the desired SOA. To see the effect of this configuration, consider as an example the nominal 20 ms H1D, 20 ms ISI, 20 ms H2D condition. Here, nominal SOA was 40 ms and, indeed, the bar probe appeared exactly 40 ms following stimulus onset. However, the 20-ms H1D was accomplished by refreshing the letter array twice (for 5.5 ms during each 10-ms period), which meant that the array was physically present on the screen only until 10 ms (the first refresh period) + 5.5 ms (the physically present time of the second refresh period) = 15.5 ms following stimulus onset; thus 15.5 ms was the functional H1D. Likewise, the remaining 24.5 ms that elapsed prior to the 40 ms SOA was functional ISI. Half-2 presentation was identical to Half-1 presentation; accordingly, functional H2D was always 15.5 ms. In short, the nominal 20-20-20 condition was functionally a 15.5-24.5-15.5 condition. The functional H1Ds, ISIs, and effective luminances for the 25 conditions of Experiments 3–7[16] are available from the second author.

REFERENCES

Allen, F. (1926). The persistence of vision. *American Journal of Physiological Optics,* **7,** 439–457.

Averbach, E., & Coriell, H. S. (1961). Short-term memory in vision. *Bell Systems Technical Journal,* **40,** 309–328.

Averbach, E., & Sperling, G. (1961). Short-term storage in Vision. In C. Cherry (Ed.), *Symposium on information theory* (pp. 196–211). London: Butterworth.

Bamber, D. (1979). State trace analysis: A method of testing simple theories of causation. *Journal of Mathematical Psychology,* **19,** 137–181.

Bowen, R. W., Pola, J., & Matin, L. (1974). Visual persistence: Effects of flash luminance, duration and energy. *Vision Research,* **14,** 295–303.

Bowling, A., & Lovegrove, W. (1982). Iconic memory: Fallacies persist (?). *Perception & Psychophysics,* **31,** 194–198.

Breitmeyer, B. G., Kropfl, W., & Julesz, B. (1982). The existence and role of retinotopic and spatiotopic forms of visual persistence. *Acta Psychologica,* **52,** 175–196.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review,* **97,** 523–547.

Busey, T. A., & Loftus, G. R. (1994). Sensory and cognitive components of visual information acquisition. *Psychological Review,* **101,** 446–469.

Busey, T. A., & Loftus, G. R. (1998). Binocular Information acquisition and visual memory. *Journal of Experimental Psychology: Human Perception and Performance,* in press.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception and Psychophysics,* **27,** 183–228.

Coltheart, M., Lea, C. D., & Thompson, K. (1974). In defence of iconic memory. *Quarterly Journal of Experimental Psychology,* **26,** 633–641.

Dick, A. O. (1974). Iconic memory and its relation to perceptual processes and other mechanisms. *Perception & Psychophysics,* **16,** 575–596.

Di Lollo, V. (1980). Temporal integration in visual memory. *Journal of Experimental Psychology: General,* **109,** 75–97.

Di Lollo, V. (1983). On laterality of visual aftereffects: A rejoinder. *Perception & Psychophysics,* **33,** 599–603.

Di Lollo, V. (1984). On the relationship between stimulus intensity and duration of visible persistence. *Journal of Experimental Psychology: Human Perception and Performance,* **10,** 144–151.

[16] Matters were further complicated in the nominal −20 ISI condition, as the refresh rate changed to accommodate the bar being presented simultaneously with the stimulus.

Di Lollo, V., & Dixon, P. (1988). Two forms of persistence in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance,* **14,** 671–681.

Di Lollo, V., & Dixon, P. (1992-a). Inverse duration effects in partial report. *Journal of Experimental Psychology: Human Perception and Performance,* **18,** 1089–1100.

Di Lollo, V., & Dixon, P. (1992-b). Is the icons worth apples and oranges? Comments on Loftus, Duncan, and Gehrig 1992. *Journal of Experimental Psychology: Human Perception and Performance,* **19,** 350–355.

Dixon, P., & Di Lollo, V. (1992). Effects of display luminance, stimulus type, and probe duration on visible and schematic persistence. *Canadian Journal of Psychology,* **45,** 54–74.

Dixon, P., & Di Lollo, V. (1994). Beyond visible persistence: An alternative account of temporal integration and segregation in visual processing. *Cognitive Psychology,* in press.

Efron, R. (1970-a). The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia,* **8,** 37–55.

Efron, R. (1970-b). Effect of stimulus duration on perceptual onset and offset latencies. *Perception and Psychophysics,* **8,** 231–234.

Eriksen, C. W., & Collins, J. F. (1967). Some temporal characteristics of visual pattern perception. *Journal of Experimental Psychology,* **74,** 476–484.

Eriksen, C. E., & Schultz, D. W. (1978). Temporal factors in visual information processing: A tutorial review. In J. Requin (Ed), *Attention and performance VII* (pp. 3–23). Hillsdale, NJ: Erlbaum.

Gardiner, J. M., & Parkin, A. J. (1990). Attention and recollective experience in recognition memory. *Memory & Cognition,* **18,** 579–583.

Gegenfurtner, K., & Sperling, G. (1993). Information transfer in iconic memory experiments. *Journal of Experimental Psychology: Human Perception and Performance,* **19,** 845–866.

Groner, M. T., Bischof, W. F., & Di Lollo, V. (1988). A model of visible persistence and temporal integration. *Spatial Vision,* **3,** 293–304.

Groner, R., Groner, M. T., Muller, P., Bischof, W. F., & Di Lollo, V. (1993). On the confounding effects of phosphor persistence in oscilloscopic displays. *Vision Research,* **33,** 913–917.

Haber, R. N. (1969). Information processing approaches to visual perception: An introduction. In R. N. Haber (Ed.), *Information processing approaches to visual perception* (pp. 1–15). New York: Holt, Rinehart, & Winston.

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *The Behavioral and Brain Sciences,* **6,** 1–10.

Hawkins, H. L., & Shulman, G. L. (1979). Two definitions of persistence in visual perception. *Perception and Psychophysics,* **25,** 348–350.

Hogben, J. H., & Di Lollo, V. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vision Research,* **14,** 1059–1069.

Holding, D. H. (1970). Guessing behavior and the Sperling store. Quarterly Journal of Experimental Psychology, **22,** 248–256.

Intraub, H. (1980). Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimental Psychology: Human Learning and Memory,* **6,** 1–12.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology,* **23,** 420–456.

Irwin, D. E. (1992). Perceiving an integrated visual world. In D. E. Meyer and S. Kornblum (Eds.), *Attention and performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 121–142). Cambridge, MA: MIT Press.

Irwin, D. E., & Brown, J. S. (1987). Tests of a model of informational persistence. *Canadian Journal of Psychology,* **41,** 317–338.

Irwin, D. E., Brown, J. S., & Sun, J.-S. (1988). Visual masking and visual integration across saccadic eye movements. *Journal of Experimental Psychology: General,* **117,** 274–285.

Irwin, D. E., Yantis, S., & Jonides, J. (1983). Evidence against visual integration across saccadic eye movements. *Perception & Psychophysics,* **34,** 49–57.

Irwin, D. E., & Yeomans, J. M. (1986-a). Sensory registration and informational persistence. *Journal of Experimental Psychology: Human Perception and Performance,* **12,** 343–360.

Irwin, D. E., & Yeomans, J. M. (1986-b). Persisting arguments about visual persistence: Reply to Long. *Perception & Psychophysics,* **39,** 225–230.

Irwin, D. E., Zacks, J. L., & Brown, J. S. (1990). Visual memory and the perception of a stable visual environment. *Perception & Psychophysics,* **47,** 35–46.

Jonides, J., Irwin, D. E., & Yantis, S. (1983). Failure to integrate information from successive saccades. *Science,* **222,** 188.

Julesz, B., & Chiarucci, E. (1973). Short-term memory for stroboscopic movement perception. *Perception,* **2,** 249–260.

Kowler, E., & Sperling, G. (1980). Transient stimulation does not aid visual search: Implications for the role of saccades. *Perception and Psychophysics,* **27,** 1–10.

Loftus, G. R. (1985). Picture perception: Effects of luminance level on available information and information-extraction rate. *Journal of Experimental Psychology: General,* **114,** 342–356.

Loftus, G. R., & Busey, T. A. (1992). Multidimensional models and Iconic Decay. *Journal of Experimental Psychology: Human Perception and Performance,* **18,** 356–361.

Loftus, G. R., Busey, T. A., & Senders, J. W. (1993). Providing a sensory basis for models of visual information acquisition. *Perception & Psychophysics,* **54,** 535–554.

Loftus, G. R., Duncan, J., & Gehrig, P. (1992). On the time course of perceptual information that results from a brief visual presentation. *Journal of Experimental Psychology: Human Perception and Performance,* **18,** 530–549.

Loftus, G. R., & Hanna, A. M. (1989). The phenomenology of spatial integration: Data and models. *Cognitive Psychology,* **21,** 363–397.

Loftus, G. R., & Hogden, J. (1988). Picture perception: Information extraction and phenomenological persistence. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 22.* New York: Academic Press.

Loftus, G. R., & Kallman, H. J. (1979). Encoding and use of detail information in picture recognition. *Journal of Experimental Psychology: Human Learning and Memory,* **5,** 197–211.

Loftus, G. R., Johnson, C. A., & Shimamura, A. P. (1985). How much is an icon worth? *Journal of Experimental Psychology: Human Perception and Performance,* **11,** 1–13.

Loftus, G. R., & Ruthruff, E. (1994). A linear-filter theory of visual information acquisition with special application to intensity-duration tradeoffs. *Journal of Experimental Psychology: Human Perception and Performance,* **20,** 33–55.

Long, G. M. (1979). Comment on Hawkins and Shulman's type I and type II visual persistence. *Perception & Psychophysics,* **26,** 412–414.

Long, G. M. (1980). Iconic memory: A review and critique of the study of short-term visual storage. *Psychological Bulletin,* **88,** 785–820.

Long, G. M. (1985). The varieties of visual persistence: Comments on Yeomans and Irwin. *Perception & Psychophysics,* **38,** 381–385.

Long, G. M., & Beaton, R. J. (1982). The case for peripheral persistence: Effects of target and background luminance on a partial-report task. *Journal of Experimental Psychology: Human Perception and Performance,* **8,** 383–391.

Massaro, D. W. (1970). Perceptual processes and forgetting in memory tasks. *Psychological Review,* **77,** 557–567.

Massaro, D., & Loftus, G. R. (1994). Sensory Storage: Icons and Echos. In E. L. Bjork and R. A. Bjork (Eds.), *Handbook of perception and cognition, Vol. 10.* New York: Academic Press (in press).

McConkie, G. W., & Rayner, K. (1976). Identifying the span of the effective stimulus in reading: Literature review and theories of reading. In H. Singer and R. B. Ruddell (Eds.), *Theoretical models and processes of reading* (pp. 137–162). Newark, DE: International Reading Association.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Nisly, S. J., & Wasserman, G. S. (1989). Intensity dependence of perceived duration: Data, theories, and neural integration. *Psychological Bulletin,* **106,** 483–496.

Rayner, K., & Pollatsek, A. (1983). Is visual information integrated across saccades? *Perception & Psychophysics,* **34,** 39–48.

Rumelhart, D. E. (1969). A multicomponent theory of the perception of briefly exposed visual displays. *Journal of Mathematical Psychology,* **7,** 191–218.

Sakitt, B., & Long, G. M. (1979). Cones determine subjective offset of a stimulus but rods determine total persistence. *Vision Research,* **19,** 1439–1441.

Shibuya, H., & Bundsen, C. (1988). Visual selection from multielement displays: Measuring and modeling effects of exposure duration. *Journal of Experimental Psychology: Human Perception and Performance,* **14,** 591–600.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs,* **74,** 1–29.

Sperling, G. (1963). A model for visual memory tasks. *Human Factors,* **5,** 19–31.

Sperling, G. (1964). Linear theory and the psychophysics of flicker. *Documenta Ophthalmologica,* **18,** 3–15.

Sperling, G. (1967). Successive approximations to a model for short-term memory. *Acta Psychologica,* **27,** 285–292.

Townsend, V. M. (1973). Loss of spatial and identity information following a tachistoscopic exposure. *Journal of Experimental Psychology,* **98,** 113–118.

Turvey, M. (1977). Contrasting orientations to a theory of visual information processing. *Psychological Review,* **84,** 67–88.

von Wright, J. M. (1972). On the problem of selection in iconic memory. *Scandinavian Journal of Psychology,* **13,** 159–171.

Watson, A. B. (1986). Temporal sensitivity. In K. R. Boff, L. Kaufman, and J. P. Thomas (Eds.), *Handbook of perception and human performance (Vol I).* New York: Wiley.

Weichselgartner, E., & Sperling, G. (1985). Continuous measurement of visible persistence. *Journal of Experimental Psychology: Human Perception and Performance,* **11,** 711–725.

Wolford, G. (1992). A model of visible persistence based on linear systems. *Canadian Psychology,* **34,** 162–165.

Yeomans, J. M., & Irwin, D. E. (1985). Stimulus duration and partial report performance. *Perception & Psychophysics,* **37,** 163–169.

Accepted January 4, 1998