

Temporal bisection with trial referents

LORRAINE G. ALLAN and KRIS GERHARDT
McMaster University, Hamilton, Ontario, Canada

In temporal bisection, the subject is required to compare the similarity of a probe duration with two referent durations. Two bisection experiments with human subjects are reported in which performance under three referent conditions was compared. In two of the conditions, no-referent and fixed-referent, the referent pair was constant throughout the session and was varied between sessions. In the no-referent condition, samples of the referents were presented at the beginning of a block of trials, but not on the probe trials. In contrast, in the fixed-referent condition, the referent pair was presented on each probe trial. In the third referent condition, roving-referent, the referent pair was also presented on each probe trial, but the values of the pair varied during a session rather than between sessions. The pseudo-logistic Model (PLM; Killeen, Fetterman, & Bizo, 1997) provided an excellent account of the data from all referent conditions and indicated that (1) the scalar variability was located in the perception of the probe and was independent of the structure of the bisection task, (2) the role of the referents was to set the criterion, (3) the probe was compared with the criterion rather than with the referents, and (4) the comparison of the probe with the criterion occurred even when trial referents were available.

There is general agreement among the quantitative models for time (see Allan, 1998) that mean perceived time (μ_t) is a power function of clock time with an exponent close to 1.0,

$$\mu_t = t, \quad (1)$$

and that variability in perceived time (σ_t) is proportional to mean perceived time,

$$\frac{\sigma_t}{\mu_t} = \frac{\sigma_t}{t} = \gamma. \quad (2)$$

The proportionality constant γ is known as the *Weber fraction*, and the proportionality relation is often referred to as *scalar variability*. This scalar property between the standard deviation and the mean results in distributions of perceived time that superpose when the temporal axis is normalized with respect to the mean of the distribution. This superposition in relative time reflects a re-scaling in time, a scale-invariant error distribution for perceived time.

Many psychophysical procedures have been developed for the study of time perception (see Allan, 1979, 1998). One relatively new procedure is temporal bisection. In the prototypic temporal bisection task, two referents, one short (S) and the other long (L), are explicitly identified either by familiarizing the subject with the ref-

erents at the beginning of a block of trials (e.g., Allan, 1999; Wearden, 1991; Wearden & Ferrara, 1995, 1996; Wearden, Rogers, & Thomas, 1997) or periodically throughout a block of trials (e.g., Allan & Gibbon, 1991; Penney, Allan, Meck, & Gibbon, 1998). On probe trials, a temporal interval t , $S \leq t \leq L$, is presented, and the subject is required to indicate whether t is more similar to S (R_S) or to L (R_L). Temporal bisection yields a psychometric function relating the proportion of long responses, $P(R_L)$, to probe duration t . The value of t at which R_S and R_L occur with equal frequency, $P(R_L) = 0.5$, is often referred to as the *bisection point* ($T_{1/2}$). One interpretation of $T_{1/2}$ is that it is the value of t that is equally confusable with S and L . For bisection, the scalar property predicts superposition of psychometric functions for all L -to- S ratios when plotted against t normalized by the bisection point (i.e., $t/T_{1/2}$).

A number of psychophysical models for temporal bisection have been proposed. These models differ with regard to the source of the scalar variability and the nature of the decision rule for categorizing t as R_L or R_S . In his now classic paper on temporal bisection, Gibbon (1981) derived the bisection function for a similarity ratio decision rule. For this rule, the decision to respond R_S or R_L is made by comparing the similarity of the perceived value of t with memories of the two referents, S and L . This comparison is based on a ratio of the similarity of t to S relative to the similarity of t to L . If that ratio is less than a criterion β , the response is R_L . Gibbon combined the similarity ratio rule with two sources for scalar variability, perception or memory. In the "referent known exactly model" (RKE), he assumed no variability in the memory for the referents and placed scalar variability in the perception of the probes. In contrast, in the "stimulus

The preparation of this paper was supported by a grant to L.G.A. from the Natural Sciences and Engineering Research Council of Canada. We are indebted to Simon Grondin, Peter Killeen, and an anonymous reviewer for their insightful comments on an earlier version of this manuscript. We also thank Peter Killeen for the derivation presented in the Appendix. Correspondence should be addressed to L. G. Allan, Department of Psychology, McMaster University, Hamilton, ON L8S 4K1, Canada (e-mail: allan@mcmaster.ca).

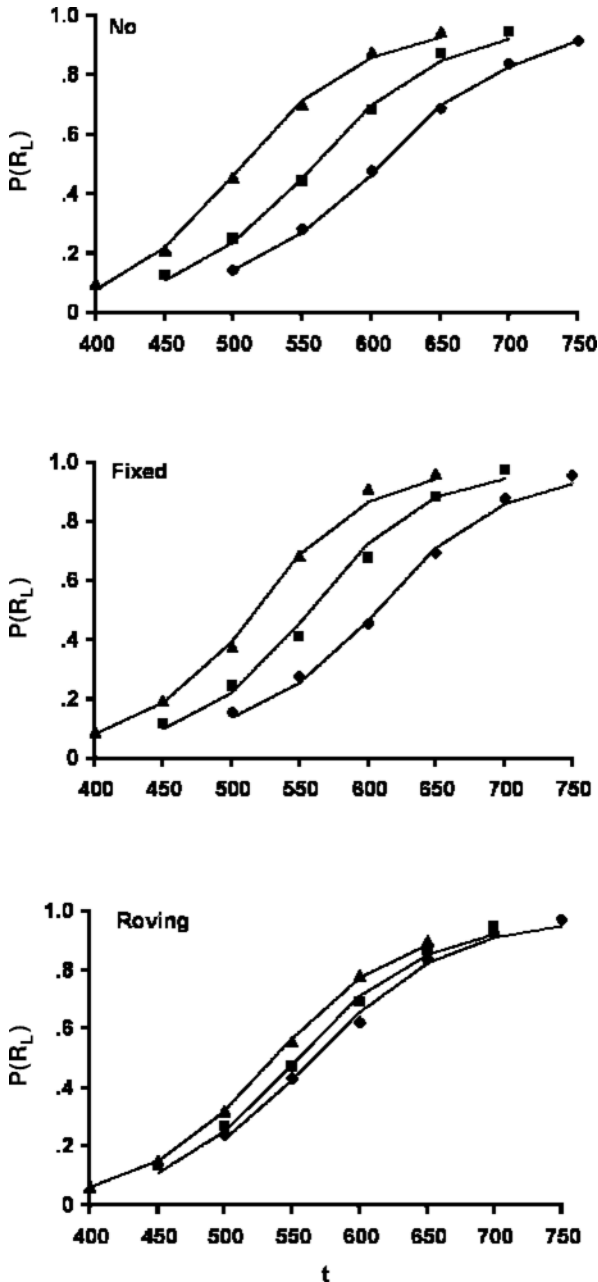


Figure 1. $P(R_L)$, averaged over subjects, as a function of t for each referent pair and each referent condition. The geometric symbols indicate the data (triangles = 650/400, squares = 700/450, and diamonds = 750/500). The lines are derived from the Pseudologic model. (Experiment 1)

known exactly model" (SKE), he assumed no variability in the perception of the probes and placed the scalar variability in the memory for the referents. The similarity ratio rule places $T_{1/2}$ at the geometric mean (GM) of S and L ,

$$T_{1/2} = \sqrt{SL} = GM. \tag{3}$$

The inclusion of the response bias β , however, provides for deviations of $T_{1/2}$ from GM ,

$$T_{1/2} = \sqrt{\frac{SL}{\beta}} = \frac{GM}{\sqrt{\beta}}. \tag{4}$$

Wearden (1991) substituted a similarity *difference* rule for Gibbon's (1981) similarity *ratio* rule. According to the difference rule, the subject responds R_L when t is more similar to L than to S and responds R_S when t is more similar to S than to L . The similarity difference rule places $T_{1/2}$ at the arithmetic mean (AM) of S and L , and the inclusion of a response bias allows $T_{1/2}$ to deviate from AM . Thus, both similarity rules (ratio and difference) involve comparisons of t with both S and L , and both rules provide for deviations of $T_{1/2}$ from their predicted locations by including a response bias parameter.

Allan and Gibbon (1991) argued that temporal bisection maximizes reliance on memory of the referents since they are not presented on probe trials, and based the analysis of their data on SKE. They concluded that although the placement of all scalar variability in the memory of the referents was undoubtedly unrealistic, SKE served remarkably well in characterizing the data. However, more recent data raise questions about the appropriateness of SKE. These new data were generated in variants of the prototypic bisection task designed explicitly to investigate the role of the referents.

Wearden and Ferrara (1995, 1996) modified the prototypic bisection task so that no members of the set were explicitly identified as referents. Rather, subjects were simply instructed to partition the values of t into two categories, R_S or R_L . To distinguish the two tasks, they labeled the task in which the referents are identified and subjects are instructed to compare the *similarity* of t with the referents as "similarity," and the task in which the referents are not identified and subjects are instructed to *partition* the t values into two categories as "partition."¹ Wearden and Ferrara (1995) compared the similarity and partition tasks and concluded that they yielded similar data. Since referents are not explicitly identified in partition, a decision rule that does not involve a direct comparison with the referents might be more appropriate than the similarity rule.

Rodriguez-Girones and Kacelnik (in press) modified the prototypic bisection task so that the referents were presented on each trial. In their bisection task, three duration values were presented sequentially on each trial: the two referents, counterbalanced for order, and then the probe t . Subjects were instructed to indicate whether the last duration (i.e., t) was more similar to the duration of the first or second trial referent. The referent pair was either constant over a series of trials or varied from trial to trial. Rodriguez-Girones and Kacelnik found steeper psychometric functions with constant referents than with variable referents, and they argued that this result indicated that the dominant source of scalar variability was in perception, not in reference memory.

Table 1
Values of γ and $T_{1/2}$ for Each Referent Condition in Experiment 1

Subject	Referent Pair	γ	$T_{1/2}$						
			Collapsed Across Referent Order			<i>SL</i>	<i>LS</i>	<i>SL</i>	<i>LS</i>
			No	Fixed	Roving	Fixed	Fixed	Roving	Roving
B.M.	650/400	.09	519	530	548	508	552	532	563
	700/450		567	580	572	559	600	554	591
	750/500		626	628	606	607	651	582	628
V.L.	650/400	.13	481	498	529	488	510	518	539
	700/450		548	544	549	525	563	533	566
	750/500		592	584	557	573	598	530	581
S.H.	650/400	.14	547	546	587	549	550	592	582
	700/450		592	568	590	599	592	593	587
	750/500		646	626	583	627	626	584	583
N.C.	650/400	.13	508	533	547	512	552	530	563
	700/450		577	581	560	548	609	543	581
	750/500		610	623	577	599	646	560	595
K.E.	650/400	.16	420	405	444	408	403	447	442
	700/450		447	451	463	450	452	461	465
	750/500		447	462	474	439	480	473	475
K.G.	650/400	.21	555	544	530	543	544	478	592
	700/450		584	571	553	521	613	486	620
	750/500		642	650	560	593	702	487	621

The partition data of Wearden and Ferrara (1995, 1996) and the trial referent data of Rodriguez-Girones and Kacelnik (in press) suggest that a model that places scalar variability in perception and that incorporates a decision rule that does not involve a direct comparison with the referents might be appropriate for bisection. One possibility is a bisection model based on signal detection theory (SDT). Gibbon (1981), in fact, did derive an SDT bisection function, in addition to the SKE and RKE functions. For this derivation, he placed scalar variability in the perception of the probe, and he assumed that the perceptual distributions were normally distributed.² More recent developments by Killeen, Fetterman, and Bizo (1997) and by Rodriguez-Girones and Kacelnik have simplified the derivation of, and the resulting equation for, the SDT bisection function.

Rodriguez-Girones and Kacelnik (in press) substituted the log normal distribution for the normal distribution. For the log normal model, the transformation of clock time to psychological time is logarithmic, and the standard deviation is constant. However, there is ample evidence in the literature, both human (see Allan, 1983) and nonhuman (see Gibbon & Church, 1981), that the relationship between clock time and psychological time is linear (Equation 1), not logarithmic. Also, in an extensive analysis of temporal magnitude estimation data, Allan (1983) found that the dispersion of log-transformed magnitude estimates were not symmetrical and that the variability of the log responses was not constant but rather tended to decrease with increasing duration. She concluded that her results did not support the log normal model for magnitude estimates of temporal intervals.

Killeen et al. (1997) maintained the linear relationship between clock time and subjective time, but approximated the normal distribution with the logistic distribution to simplify the mathematical derivation of the bisection

function. For the logistic approximation to the normal, the bisection function is

$$P(R_L) = \left[1 + \exp \left(\frac{T_{1/2} - t}{\frac{\sqrt{3}}{\pi} \sigma_t} \right) \right]^{-1}, \quad (5)$$

where $T_{1/2}$, the bisection point, is the SDT criterion and

$$\sigma_t = \sqrt{(\gamma t)^2 + pt + c^2}. \quad (6)$$

Equation 6 is the Weber function derived by Killeen and Weiss (1987) that provides for scalar (γ), nonscalar (p), and constant (c) sources of variability. When scalar variability dominates, Equation 6 reduces to Equation 2. Killeen et al. noted that Equation 5 is not a logistic function even though it is derived from one, and that Equation 5 is not a distribution function since it asymptotes at a value less than 1.0. Therefore, they referred to their bisection function as a pseudo-logistic function and their SDT model as the pseudo-logistic model (PLM).

Killeen et al. (1997) reanalyzed the bisection data reported by Allan and Gibbon (1991) and found that they were consistent with PLM. Later, Allan (1999) demonstrated the success of PLM with a more extensive data set than the one reported in Allan and Gibbon. In the Allan and Gibbon (1991) and Allan (1999) experiments, the data were generated in the prototypic bisection task; that is, the referents were not presented on probe trials. The major purpose of the present experiments is to evaluate the applicability of PLM to data generated in bisection tasks involving trial referents. We compared psychometric functions generated when the referents were presented on probe trials with psychometric functions generated when the referents were not presented on probe trials. This direct comparison could not be made with the

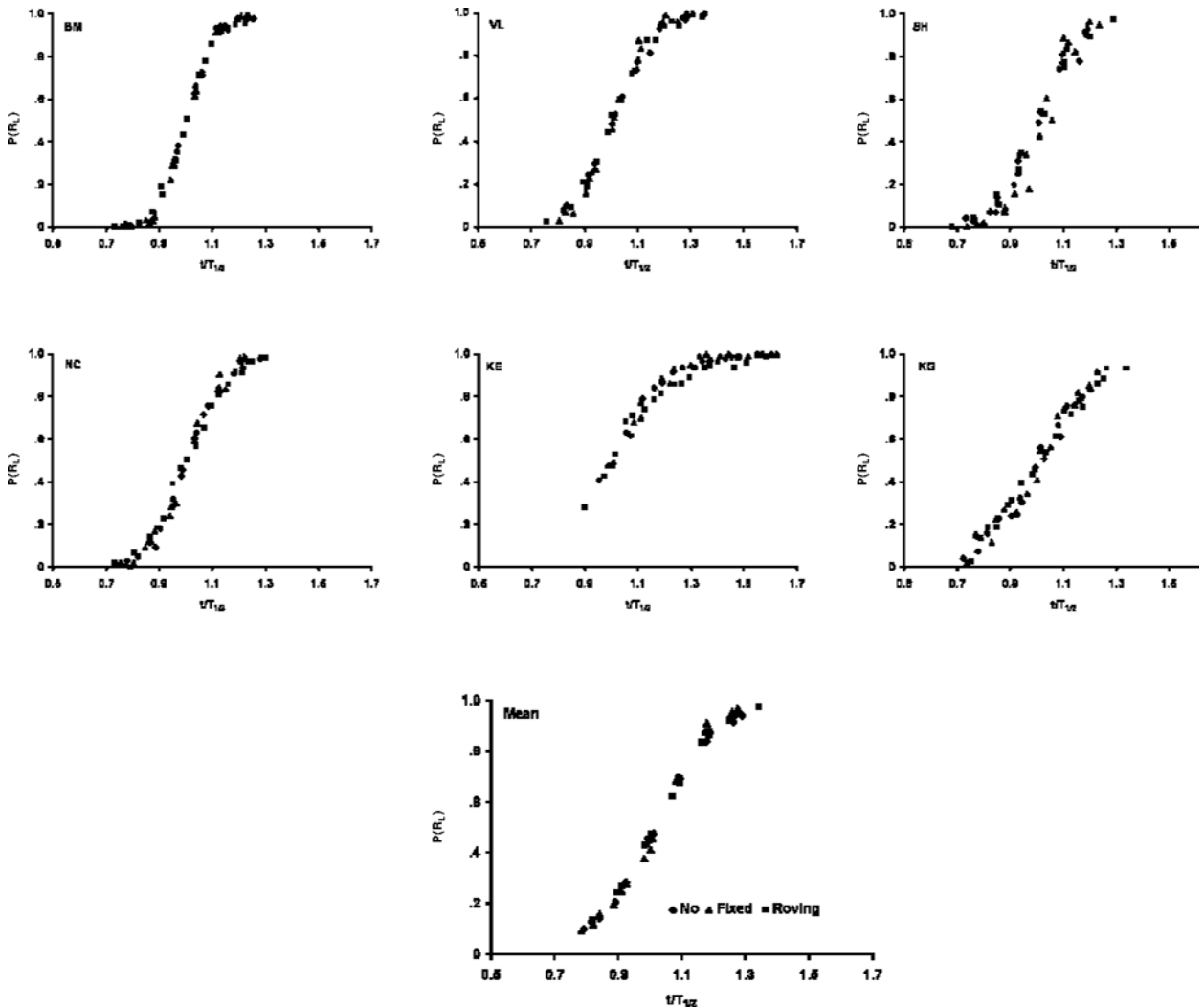


Figure 2. $P(R_L)$ as a function of $t/T_{1/2}$. The psychometric functions for the three referent pairs from the no-referent condition are indicated by diamonds, from the fixed-referent condition by triangles, and from the roving-referent condition by squares. The functions are shown for each subject and also averaged over subjects. (Experiment 1)

data reported by Rodriguez-Girones and Kacelnik (in press) since they did not include the prototypic bisection task (i.e., absence of trial referents) in their experiment.

PLM for the prototypic bisection tasks specifies that the role of the referents is to set the value of the criterion $T_{1/2}$. On each trial, the perceived value of the probe is compared with $T_{1/2}$ (rather than directly with the referents, as in SKE and RKE), and the decision is R_L if the perceived value is larger than $T_{1/2}$. Of particular interest in the present experiments is determining the nature of the comparison process when the referents are presented on each trial. Our earlier research (Allan, 1977; Allan, Kristofferson, & Rice, 1974) suggests that the referents will continue to set the value of $T_{1/2}$ even when they are available on each trial for direct comparison. In these earlier

studies, we concluded that when two temporal intervals were presented on a trial, they were not directly compared. Rather, each interval was compared with a single criterion value and separately categorized as S or L .

EXPERIMENT 1

We compared three referent conditions in Experiment 1. In two of the conditions, no-referent and fixed-referent, the referent pair was constant throughout the session and was varied between sessions. In the no-referent condition, samples of the referents were presented at the beginning of a block of trials, but not on the probe trials. In contrast, in the fixed-referent condition, the referent pair was presented on each probe trial. In the third refer-

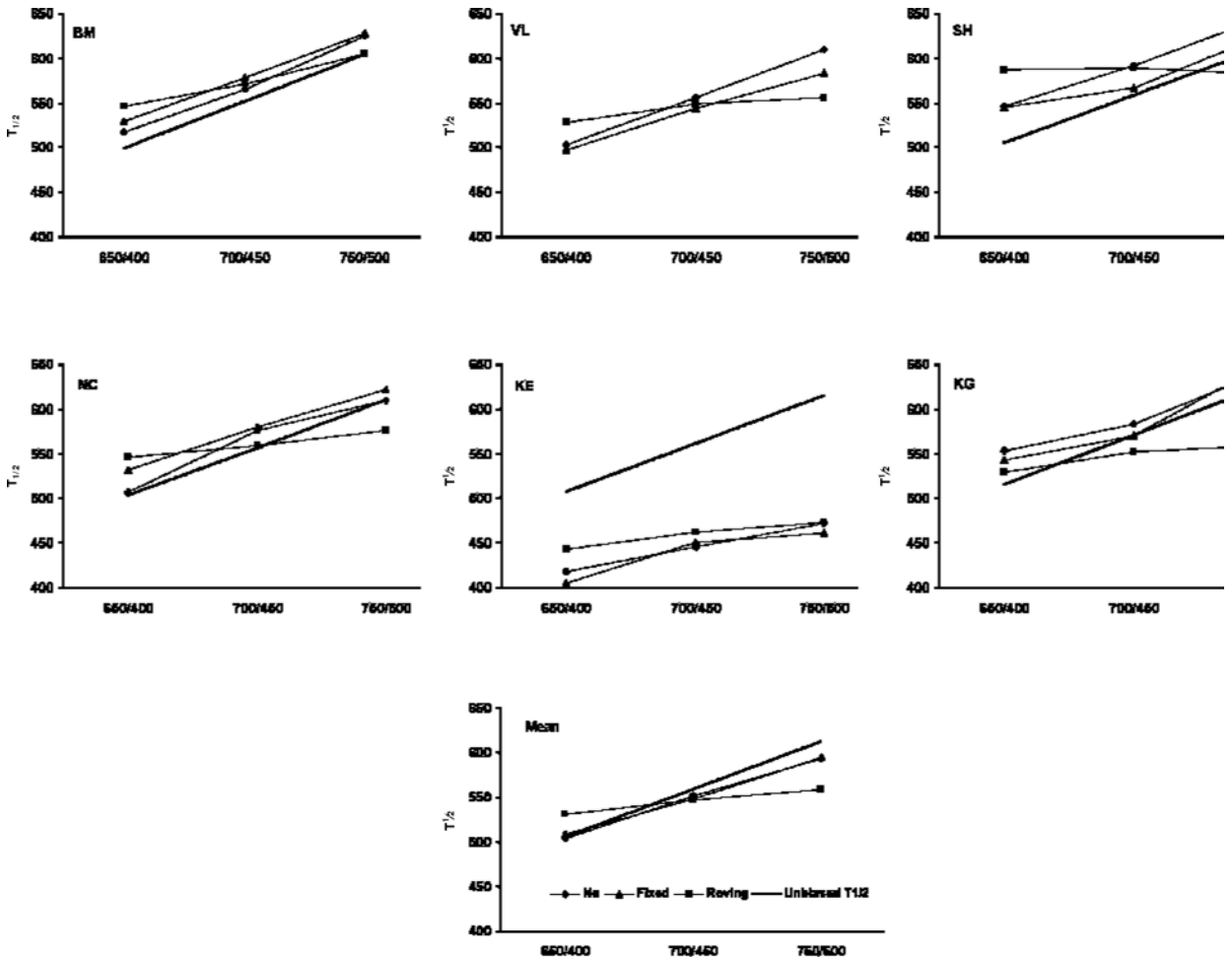


Figure 3. $T_{1/2}$ as a function of referent pair for each referent condition (diamonds = no, triangles = fixed, squares = roving). Also shown is the unbiased $T_{1/2}$ (continuous line) for each referent pair. The data are presented for each subject and also averaged over subjects. (Experiment 1)

ent condition, roving-referent, the referent pair was also presented on each probe trial, but the values of the pair varied during a session rather than between sessions.

Our fixed and roving conditions were similar to those in Rodriguez-Girones and Kacelnik (in press) in that the trial referents were present on probe trials, and were constant in the fixed condition and varied in the roving condition. They differed, however, in many important details. Rodriguez-Girones and Kacelnik did not control for the spacing of the probe durations. In their fixed-referent condition, there were four referent pairs, and the distribution of the probes differed among the four pairs. In addition, the fixed probe distributions differed from the roving probe distribution. Also, Rodriguez-Girones and Kacelnik did not control for range effects. The range of the L referent differed between fixed and roving: For fixed, the range for L was 3.0–4.5 sec, whereas for roving it was .75–5.5 sec. There is evidence in the literature that probe spacing and referent range influence bisection

(see Allan, 1998, 1999). In the present experiments, we controlled for probe distribution and referent range.

The most notable difference is the structure of the roving condition. In the Rodriguez-Girones and Kacelnik (in press) roving condition, S was uniformly distributed between .5 and 2 sec. For any value of S , $L = rS$ for $1.5 \leq$

Table 2
Values of the Slope of the Linear Function Relating $T_{1/2}$ to Referent Pair for Each Referent Condition in Experiment 1

Subject	Condition		
	No	Fixed	Roving
B.M.	53.70	49.10	29.05
V.L.	53.24	43.25	13.80
S.H.	49.50	40.40	-2.20
N.C.	51.15	44.95	14.95
K.E.	26.90	28.20	14.90
K.G.	43.65	53.15	14.90
<i>M</i>	46.36	43.18	14.23

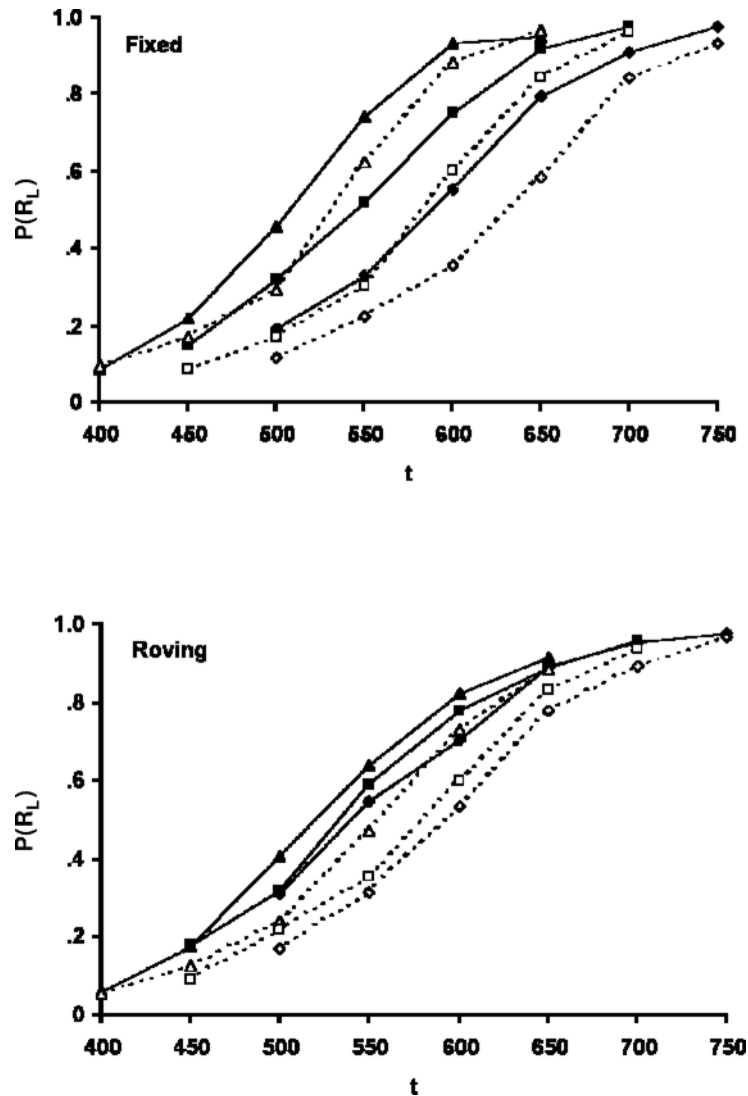


Figure 4. $P(R_L)$, averaged over subjects, as a function of t for each referent pair (triangles = 650/400, squares = 700/450, and diamonds = 750/500) and each referent presentation order (SL = filled symbols and continuous lines, LS = unfilled symbols and dotted lines). S , short; L , long. (Experiment 1)

$r \leq 8$, with an upper limit of $L = 5.5$ sec. For example, for $S = 1$ sec, L was equally likely to take on any value between 1.5 and 5.5 sec. Although this structure introduced extensive variability in the value of the trial referents, it also introduced complications in the analysis of the data. As Rodriguez-Girones and Kacelnik noted, only normalized psychometric functions could be constructed, which required an a priori assumption about the value of $T_{1/2}$. Our roving condition was modeled after the roving-standard designs used in our earlier duration discrimination studies (e.g., Allan, 1977).

Our experimental design also differed from that of Rodriguez-Girones and Kacelnik (in press). They used a between-subjects design with each subject being tested

under only one condition (roving or fixed), and in the fixed condition with only one of the four referent pairs. One purpose of our research was to fit PLM to individual subject data, thereby allowing comparison of parameter values across referent conditions. Therefore, we used a within-subjects design, with each subject running under all experimental conditions. In order to avoid explicit counting, our duration values were lower than 750 msec, and therefore considerably shorter than those in Rodriguez-Girones and Kacelnik.

Method

Subjects. The seven subjects (4 females and 3 males) were graduate students in the psychology department at McMaster University

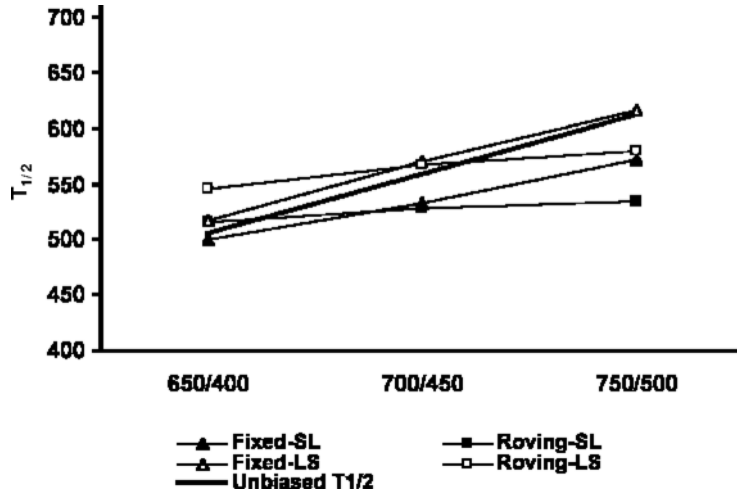


Figure 5. $T_{1/2}$, averaged over subjects, as a function of referent pair separately for each referent presentation order (SL = filled symbols, LS = unfilled symbols) from the fixed (triangles) and roving (squares) conditions. Also shown is the unbiased $T_{1/2}$ (continuous line) for each referent pair. S , short; L , long. (Experiment 1)

who were paid for their participation. The psychometric functions for 1 subject were quite flat and nonmonotonic, and these data were not included in the data analysis. Subjects were required to complete a minimum of five sessions per week, with the restriction of a maximum of two sessions (separated by at least 1 h) per day.

Apparatus. Temporal parameters, stimulus presentation, and recording of responses were controlled by a Macintosh computer.

The temporal intervals were visual and filled. They were marked by clearly visible geometric forms displayed on an Apple color monitor.

Procedure. There were three referent pairs: $L = 650$ msec and $S = 400$ msec (650/400); $L = 700$ msec and $S = 450$ msec (700/450); and $L = 750$ msec and $S = 500$ msec (750/500). For each referent pair, there were six probe durations, $S \leq t \leq L$, arithmetically spaced in 50-msec steps. S referent values were always marked by a red cir-

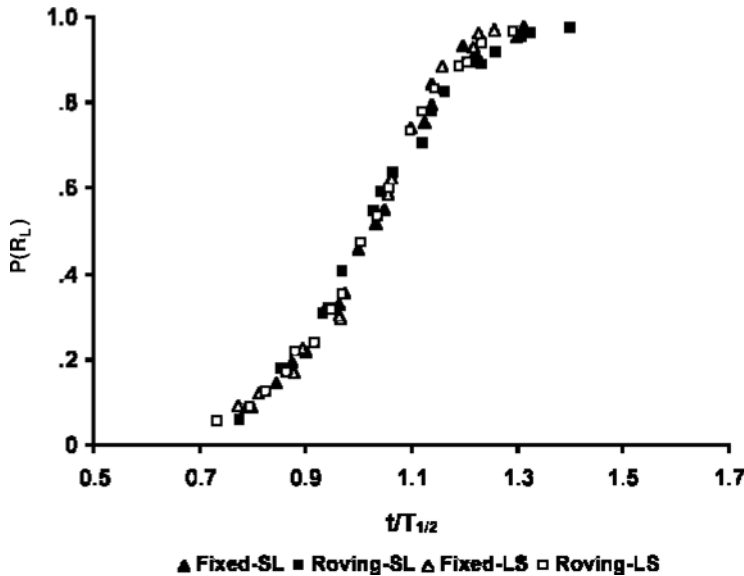


Figure 6. $P(R_L)$, averaged over subjects, as a function of $t/T_{1/2}$ for each referent presentation order (SL = filled symbols, LS = unfilled symbols). The psychometric functions for the three referent pairs from the fixed-referent condition are indicated by triangles and from the roving-referent condition by squares. S , short; L , long. (Experiment 1)

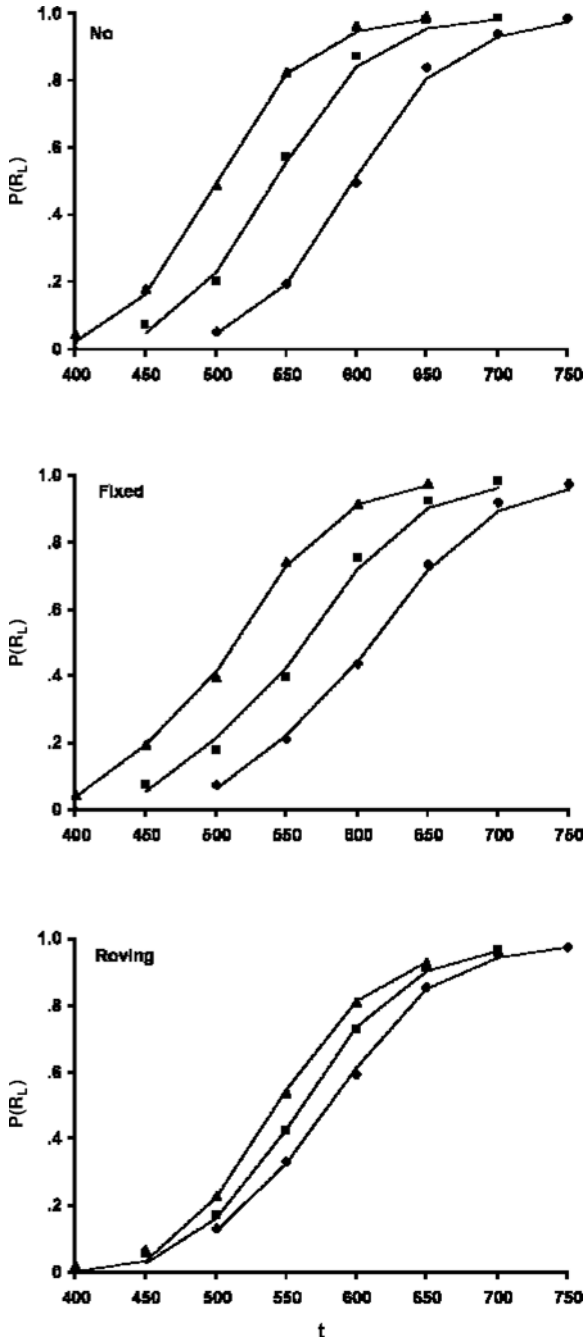


Figure 7. $P(R_L)$, averaged over subjects, as a function of t for each referent pair and each referent condition. The geometric symbols indicate the data (triangles = 650/400 pair, squares = 700/450 pair, and diamonds = 750/500 pair). The lines are derived from the pseudo-logistic model. (Experiment 2)

cle, L referent values by a green circle, and probe values by a black square.

There were three referent conditions: no-referent, fixed-referent, and roving-referent. In the no-referent condition, the referent pair was constant throughout a session. There were eight sample trials at the beginning of each block of trials. On these eight trials, each of the two referent durations were presented four times each in random

order. The eight sample trials were followed by 108 probe trials. On each probe trial, one of the six probe durations, t , was presented. The order of the six t values was random with the restriction that each t value was presented 18 times during the block of 108 probe trials. The referent pair was also constant throughout a session in the fixed-referent condition. There were no sample trials. On each of the 108 probe trials, t was preceded by the two referents, counterbalanced for order: S, L, t , or L, S, t , with an interstimulus interval (ISI) of 1 sec. The roving-referent condition was similar to the fixed-referent condition except that the pair of referents varied from trial to trial. The order of the three referent pairs was random with the restriction that each referent pair was presented 36 times during the block of 108 probe trials. In all referent conditions, the subject's task was to decide whether t was more similar to S or to L . At the termination of the probe interval, the subject indicated the decision by pressing S or the L on the computer keyboard. There was no feedback, and the next probe trial began 1 sec after the response was made.

Referent condition was constant during a session. Each subject participated in 20 sessions (six no-referent, six fixed-referent, and eight roving-referent), each session consisting of four blocks of 108 probe trials. A session lasted about 45 min, depending on referent condition and the self-pacing of the subject's response. For the no- and fixed-referent sessions, the order of the referent pairs over sessions was random, with the restriction that each referent pair was in effect during two sessions. Thus, for each referent pair there was a total of 144 presentations of each of the six probes (18 per block \times 4 blocks \times 2 sessions). For the roving condition, for each referent pair there was a total of 192 presentations of each of the six probes (6 per block \times 4 blocks \times 8 sessions).

Results

For each referent condition, psychometric functions were determined for each of the three referent pairs. The psychometric functions plotted against t , averaged over subjects, are presented in Figure 1. Although the slopes of all nine functions appear similar, the spacing of the functions for the three referent pairs was notably different for the roving condition compared with the no- and fixed-referent conditions.

Our strategy in applying PLM to the data was to determine how well the model could account for the functions from the three referent conditions with a restricted set of parameter values. We assumed that the comparison process was identical for the three referent conditions; that is, on each trial, the probe was compared with $T_{1/2}$. We also assumed that the scalar sources of variance dominated (i.e., $p = 0$ and $c = 0$ in Equation 6). The similar slopes for the nine functions in Figure 1 suggested that γ was constant across the nine functions. Thus, we fit Equation 5 to the data of each subject (using the non-linear fit algorithm from Mathematica), allowing $T_{1/2}$ to vary but keeping γ constant across the nine functions. Table 1 gives, for each subject, the resulting parameter values. The lines in Figure 1 are the values of $P(R_L)$ predicted by PLM. It is clear that PLM provided an excellent fit to the data. Averaged over the nine functions, the sum of the squared deviations between the data and the predictions of PLM (ω^2) ranged from .980 to .999 for the 6 subjects.

Figure 2 plots the nine psychometric functions against t normalized by the bisection point ($t/T_{1/2}$). The functions are shown separately for each subject and also averaged

Table 3
Values of γ and $T_{1/2}$ for Each Referent Condition in Experiment 2

Subject	Referent Pair	γ	$T_{1/2}$						
			Collapsed Across Referent Order			<i>SL</i>	<i>LS</i>	<i>SL</i>	<i>LS</i>
			No	Fixed	Roving	Fixed	Fixed	Roving	Roving
D.F.	650/400	.14	480	521	575	490	558	530	625
	700/450		513	570	584	553	591	545	625
	750/500		572	619	598	591	644	564	641
M.E.	650/400	.09	521	536	524	521	555	513	537
	700/450		569	580	565	569	592	554	574
	750/500		623	628	593	616	642	595	592
C.T.	650/400	.09	471	437	504	439	435	507	502
	700/450		511	482	503	495	470	506	499
	750/500		577	533	509	538	528	511	508
N.A.	650/400	.08	520	534	573	521	548	560	583
	700/450		560	591	585	580	602	570	599
	750/500		610	638	607	628	647	594	521

over subjects. The nine functions superposed across both referent pair and referent condition. The superposition across referent pair is consistent with the scalar property. The superposition across referent condition indicates that scalar variability was identical for the three referent conditions.

$T_{1/2}$ estimated from the data (Table 1) is plotted in Figure 3 as a function of referent pair for each referent condition. The data are shown separately for each subject and also averaged over subjects. In PLM, if responding is unbiased, $T_{1/2}$ should be located at the value of t where the S distribution crosses the L distribution. Killeen et al. (1997) provided an approximated solution for unbiased $T_{1/2}$,

$$\text{unbiased } T_{1/2} = (1+k) \frac{2}{\frac{1}{S} + \frac{1}{L}},$$

$$\text{where } k = \frac{\gamma}{2} \ln\left(\frac{L}{S}\right). \quad (7)$$

Since then, Killeen (personal communication, 1999) has provided an exact solution (see the Appendix for derivation of Equation 8),

$$\text{unbiased } T_{1/2} = \frac{(S - RL) + \sqrt{(S - RL)^2 - (1 - R)[(\gamma S)^2 \ln(R)]}}{1 - R},$$

$$\text{where } R = \left(\frac{S}{L}\right)^2. \quad (8)$$

Unbiased $T_{1/2}$ (Equation 8) is also plotted in Figure 3. It is clear that $T_{1/2}$ deviated from unbiased $T_{1/2}$ and that there was considerable between-subjects variability in the placement of $T_{1/2}$ relative to unbiased $T_{1/2}$, with some subjects setting $T_{1/2}$ larger than unbiased $T_{1/2}$, and other subjects setting $T_{1/2}$ smaller than unbiased $T_{1/2}$. However, there was between-subjects consistency across referent conditions, which is reflected in $T_{1/2}$ averaged over subjects. For the no- and fixed-referent conditions, as the refer-

ents (and therefore unbiased $T_{1/2}$) increased in duration, so did $T_{1/2}$. In contrast, $T_{1/2}$ was relatively flat across the three referent pairs for the roving condition. The trends seen in Figure 3 were confirmed statistically. For each subject, the best linear fit to the function relating $T_{1/2}$ to referent pair was determined for each referent condition. These slopes are shown in Table 2. A within-subjects analysis of variance (ANOVA) on the slopes revealed a significant effect of reference condition [$F(2,10) = 31.38, p < .001$]. Post hoc Tukey tests indicated that the roving slopes differed from both the no-referent ($p < .001$) and the fixed-referent ($p < .001$) slopes, and that the no- and the fixed-referent slopes did not differ ($p > .05$).

Thus far, the data from the fixed and roving conditions have been collapsed over the presentation order of the referents. There is considerable evidence in the literature, however, that presentation order of temporal stimuli influences responding (e.g., Allan, 1977, 1979; Allan & Gibbon, 1994; Hellström, 1985; Jamieson & Petrusic, 1975). Figure 4 presents the psychometric functions, averaged over subjects, separately for each referent presentation order, *SL* and *LS*. For each referent pair, $P(R_L)$ was larger for the *SL* order than for the *LS* order. We refit PLM to the fixed and roving data, separately for each referent order, using the value of γ from the collapsed functions but allowing $T_{1/2}$ to vary across referent order. The values of $T_{1/2}$ are shown in Table 1 for each subject, averaged over subjects, in Figure 5. For both referent conditions, the value of $T_{1/2}$ was larger for *LS* than for *SL*, indicating that a probe was more likely to be categorized as similar to the second trial referent— R_L on *SL* trials and R_S on *LS* trials. It should also be noted that the pattern seen in Figure 3 for the collapsed data is present for each referent order. That is, for the fixed condition $T_{1/2}$ increased as the referents (and therefore unbiased $T_{1/2}$) increased, whereas for the roving condition $T_{1/2}$ was relatively flat across the three referent pairs.

Figure 6 shows the 12 functions (two referent orders, two referent conditions, and three referent pairs) plotted against $t/T_{1/2}$. It is clear that there is superposition of the

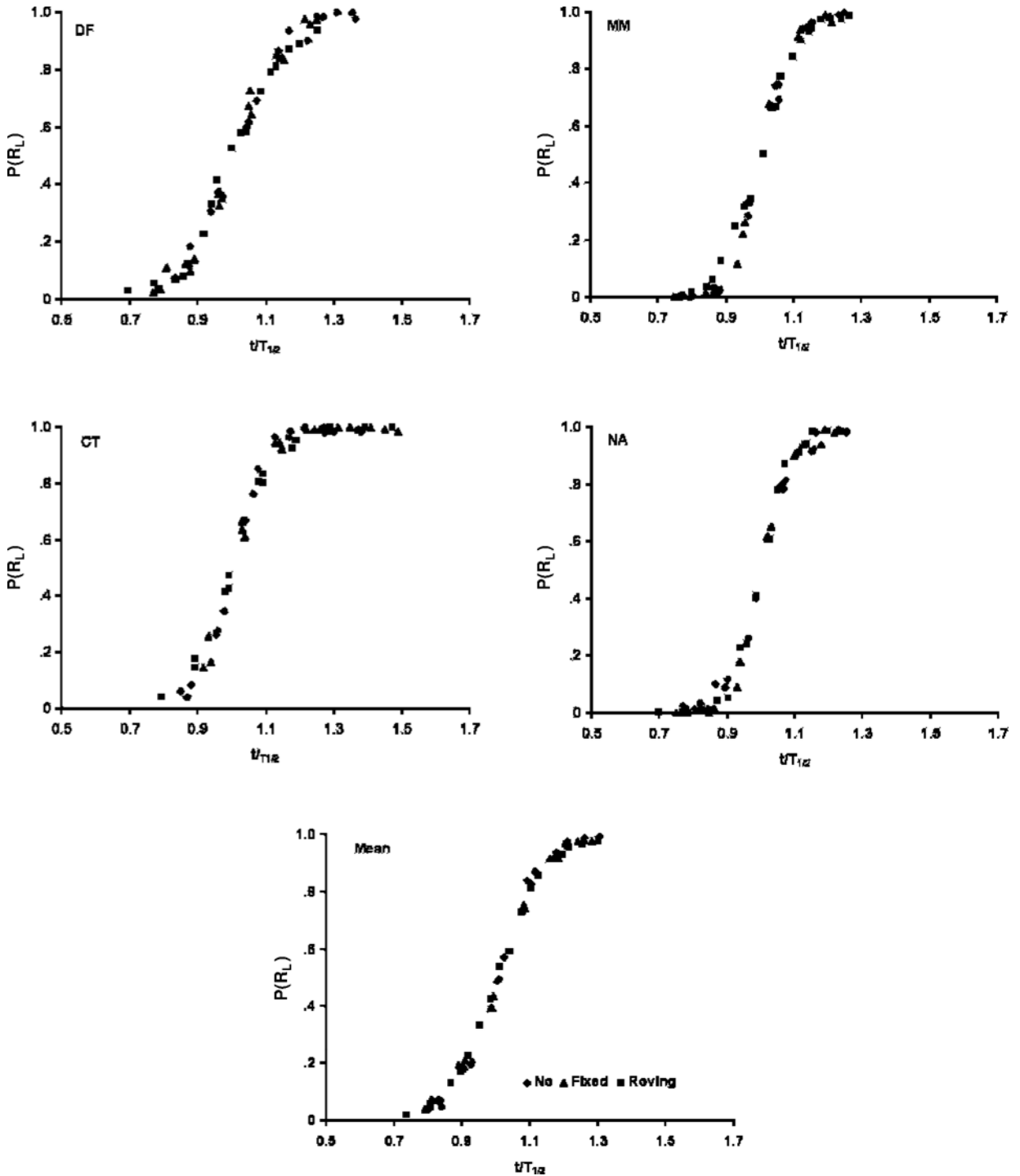


Figure 8. $P(R_L)$ as a function of $t/T_{1/2}$. The psychometric functions for the three referent pairs from the no-referent condition are indicated by diamonds, from the fixed-referent condition by triangles, and from the roving-referent condition by squares. The functions are shown for each subject and also averaged over subjects. (Experiment 2)

12 functions. As was seen earlier with the data collapsed across referent order (Figure 2), the functions superposed across referent condition and referent pair. Figure 6 reveals superposition in relative time for the two referent orders as well.

Discussion

In fitting PLM to the data from the three referent conditions, we assumed that the only source of variability was scalar and that the scalar variability was located in the probe. The excellent fit of PLM, with the same value

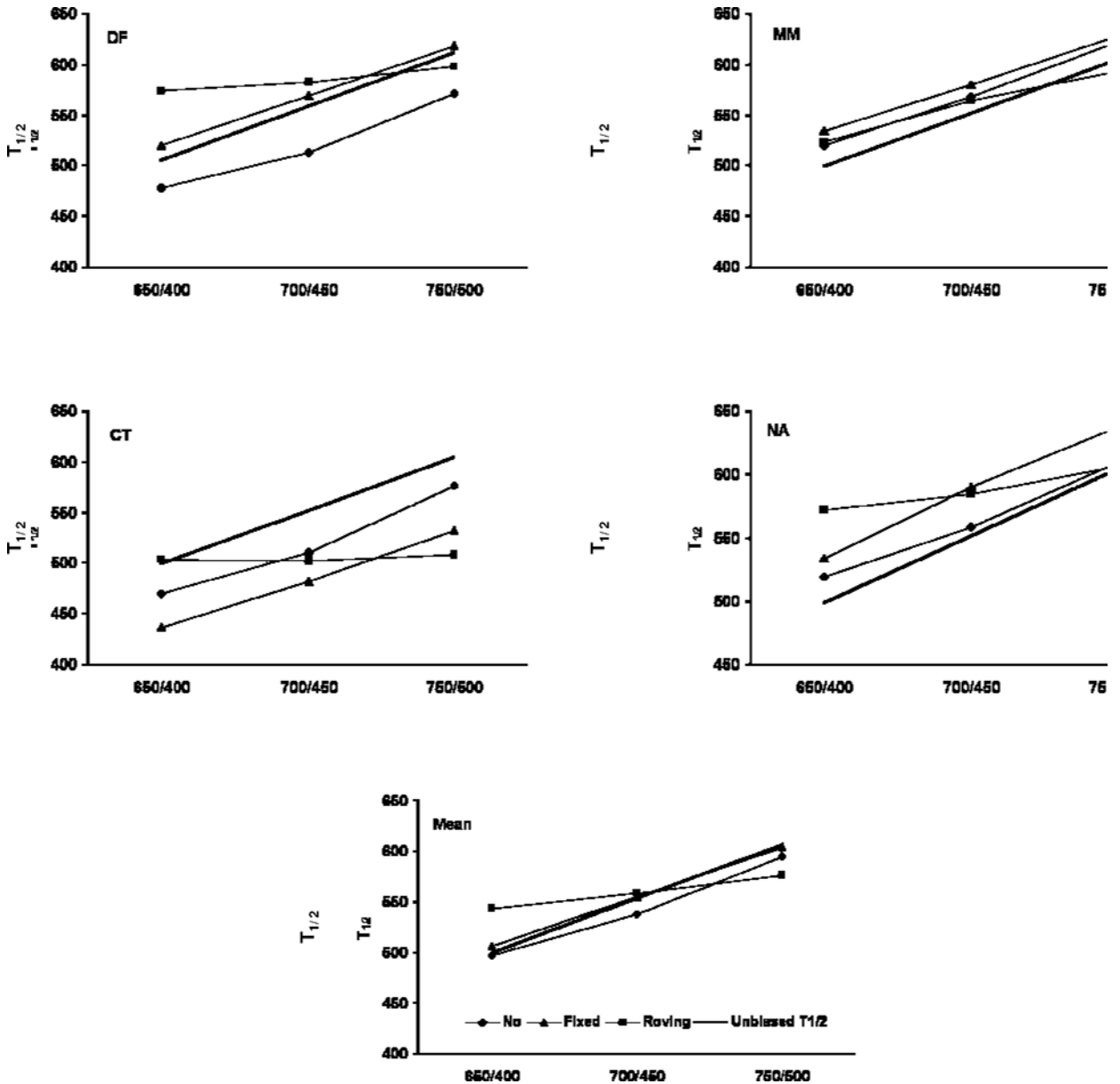


Figure 9. $T_{1/2}$ as a function of referent pair for each referent condition (diamonds = no, triangles = fixed, squares = roving). Also shown is the unbiased $T_{1/2}$ (continuous line) for each referent pair. The data are presented for each subject and also averaged over subjects. (Experiment 2)

of γ , to the data from the three referent conditions provides strong support for these assumptions. If the absence of trial referents or the variation in trial referents was an important contribution to the overall variability, then a good fit would not have been obtained with a constant value of γ .

Although the referents did not affect variability as reflected in γ , they did influence the placement of the criterion $T_{1/2}$. When the referents were constant (no- and fixed-referent), the placement of the criterion was dependent on the referent pair in that $T_{1/2}$ increased with

referent pair. When the referents varied from trial to trial (roving), very similar criteria were set for the three referent pairs. In the fixed and roving conditions, referent order also influenced the placement of the criterion.

EXPERIMENT 2

In Experiment 1, we marked the temporal intervals in the visual modality, as did Rodriguez-Girones and Kacelnik (in press). In Experiment 2, we marked the temporal intervals in the auditory modality, as we did in Allan

Table 4
Values of the Slope of the Linear Function Relating $T_{1/2}$ to Referent Pair for Each Referent Condition in Experiment 2

Subject	No	Fixed	Roving
D.F.	46.25	48.80	11.90
M.E.	50.60	46.30	34.65
C.T.	53.20	48.15	2.25
N.A.	44.90	51.90	17.30
<i>M</i>	48.74	48.79	16.53

and Gibbon (1991). Experiment 2, thus, allows the results of Experiment 1 to be generalized for different duration markers.

Method

Four new subjects (2 females and 2 males) participated in this experiment. As in Experiment 1, they were graduate students in the psychology department at McMaster University who were paid for their participation.

Except for the duration markers, Experiment 2 was identical to Experiment 1. In Experiment 1, the temporal intervals were filled and were marked by the presentation of a geometric form at the center of the computer monitor. In Experiment 2, the temporal intervals were also filled but were marked by a 1000-Hz pure tone presented through headphones (monaural) at a comfortable listening level. The tones had a 10-msec rise-decay time.

In Experiment 1, *S* referent values were marked by a red circle and *L* referent values by a green circle. In the eight sample trials in the no-referent condition, the color of the circle identified to the subject whether the referent was *S* or *L*. In the fixed and roving conditions, the color of the circles distinguished *SL* trials from *LS* trials. In Experiment 2, this information about the referents was provided by words printed in the center of the computer monitor. On the eight sample trials in the no-referent condition, the word *short* accompanied *S* and the word *long* accompanied *L*. In the fixed and roving conditions, either *SL* or *LS* was presented on each trial. This visual information appeared simultaneously with the first referent and remained on the screen until the subject responded.

Results and Discussion

The psychometric functions plotted against *t*, averaged over the 4 subjects, are presented in Figure 7, with the fixed and roving functions collapsed over referent order. As in Experiment 1, the spacing of the three roving functions differed from the spacing for the no- and fixed-referent functions.

As in Experiment 1, we fit Equation 5, ($p = 0$ and $c = 0$) to the data of each subject with the constraint that γ was constant across the nine functions. Table 3 gives, for each subject, the resulting parameter values. The lines in Figure 8 are the values of $P(R_L)$ predicted by PLM. Again, PLM provided an excellent fit to the data. Averaged over the nine functions, ω^2 ranged from .986 to .996 for the 4 subjects. The superposition in relative time across referent pair and referent condition is shown in Figure 8.

Figure 9 shows $T_{1/2}$ estimated from the data (Table 3) and unbiased $T_{1/2}$ specified by the model (Equation 8) for each referent pair, separately for each referent condition. As in Experiment 1, $T_{1/2}$ deviated from unbiased $T_{1/2}$, and there was between-subjects variability in the placement of

$T_{1/2}$ relative to unbiased $T_{1/2}$. However, again there was between-subjects consistency across referent condition, which is shown in $T_{1/2}$ averaged over subjects. $T_{1/2}$ increased with unbiased $T_{1/2}$ for the no- and fixed-referent conditions, whereas $T_{1/2}$ was relatively flat across the three referent pairs for the roving condition. As in Experiment 1, the best linear fit to the function relating $T_{1/2}$ to referent pair was determined for each referent condition. These slopes are shown in Table 4. A within-subjects ANOVA on the slopes revealed a significant effect of reference condition [$F(2,6) = 18.09, p < .01$]. Post hoc Tukey tests indicated that the roving slopes differed from both the no- ($p < .01$) and the fixed- ($p < .01$) referent slopes, and that the no- and fixed-referent slopes did not differ ($p > .05$).

Figure 10 presents the psychometric functions, averaged over subjects, separately for each referent order, *SL* and *LS*. As in Experiment 1, we refit PLM to the fixed and roving data, separately for each referent order, using the value of γ from the collapsed functions. The values of $T_{1/2}$ are shown in Table 3 for each subject, and are shown, averaged over subjects, in Figure 11. The value of $T_{1/2}$ depended on the order of the referents, being larger for *LS* than for *SL*. Figure 12 shows the 12 functions (two referent orders, two referent conditions, and three referent pairs) plotted against $t/T_{1/2}$. It is clear that there is superposition of the 12 functions.

The results from Experiment 2 replicate those from Experiment 1. The only notable difference is that, in general, γ was smaller in Experiment 2. Within the context of PLM, γ reflects the variability in the perception of the probe. The relatively greater variability for visual markers than for auditory markers is consistent with data reported by others (e.g., Penney et al., 1998; Wearden, Edwards, Fakhri, & Percival, 1998).

GENERAL DISCUSSION

PLM provided an excellent account of the data from the two experiments. In our analysis, the dominant source of variability was scalar. This scalar variability was located in the perception of the probe and was independent of the structure of the bisection task. The probe was compared with a criterion rather than directly with the referents, even when trial referents were available. The role of the referents was to set the criterion. When the referent pair was constant throughout a session, as in the no- and fixed-referent conditions, the value of the criterion depended on referent pair, increasing as the values of the referents increased. When the referent pairs varied within a session, as in the roving condition, a relatively constant criterion was set for all referent pairs. Any variability in the location of the criterion for a referent pair (e.g., c in Equation 6) was negligible relative to the scalar variability given the excellent fit of PLM to the data with $c = 0$.

Our results differ from those reported by Rodriguez-Girones and Kacelnik (in press). Their fixed psychometric functions were steeper than their roving functions,

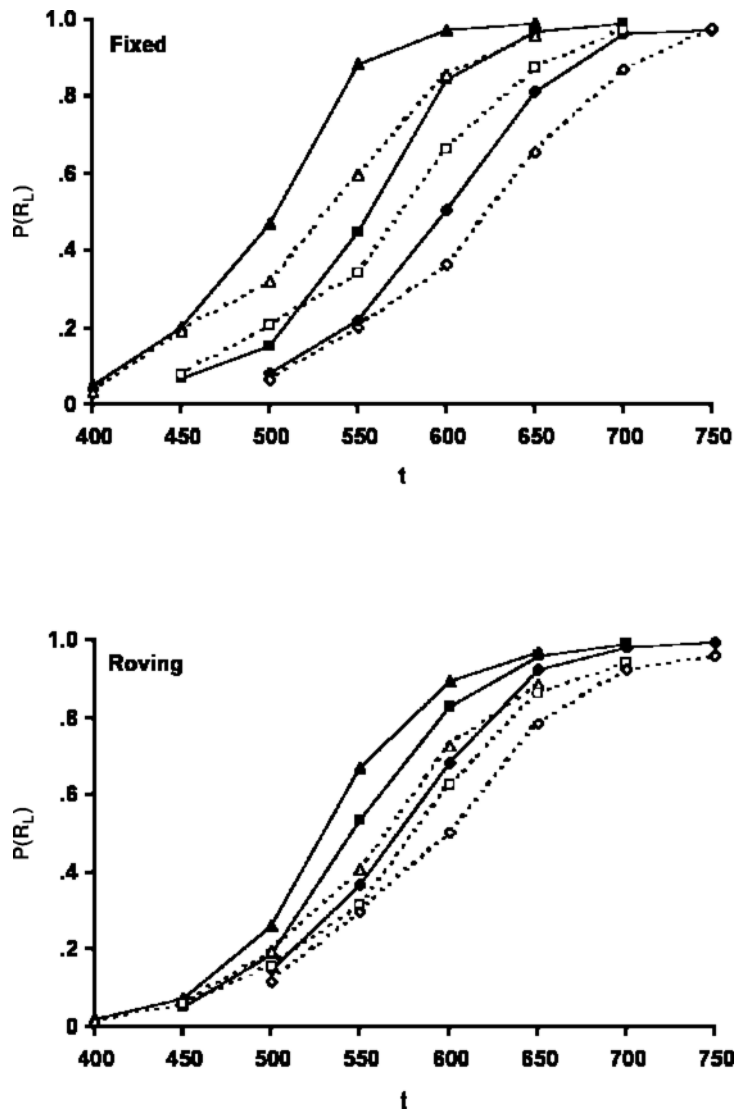


Figure 10. $P(R_L)$, averaged over subjects, as a function of t for each referent pair (triangles = 650/400, squares = 700/450, and diamonds = 750/500) and each referent presentation order (SL = filled symbols and continuous lines, LS = unfilled symbols and dotted lines). S, short; L, long. (Experiment 2)

whereas our functions had similar slopes. This difference across studies might be due to the greater variability in the referent pairs in the Rodriguez-Girones and Kacelnik experiment. However, it might also be because Rodriguez-Girones and Kacelnik did not equate probe spacing and referent range across their roving and fixed conditions. As noted, there is evidence that these variables influence bisection (see Allan, 1998, 1999). Clearly, future research needs to systematically increase the number of referent pairs in the roving condition while controlling for spacing and range across bisection conditions.

Although our results are consistent with PLM, they are not consistent with the log-normal model proposed by Rodriguez-Girones and Kacelnik (in press). According to that model, the psychometric functions should not have identical slopes across the three referent conditions. In deriving their model, Rodriguez-Girones and Kacelnik assumed that in the no-referent condition, the probe would be compared with memories of the referents, whereas in the roving condition, the probe would be compared with the trial referents. If the dominant source of scalar variability was in perception, then the fixed-referent would

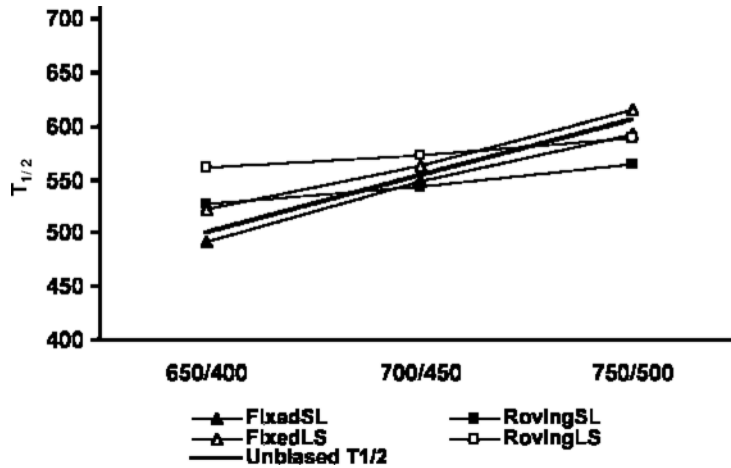


Figure 11: $T_{1/2}$, averaged over subjects, as a function of referent pair separately for each referent presentation order (SL = filled symbols, LS = unfilled symbols) from the fixed (triangles) and roving (squares) conditions. Also shown is the unbiased $T_{1/2}$ (continuous line) for each referent pair. (Experiment 2)

be equivalent to the no-referent condition, and the slopes for no- and fixed-referent conditions would not differ, but both would be steeper than that for roving. In contrast, if the dominant source of scalar variability was in memory for the referents, then fixed would be equivalent to roving, and the slopes for fixed and roving would not differ, but both would be steeper than in the no-referent condition.

The presence of trial referents in the fixed and roving conditions introduces referent order (SL or LS) as a vari-

able in bisection. The data from both experiments revealed large and consistent referent order effects. The same probe value was more likely to be categorized as similar to the second trial referent— R_L on SL trials and R_S on LS trials. The same pattern of results was reported by Rodriguez-Girones and Kacelnik (in press) with regard to referent order effects. PLM located the referent order effect in the SDT criterion. According to that analysis, referent order affected the response to the probe, but not the perception

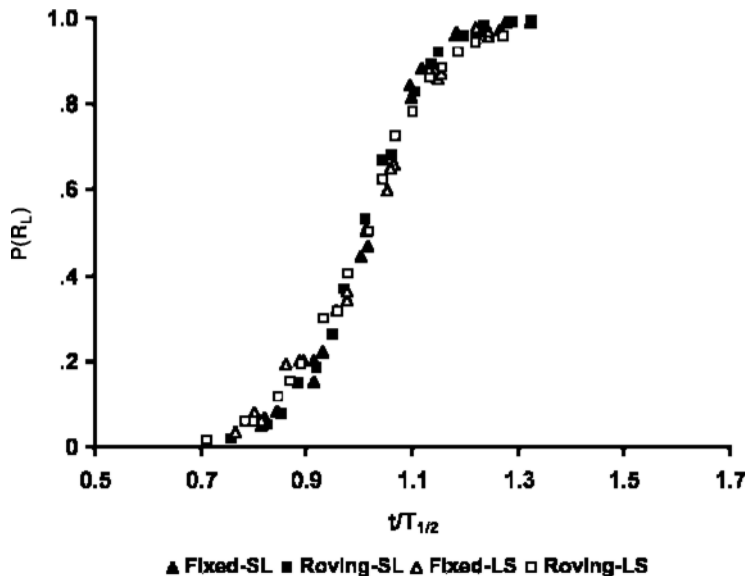


Figure 12. $P(R_L)$, averaged over subjects, as a function of $t/T_{1/2}$ for each referent presentation order (SL = filled symbols, LS = unfilled symbols). The psychometric functions for the three referent pairs from the fixed-referent condition are indicated by triangles and from the roving-referent condition by squares. (Experiment 2)

of the probe. Placing the referent order effect in the SDT criterion is consistent with our earlier position on presentation order effects (e.g., Allan, 1977, 1979; Allan & Gibbon, 1994). In the present experiments, referent presentation order was marked by color in Experiment 1 and by labels in Experiment 2. This explicit identification of referent order might be responsible for the dependence of criterion location on referent order. If so, the presentation order effect should be attenuated if the order markers are removed.

The finding in the present experiments that the probe is compared with a criterion is consistent with our earlier results using different psychophysical procedures (Allan, 1977; Allan et al., 1974). Allan et al. compared performance in two duration discrimination tasks, single stimulus (SS) and forced choice (FC). In SS, one of two duration values, S or L , was presented on each trial, and the task was to categorize these values as short (R_S) or long (R_L). In FC, both duration values were presented sequentially on each trial, counterbalanced for order (SL or LS), and the task was to indicate the order, $R_S R_L$ or $R_L R_S$. As in the present analysis, Allan et al. also conducted their analysis within a SDT framework. According to standard SDT analysis (see Green & Swets, 1966; Kinchla & Allan, 1969), if the two durations on an FC trial are directly compared, FC performance (d'_{FC}) should be better than SS performance (d'_{SS}). Allan et al., however, found that d' did not differ for the two tasks. Allan et al. also varied the ISI between the two duration values in the FC task. A number of earlier studies across a variety of sensory dimensions (e.g., auditory intensity, auditory frequency, visual movement, spatial position) indicated that d'_{FC} decreased as ISI increased (e.g., Allan, 1968; Bull & Cuddy, 1972; Kinchla & Allan, 1969, 1970; Kinchla & Smyzer, 1967; Moss, Myers, & Filmore, 1970). Allan et al., however, found that d'_{FC} for the temporal dimension remained constant over these same ISI values. They concluded that for temporal judgments, the FC comparison was similar to the SS comparison. In both tasks, each trial value was compared with a criterion duration value. Thus, in the FC task, rather than the two trial durations being compared with each other, each was compared with an internal criterion value and each was categorized as R_S or R_L . If the two categorizations were the same, the subject guessed. Later, Allan (1977) reported the same result for the roving-standard FC task, where the pair of duration values varied from trial to trial.

Allan et al. (1974) and Allan (1977) concluded that when multiple temporal intervals were presented on a trial, each interval is coded categorically. The data from the present experiment are compatible with that conclusion. Other data exist, however, which are not. For example, nonhuman data indicate that coding is analogical rather than categorical. Such nonhuman data have been generated, for example, in a symbolic temporal matching-to-sample task (e.g., Spetch, Grant, & Kelly, 1996) where an animal is trained with one response reinforced after a short-duration sample and a different response reinforced

after a long-duration sample. To test for the memory of the sample duration, delays are introduced between the termination of the sample duration and the availability of the choice opportunity. Numerous (but not all) experiments have indicated that when the test delay is longer than the training delay, pigeons typically select the choice associated with the short sample. This "choose-short" effect has been taken to indicate that pigeons code durations analogically and that the analogical representation foreshortens during the delay interval.

Wearden and Ferrara (1993) were interested in developing a human analogue for the nonhuman temporal matching-to-sample task. The task that they used was similar to the Allan (1977) roving-standard FC task. Unlike Allan (1977), however, they concluded that their data revealed subjective shortening and therefore analogical temporal memory. Thus, the roving-standard FC task appears to result sometimes in categorical (Allan, 1977) and sometimes in analogical (Wearden & Ferrara, 1993) coding. The two sets of experiments differed in numerous ways, however, and systematic experiments need to be conducted to identify which of these methodological differences are important. For example, the Wearden and Ferrara (1993) subjects participated in a short session consisting of few trials, whereas the Allan (1977) subjects participated in multiple sessions consisting of many trials. It could be that temporal coding becomes categorical with practice.

The bisection data from the present experiments are consistent with categorical coding of multiple trial durations in that the probe was not directly compared with the trial referents. It would be of theoretical interest to vary the referent-probe interval as well as to present the referent pair after the probe. If temporal coding is categorical, then variations in referent-probe interval and order should have no effect on γ .

Categorical coding of trial durations would suggest that human memory of a temporal interval is poor. In fact we know little about this aspect of time perception. Although the original purpose of the bisection task was to study temporal memory, the data from the present experiments suggest that referent memory is not a dominant source of variability in bisection. In fact, relatively few studies have been directed at the study of human temporal memory, and, as noted, the scant data that do exist do not paint a clear and consistent picture.

REFERENCES

- ALLAN, L. G. (1968). Visual position discrimination: A model relating temporal and spatial factors. *Perception & Psychophysics*, *4*, 267-278.
- ALLAN, L. G. (1977). The time-order error in judgments of duration. *Canadian Journal of Psychology*, *31*, 24-31.
- ALLAN, L. G. (1979). The perception of time. *Perception & Psychophysics*, *26*, 340-354.
- ALLAN, L. G. (1983). Magnitude estimation of temporal intervals. *Perception & Psychophysics*, *33*, 29-42.
- ALLAN, L. G. (1998). The influence of scalar timing on human timing. *Behavioural Processes*, *44*, 101-117.
- ALLAN, L. G. (1999). Understanding the bisection psychometric func-

- tion. In W. Uttal and P. Killeen (Eds.), *Fechner Day 99: Proceedings of the Fifteenth Annual Meeting of the International Society for Psychophysics*. Tempe: Arizona State University.
- ALLAN, L. G. & GIBBON, J. (1991). Human bisection at the geometric mean. *Learning & Motivation*, **22**, 39-58.
- ALLAN, L. G., & GIBBON, J. (1994). A new temporal illusion or the TOE once again? *Perception & Psychophysics*, **55**, 227-229.
- ALLAN, L. G., KRISTOFFERSON, A. B., & RICE, M. E. (1974). Some aspects of perceptual coding of duration in visual duration discrimination. *Perception & Psychophysics*, **15**, 83-88.
- BULL, A. R., & CUDDY, L. L. (1972). Recognition memory for pitch of fixed and roving stimulus tones. *Perception & Psychophysics*, **11**, 105-109.
- CREELMAN, D. D. (1962). Human discrimination of auditory duration. *Journal of the Acoustical Society of America*, **34**, 582-593.
- GIBBON, J. (1981). On the form and location of the bisection function for time. *Journal of Mathematical Psychology*, **24**, 58-87.
- GIBBON, J., & CHURCH, R. M. (1981). Time left: Linear vs. logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes*, **7**, 87-108.
- GREEN, D. M., & SWETS, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- HELLSTRÖM, A. (1985). The time-order error and its relatives: Mirrors of cognitive processes in comparing. *Psychological Bulletin*, **97**, 35-610.
- JAMIESON, D. G., & PETRUSIC, W. M. (1975). Presentation order effects in duration discrimination. *Perception & Psychophysics*, **17**, 197-202.
- KILLEEN, P. R., FETTERMAN, J. G., & BIZO, L. A. (1997). Time's causes. In C. M. Bradshaw & E. Szabadi (Eds.), *Time and behaviour: Psychological and neurological analyses* (pp. 79-131). Amsterdam: Elsevier.
- KILLEEN, P. R., & WEISS, N. A. (1987). Optimal timing and the Weber function. *Psychological Review*, **94**, 455-468.
- KINCHLA, R. A., & ALLAN, L. G. (1969). A theory of visual movement perception. *Psychological Review*, **76**, 537-558.
- KINCHLA, R. A., & ALLAN, L. G. (1970). Visual movement perception: A comparison of sensitivity to vertical and horizontal movement. *Perception & Psychophysics*, **8**, 399-405.
- KINCHLA, R. A., & SMYZER, F. (1967). A diffusion model of perceptual memory. *Perception & Psychophysics*, **2**, 219-229.
- MOSS, S. M., MYERS, J. L., & FILMORE, T. (1970). Short-term recognition memory of tones. *Perception & Psychophysics*, **7**, 369-373.
- PENNEY, T. B., ALLAN, L. G., MECK, W. H., & GIBBON, J. (1998). Memory mixing in duration bisection. In D. A. Rosenbaum & C. E. Collyer (Eds.), *Timing of behavior: Neural, psychological, and computational perspectives* (pp. 165-193). Cambridge MA: MIT Press.
- RODRIGUEZ-GIRONES, M. A., & KACELNIK, A. (in press). Relative importance of perceptual and mnemonic variance in human temporal bisection. *Quarterly Journal of Experimental Psychology*.
- SPETCH, M. L., GRANT, D. S., & KELLY, R. (1996). Procedural determinants of coding processes in pigeons' memory for duration. *Learning & Motivation*, **27**, 179-199.
- WEARDEN, J. H. (1991). Human performance on an analogue of an interval bisection task. *Quarterly Journal of Experimental Psychology*, **43B**, 59-81.
- WEARDEN, J. H., EDWARDS, H., FAKHRI, M., & PERCIVAL, A. (1998). Why "sounds" are judged longer than "lights": Application of a model of the internal clock in humans. *Quarterly Journal of Experimental Psychology*, **51B**, 97-120.
- WEARDEN, J. H., & FERRARA, A. (1993). Subjective shortening in humans' memory for stimulus duration. *Quarterly Journal of Experimental Psychology*, **46B**, 163-186.
- WEARDEN, J. H., & FERRARA, A. (1995). Stimulus spacing effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology*, **48B**, 289-310.
- WEARDEN, J. H., & FERRARA, A. (1996). Stimulus range effects in temporal bisection by humans. *Quarterly Journal of Experimental Psychology*, **49B**, 24-44.
- WEARDEN, J. H., ROGERS, P., & THOMAS, R. (1997). Temporal bisection in humans with longer stimulus durations. *Quarterly Journal of Experimental Psychology*, **50B**, 79-94.

NOTES

1. The partition task is similar to the many-to-few task described by Allan (1979) in that there are multiple stimuli and two responses, but the two tasks differ in a fundamental way. In partition, there is no trial-by-trial feedback, whereas as many-to-few, there is.

2. Gibbon's (1981) SDT model differed from the first SDT model for time proposed by Creelman (1962). In the Creelman model, the variance was proportional to the mean. Also, Creelman did not derive a psychometric function.

APPENDIX

$$\left(\frac{T_{V/2}-S}{\sqrt{2}\sigma_S}\right)^2 - \ln\left(\frac{1}{\sigma_S}\right) = \left(\frac{T_{V/2}-L}{\sqrt{2}\sigma_L}\right)^2 - \ln\left(\frac{1}{\sigma_L}\right)$$

$$\left(\frac{T_{V/2}-S}{\sqrt{2}\sigma_S}\right)^2 = \left(\frac{T_{V/2}-L}{\sqrt{2}\sigma_L}\right)^2 - \left[\ln\left(\frac{1}{\sigma_L}\right) - \ln\left(\frac{1}{\sigma_S}\right)\right]$$

$$\left(\frac{T_{V/2}-S}{\sqrt{2}\sigma_S}\right)^2 = \left(\frac{T_{V/2}-L}{\sqrt{2}\sigma_L}\right)^2 - \ln\left(\frac{\sigma_S}{\sigma_L}\right)$$

$$\left(\frac{T_{V/2}-S}{\sigma_S}\right)^2 = \left(\frac{T_{V/2}-L}{\sigma_L}\right)^2 - 2\ln\left(\frac{\sigma_S}{\sigma_L}\right)$$

$$(T_{V/2}-S)^2 = \frac{\sigma_S^2}{\sigma_L^2}(T_{V/2}-L)^2 - 2\sigma_S^2 \ln\left(\frac{\sigma_S}{\sigma_L}\right)$$

Let $R = \left(\frac{S}{L}\right)^2 = \left(\frac{\sigma_S}{\sigma_L}\right)^2 \neq 1$

$$(T_{V/2}-S)^2 = R(T_{V/2}-L)^2 - 2\sigma_S^2 \ln\sqrt{R}$$

$$\left(T_{V/2}^2 - 2T_{V/2}S + S^2\right) = R\left(T_{V/2}^2 - 2T_{V/2}L + L^2\right) - 2\sigma_S^2 \ln\sqrt{R}$$

$$T_{V/2}^2(1-R) + T_{V/2}(2RL-2S) + S^2 - RL^2 + 2\sigma_S^2 \ln\sqrt{R} = 0$$

$$T_{V/2}^2(1-R) + T_{V/2}(2RL-2S) + 2\sigma_S^2 \ln\sqrt{R} = 0$$

let

$$a = (1-R), \quad b = 2(RL-S), \quad c = 2\sigma_S^2 \ln\sqrt{R}$$

then

$$T_{V/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

$$T_{V/2} = \frac{2(S-RL) \pm \sqrt{4(S-RL)^2 - 4(1-R)(2\sigma_S^2 \ln\sqrt{R})}}{2(1-R)}$$

$$T_{V/2} = \frac{(S-RL) \pm \sqrt{(S-RL)^2 - (1-R)(2\sigma_S^2 \ln\sqrt{R})}}{(1-R)}$$

for $\sigma_t = \gamma t$ (Equation 1)

$$T_{V/2} = \frac{(S-RL) \pm \sqrt{(S-RL)^2 - (1-R)(2\gamma^2 S^2 \ln\sqrt{R})}}{(1-R)}$$

$$T_{V/2} = \frac{(S-RL) \pm \sqrt{(S-RL)^2 - (1-R)(\gamma^2 S^2 \ln R)}}{(1-R)}$$