BIAS IN SELF-EVALUATION: SIGNAL PROBABILITY EFFECTS

THOMAS S. CRITCHFIELD

AUBURN UNIVERSITY

Two experiments examined apparent signal probability effects in simple verbal self-reports. After each trial of a delayed matching-to-sample task, young adults pressed either a "yes" or a "no" button to answer a computer-presented query about whether the most recent choice met a point contingency requiring both speed and accuracy. A successful matching-to-sample choice served as the "signal" in a signal-detection analysis of self-reports. Difficulty of matching to sample, and thus signal probability, was manipulated via the number of nonmatching sample and comparison stimuli. In Experiment 1, subjects exhibited a bias ($\log b$) for reporting matching-to-sample success when success was frequent, and no bias or a bias for reporting failure when success was infrequent. Contingencies involving equal conditional probabilities of point consequences for "I succeeded" and "I failed" reports had no systematic effect on this pattern. Experiment 2 found signal probability effects to be evident regardless of whether referent-response difficulty was manipulated in different conditions or within sessions. These findings indicate that apparent signal probability effects in self-report bias that were observed in previous studies probably were not an artifact of contingencies intended to improve self-report accuracy or of the means of manipulating signal probability. The findings support an analogy between simple self-reports and psychophysical judgments and bolster the conclusion of Critchfield (1993) that signal probability effects can influence simple self-reports much as they do reports about external stimuli in psychophysical experiments.

Key words: self-reports, verbal behavior, matching to sample, signal detection, discriminability, bias, signal probability effects, button press, button release, college students

In psychophysical studies, reports may not correspond perfectly to the experimental stimuli that presumably serve as their referents (e.g., lights, tones, vibrotactile stimulation). Part of the variation in subject reports is accounted for by genuine failures to detect the relevant stimulus—for example, in cases in which stimulus intensity is below the momentary threshold of the relevant sensory system, or the stimulus is essentially indistinguishable from background stimulation. Another portion of the variability in subject reports in psychophysical experiments may be accounted for by bias, a predisposition to make certain kinds of reports regardless of the actual status of the stimulus. Bias has been found to be especially sensitive to reinforcement contingencies and stimulus (or signal) probability (e.g., Gescheider, 1985; Green & Swets, 1966; McCarthy & Davison, 1981a). In the case of signal probability, the more frequently a stimulus occurs, the greater the likelihood of a bias for reporting the presence of that stimulus (Craig,

1978, 1979; Hume, 1974a, 1974b; Schusterman & Johnson, 1975).

A recent series of studies on the experimental analysis of "self-awareness" revealed possible parallels between simple verbal self-reports and the reports about external stimuli made by subjects in psychophysical experiments (Critchfield, 1993; Critchfield & Perone, 1990a, 1990b, 1993). The procedure required subjects, after each trial of a delayed matching-to-sample (DMTS) task, to report whether their last response was successful or unsuccessful in meeting a point contingency that required responses to be both accurate and faster than a time limit. The stringency of the time limit and the number of DMTS sample and comparison stimuli were manipulated to produce a broad range of success rates across experimental conditions. A signal-detection analysis (Green & Swets, 1966), conducted using the response categories shown in Figure 1, revealed that self-report bias tended to track DMTS success rate. That is, when success was frequent, subjects tended to exhibit a bias for reporting success. As success became less frequent, the bias became less pronounced, and at extremely low DMTS success rates, some subjects exhibited a bias for reporting failure. Considering a successful DMTS response as the "signal," these results are consistent with

Thanks are due to Peter Harzem and Larry Alferink for use of laboratory space and Michael Perone for use of equipment. Portions of the data were presented at the 1992 Association for Behavior Analysis convention in San Francisco. Address reprint requests and correspondence to the author at the Department of Psychology, Auburn University, Auburn, Alabama 36849.

	DMTS RESPONSE	
	Successful	Unsuccessful
"I succeeded"	HIT	FALSE ALARM
"I failed"	MISS	CORRECT REJECTION

Fig. 1. Contingency matrix showing self-reports of delayed matching-to-sample success as a function of actual success; cells are labeled using response classes derived from signal-detection theory.

signal probability effects like those observed in psychophysical studies. Critchfield (1993) concluded that self-evaluation, in the form of verbal self-reports, can be subject to the same biasing effects that influence psychophysical judgments.

At a surface level, there is nothing novel about drawing parallels between verbal self-reports and psychophysical judgments (in fact, the historical roots of psychophysics lie partly in the search for a technology of self-observation; Boring, 1953; Danziger, 1980). But the analogy takes on new utility in the context of experimental procedures that permit the systematic manipulation of referent self-events. This is especially true given the relative paucity of research on verbal behavior conducted from an operant perspective (e.g., Oah & Dickinson, 1989), and the central role that self-evaluation plays in many current theoretical accounts of human behavior. In the former case, contact with established research traditions may suggest experimental methods adaptable to the analysis of verbal behavior. In the latter case, many psychologists assume that moment-to-moment acts of self-evaluation serve as an important determinant of, or precursor to, other behavior. Examples include cognitive dissonance theory (Festinger, 1957), control theory (Carver & Scheier, 1981), and some aspects of cognitive theories of information processing (e.g., metacognition; Flavell & Wellman, 1977). Variables likely to affect the correspondence between self-evaluation and actual performance should be of special inter-

est to scholars approaching human behavior from these points of view.

Two experiments were conducted to extend the research by Critchfield (1993) in which bias in verbal self-reports about DMTS success was studied across several levels of actual success, apparently showing self-report bias to be a function of signal probability. Experiment 1 examined whether apparent bias patterns in the Critchfield (1993) study were an artifact of a point contingency intended to promote accuracy in self-reports. Experiment 2 looked at the effects on self-report bias of mixing different types of DMTS trials, and thus different rates of DMTS success, within sessions (the Critchfield, 1993, study manipulated success rates only across conditions, whereas previous studies manipulated success within sessions).

EXPERIMENT 1

In Critchfield's (1993) study, in which subjects apparently exhibited signal probability effects in verbal self-reports, the referent performance (DMTS) was instated and maintained via a contingency of point reinforcement. Points also depended on self-reports: On a random-ratio (RR) 3 schedule, accurate self-reports produced point gains, and inaccurate self-reports produced point losses. This type of contingency typically does not bias reports in psychophysical procedures because it is symmetrical—that is, the conditional probability of point consequences is the same for reports of the presence and absence of the stimulus (Gescheider, 1985; McCarthy & Davison, 1979). Nevertheless, because the point contingency was not manipulated, it could not be ruled out as a source of self-report bias. The present experiment examined effects of the presence versus absence of report-contingent point gains and losses. In Experiment 1A, self-reports produced either (a) point gains when accurate and point losses when inaccurate or (b) point gains when accurate and no consequences when inaccurate. In Experiment 1B, self-reports produced either (a) point gains when accurate and point losses when inaccurate or (b) no point consequences. In both cases, comparisons were made across a range of DMTS success rates to determine whether signal probability effects could be attributed to the point contingencies.

METHOD

Subjects

Two female (Subjects S11 and S14) and 2 male (Subjects S12 and S13) undergraduate students volunteered to participate in a laboratory experiment on "Human Performance and Decision Making." Subjects received bonus credit in psychology classes based on their hours of participation; during sessions they accumulated points that served as chances in a drawing for cash prizes.

Apparatus

Subjects worked alone in a small room containing a table, chair, and a response console with a monochrome video monitor resting on it (for details, see Critchfield & Perone, 1990b). Subjects performed the DMTS task using four round illuminable response keys arranged horizontally near the bottom of the console's sloping front panel. Self-reports were made using two push buttons, each mounted to a small box extending from each side of the console. The video screen was divided into an upper box, used in conjunction with the DMTS task, and a lower box, used in conjunction with the self-report portion of the trial. A microcomputer outside the workroom controlled experimental events and collected the data.

Procedure

Trial format. During the main experiment each trial consisted of one DMTS response followed immediately, when scheduled, by a self-report, feedback about the success of the DMTS response, and consequences contingent on the self-report. Throughout the trial, error messages discouraged responses not conforming to the experimental protocol (see Critchfield & Perone, 1990b).

Subjects initiated each trial at the end of a 1-s intertrial interval (ITI) by simultaneously depressing the four round buttons on the front of the console, producing a sample-stimulus display that remained for 800 ms in the center of the DMTS box on screen. Following a 1-s delay, comparison stimuli appeared in locations corresponding to at least two of the depressed buttons. One comparison stimulus matched one sample element, and the others were randomly generated. Subjects attempted to select the matching comparison stimulus by

releasing the round button corresponding to it (see DMTS Stimuli section below). A successful response was recorded if a correct choice occurred within a time limit, normally 800 ms after presentation of the comparison stimuli. No stimulus change indicated when the time limit had elapsed.

Immediately after the choice, the DMTS box on the screen cleared and the center of the self-report box displayed the query, "Did you score?" (the word "score" had been used during preliminary training to signal point delivery). Below it, the labels "←YES" and "NO→" appeared 1 cm from the right and left sides of the self-report box, respectively. Pressing the button attached to the console's left side registered a "yes" report, and pressing the button attached to the console's right side registered a "no" report. Pressing either of these side buttons cleared the screen and advanced the trial to the next scheduled event. Other buttons were ineffective.

When scheduled, feedback messages describing the success of the DMTS response immediately followed the self-report. Three messages appeared simultaneously for 1 s in the DMTS area of the screen. The first message stated, "Your choice was CORRECT [or WRONG]." The second message stated, "Your choice was FAST ENOUGH [or TOO SLOW]." The third message summarized the implications of the other messages for point reinforcement, stating either "YOU SCORED! x points added to your total," or "NO SCORE" ($x = 1$ or 2, depending on the session; see below). When no DMTS feedback was scheduled, the trial advanced immediately to the next event.

When scheduled, a 1-s message describing the accuracy and point consequences of the self-report occurred next. In the self-report area of the screen a message stated, "RESULTS OF YOUR REPORT," accompanied by either "Correct— x point bonus" or "Wrong— x point penalty" as appropriate to the preceding self-report ($x =$ either 1 or 3 points, depending on the session; see below under Session and Condition Format). When no feedback about self-reports was scheduled, the trial advanced immediately to the ITI.

DMTS stimuli. Each sample and comparison stimulus consisted of a six-by-three matrix of rectangular cells, of which as few as three

or as many as 18 could be illuminated (similar stimuli were described by Baron & Menich, 1985). An element could be as large as 10 mm by 7 mm, depending on how many cells were illuminated. On each trial, stimuli were drawn randomly from a pool of several thousand unique shapes, without replacement except that one sample stimulus always matched one comparison stimulus.

Across conditions, the number of sample stimuli displayed on each trial could vary from one to four, and the number of comparison stimuli from which subjects chose could vary from two to four. Each sample and comparison stimulus was displayed in one of four possible locations, arranged horizontally about 2 cm apart, within the DMTS box on screen. If the number of sample or comparison stimuli was less than four, unused locations were left blank. On such occasions, the locations actually used were randomly determined on each trial. The entire stimulus array appeared centered within the DMTS box.

During comparison-stimulus display, each of the four stimulus locations corresponded to one of the round illuminated buttons being depressed on the console. Subjects indicated their choice of a comparison stimulus by releasing the round button corresponding to the location of the matching stimulus. If the button released corresponded to an unused stimulus location, the trial was canceled, the screen cleared, and a 4-s message staged, "Illegal action! You cannot choose a blank." The trial then restarted using new stimuli.

Session and condition format. During each 2-hr visit to the laboratory, subjects completed eight sessions, each lasting 100 trials (8 to 12 min). At the end of each session, a message on the subject's screen displayed the number of points accumulated during that session. The message included an overall session total and subtotals reflecting the number of points (out of 100) earned from DMTS and the number of points accumulated from self-reports. The self-report total was further broken down into total point gains and total point losses.

Each experimental condition, defined according to the number of sample and comparison stimuli present on each trial, lasted eight sessions. Sessions 1 through 3 consisted solely of DMTS trials without self-reports; each DMTS response was followed by the outcome feedback described previously. Successful

DMTS responses (those that were both correct and faster than the time limit) earned 2 points. Session 4 was intended to enhance the correspondence between self-reports and DMTS outcomes. Each DMTS choice was followed by a self-report and then feedback messages describing, in sequence, the success of the DMTS response and the consequences of the self-report. Successful DMTS responses earned 1 point. Accurate self-reports earned 1 point, and inaccurate ones resulted in a 1-point deduction from the subject's total. Sessions 5 through 8 provided the main data for the experiment and differed from the fourth session in that no feedback messages described DMTS performance after any trial. The point consequences for self-reports were manipulated across phases (described below).

Instructions. Immediately before the first session, subjects read printed instructions (identical to those used by Critchfield, 1993) describing the operation of the DMTS task and the point-reinforcement contingency requiring prompt, accurate DMTS responses. Each session began with messages on the computer screen describing the point contingencies operating in that session. For "matching decisions," the message stated the number of points earned per "score" and whether scores would be signaled on the screen. For "reports," the message stated the point value of each "bonus" and "penalty," and noted that point consequences for self-reports occurred only when indicated by feedback messages on the screen. Subjects cleared these messages and began the session by pressing a button located near the top of the console's front panel.

Preliminary training. Subjects first participated in an eight-session preliminary training phase (described by Critchfield, 1993), in which the stimuli were alphanumeric characters instead of the geometric shapes used in the main experiment; feedback messages lasted 2 s instead of 1 s; and the time limit on DMTS choices began at 3,000 ms and was gradually reduced to 800 ms across the first 200 trials. DMTS trials always consisted of two sample stimuli and three comparison stimuli.

Experimental manipulations. Stimulus conditions were defined, and named, according to the number of sample and comparison stimuli appearing on each trial. For example, when two sample stimuli were presented, one of which appeared among three comparison stim-

Table 1

Experiment 1: Raw delayed matching-to-sample and self-report data. See text for explanation of stimulus condition names.

Subject	Phase	Stimulus condition	Position in sequence of conditions	% successful DMTS responses	Trials	Self-reports			
						Hits	Misses	False alarms	Correct rejections
S11	gain-loss	14	3	88	400	348	5	6	41
		23	2 ^b , 5	73	400	280	12	21	87
		33	9	58	400	204	26	38	132
		44 ^a	6	30	400	63	57	16	264
	gain only	14	4	83	400	323	10	14	53
		23	1	55	400	207	13	35	145
		33	8	58	400	211	20	45	124
		44 ^a	7	38	400	82	70	4	184
S12	gain-loss	13	6	93	400	371	0	16	13
		22	7	81	400	321	4	43	32
		23	2	58	400	216	16	56	112
		43	3	44	400	142	34	100	124
	gain only	44 ^a	10	36	400	110	32	43	215
		13	5	92	400	367	2	11	20
		22	8	86	400	336	6	35	23
		23	1	51	400	178	24	57	141
		43	4	42	400	135	32	92	141
		44 ^a	9	34	400	93	41	50	216
S13	gain-loss	13	4	91	400	360	2	4	33
		22	8	79	400	312	5	52	31
		23	1	63	400	237	12	52	99
		43	5	28	400	81	31	59	228
	no consequences	13	3	91	400	363	1	12	24
		22	7	81	400	317	5	58	20
		23	2	45	400	173	9	37	180
		43	6	41	400	133	31	95	140
S14	gain-loss	13	5	96	400	383	2	3	18
		22	8	85	400	335	4	21	40
		23	2	60	400	222	16	21	141
		33	10	57	300 ^c	148	23	45	84
		43	3	47	300 ^c	101	39	35	125
	no consequences	13	6	95	400	378	1	12	9
		22	7	85	400	338	1	60	1
		23	1	56	400	196	28	63	113
		33	9	57	400	226	0	172	1
		43	4	48	400	145	46	97	112

^a Time limit on delayed matching-to-sample choices set at 600 ms.

^b Data from entire condition lost due to disk error. Condition repeated as soon as scheduling would allow.

^c Data from one session (100 trials) lost due to disk error.

uli, the condition was designated as "23," with the first digit describing the number of samples and the second the number of comparisons. Each subject participated in a different sequence of at least four stimulus conditions that were selected, based on individual performance and the results of Critchfield (1993), to produce a broad range of DMTS success rates. Table 1 lists the stimulus conditions used for each subject.

In Experiment 1A (Subjects S11 and S12), each stimulus condition occurred twice, once as part of the gain-loss phase and once as part of the gain-only phase. A coin flip determined which occurred first for each stimulus condition for each subject. The top portion of Table 2 summarizes the consequences produced by self-reports in these two phases. In the gain-loss phase, in the fourth through eighth sessions of each condition, self-reports produced

Table 2

Experiment 1: Summary of point consequences for self-reports. Plus signs denote a point gain, and minus signs denote a point loss. A zero indicates that no points and no feedback messages were contingent on self-reports. Delayed matching-to-sample choices, if accurate and faster than the time limit, earned 1 point on an FR 1 schedule. FR = fixed ratio; RR = random ratio.

			Status of self-report	
	Session	Schedule	Accurate	Inaccurate
Experiment 1A				
gain-loss phase	4	FR1	+1	-1
	5-8	RR3	+3	-3
gain-only phase	4	FR1	+1	0
	5-8	RR3	+3	0
Experiment 1B				
gain-loss phase	4	FR1	+1	-1
	5-8	RR3	+3	-3
no-consequence phase	4	FR1	0	0
	5-8	RR3	0	0

a point gain if accurate, and a point penalty subtracted from the subject's total if inaccurate. In Session 4, each accurate self-report produced a gain of 1 point, and each inaccurate self-report produced a loss of 1 point. In Sessions 5 through 8, correct self-reports produced a 3-point gain, and inaccurate self-reports produced a 3-point loss, both on an RR 3 schedule (due to characteristics of the computer's random-number generator, obtained probabilities of consequence delivery were slightly higher than programmed, condition median = .38, range, .29 to .39). In the gain-only phase, point gains occurred contingent on accurate self-reports exactly as in the gain-loss phase, but no point losses (and no feedback messages) were contingent on inaccurate self-reports.

In Experiment 1B (Subjects S13 and S14), each stimulus condition occurred twice, once as part of the gain-loss phase and once as part of the no-consequences phase. A coin flip determined which occurred first for each stimulus condition for each subject. The bottom portion of Table 2 summarizes the consequences produced by self-reports. The gain-loss phase was identical to that occurring in Experiment 1A. In the no-consequences phase, self-reports produced no point losses or gains and no feedback messages.

RESULTS AND DISCUSSION

Table 1 shows the percentage of successful DMTS responses (those that were both correct and faster than the time limit) in each of the stimulus conditions in both phases. For each subject, the stimulus conditions produced a broad range of DMTS success rates, and replications of stimulus conditions across phases tended to produce similar DMTS success rates.

Figure 2 shows rates of two types of self-report errors—false alarms (inaccurate self-reports of DMTS success) and misses (inaccurate self-reports of DMTS failure)—plotted as a function of DMTS success rate in each stimulus condition. Table 1 shows the raw self-report data from which these rates were calculated. In general, miss rates tended to increase, and false alarm rates tended to decrease, as DMTS success became less frequent. The two left panels (S11 and S12) compare self-reports in the gain-loss phase (open circles) with those in the gain-only phase (filled circles). There was no systematic difference in self-report error rates as a function of phase for these 2 subjects. The two right panels (S13 and S14) compare self-reports in the gain-loss phase (open circles) with those in the no-consequences phase (filled triangles). For both subjects, there was a tendency for false alarm rates to be higher in the no-consequence phase. For S14, miss rates also were slightly higher in the no-consequences phase.

Figure 3 shows the bias ($\log b$, bottom panels) and discriminability ($\log d$, top panels) of self-reports (McCarthy & Davison, 1981b, Equations 7 and 6, respectively), calculated from the data in Table 1 and plotted as a function of DMTS success rate in each stimulus condition. Here, positive $\log b$ scores represent a bias for reporting DMTS failure, whereas negative scores represent a bias for reporting success, with a score of zero (horizontal line in the bottom row of panels in Figure 3) representing no self-report bias. For all subjects, bias for reporting DMTS success was most pronounced when DMTS success was frequent and became less extreme as DMTS success became less frequent. Only 1 subject (S11) exhibited a clear bias for reporting DMTS failure coincident with extremely low DMTS success rates.

Higher $\log d$ scores represent better discriminability. For all subjects, discriminability tended to decrease with reductions in DMTS

Fig. 2. Experiment 1: Rates of self-report errors, false alarms, and misses, as a function of DMTS success.

success. (Because the manipulations designed to alter DMTS success involved simultaneous changes in both the number of sample stimuli and the number of comparison stimuli, this

outcome cannot be compared with the finding by Critchfield, 1993, that discriminability was negatively correlated with the number of DMTS sample stimuli and positively corre-

Fig. 3. Experiment 1: Self-report discriminability and bias as a function of DMTS success. In the bottom row of panels a horizontal line indicates zero bias. Data could be plotted only for those conditions producing at least one self-report in each of the categories in Figure 1 (see Table 1).

lated with the number of DMTS comparison stimuli.)

There were no systematic differences in bias or discriminability across phases for S11 or S12 (Experiment 1A). Thus, subtracting points contingent on accurate self-reports had no consistent effect on the signal-detection properties of self-reports. Phase-related differences were apparent, however, for S13 and S14 (Experiment 1B). In general, discriminability scores were higher in the phase in which both point gains and point losses were contingent on self-report accuracy than in the phase in which there was no contingency on self-report accuracy. Bias scores were less pronounced in the point-contingency phase for S14, but not consistently so for S13.

Comparing the results of Experiment 1A with those of Experiment 1B, it appears that the point-gain portion of the self-report contingency, but not the point-loss portion, influenced characteristics of self-reports. The effect appeared as a change in the height, but not the general shape, of functions relating self-report characteristics to DMTS success. Of primary importance to the present discussion is the fact that bias scores covaried with DMTS success rates in a fashion consistent with signal probability effects, just as in the previous study, regardless of the presence or absence of the point contingencies. Thus, the present results systematically replicate those of Critchfield (1993) and suggest that the apparent signal probability effects observed in that experiment were not an artifact of the point contingencies on self-reporting.

McCarthy and Davison (1979, 1981b) proposed that apparent signal probability effects in animal psychophysical experiments are an artifact of covariation between signal probability and the rate of reinforcement for "reporting" the presence of the stimulus. Bias remains relatively unchanged during signal probability manipulations when reinforcement rate is held constant (McCarthy & Davison, 1979, Procedure 2). In the gain-loss and gain-only phases of the present experiment, the confounding element identified by McCarthy and Davison was apparent in a strong correlation ($r \geq .98$ for all subjects) between signal probability (DMTS success rate) and rate of point gains for "I succeeded" self-reports (shown in the Appendix). This association cannot, however, directly explain the bias

functions of the present experiment, because bias varied with signal probability even when no points were contingent on self-reports (no-consequences phase).

Two procedural details merit investigation to determine whether the findings of Experiment 1A indeed diverge from those of McCarthy and Davison (1979). Self-report bias in the no-consequences phase could be a function of either (a) feedback and points delivered contingent on self-reports during the preliminary training sessions of each stimulus condition or (b) intermittent feedback and points delivered contingent on DMTS performance during the self-report sessions of each stimulus condition. In the former case, bias would reflect reinforcement history, whereas in the latter case, it might reflect implicit instructional control by the DMTS feedback (e.g., see Craig, 1980). The present procedures can be readily modified to eliminate these factors.

The contingencies on self-report accuracy employed here and in two previous studies (Critchfield, 1993; Critchfield & Perone, 1993) are symmetrical in the sense that conditional probabilities of point consequences are identical for "I succeeded" and "I failed" self-reports. Symmetrical contingencies normally do not have profound effects on bias in psychophysical studies (e.g., Gescheider, 1985), as might be predicted based on various permutations of the matching law (e.g., McCarthy & Davison, 1979; Nevin, Jenkins, Whitaker, & Yarensky, 1982). By contrast, asymmetrical contingencies—in the present context, arranging differential probabilities of consequences for "I succeeded" and "I failed" self-reports—can be a potent source of bias in psychophysical judgments (e.g., Davison & McCarthy, 1980; Goldiamond, 1964; Hobson, 1978; Hume, 1974b; McCarthy & Davison, 1981b; Nevin, 1970; Nevin et al., 1982; Nevin, Olson, Mandell, & Yarensky, 1975). In his conceptual analysis of verbal behavior, Skinner (1957) described self-reports as possessing properties of the tact, a verbal operant defined partly as under discriminative control of an object or event or property of an object or event. Although not explicitly invoking a psychophysical analogy, Skinner proposed that asymmetrical contingencies are likely to disrupt the normal stimulus control of tacts in general and of self-reports specifically (e.g., Skinner, 1957, pp. 149–150). Asymmetrical contingencies thus

are likely to constitute a useful focus for future research on the signal-detection properties of verbal self-reports.

EXPERIMENT 2

The relationship between self-report bias and DMTS success rates found in Critchfield's (1993) experiment and in Experiment 1 above was more consistent than that found in an earlier study by Critchfield and Perone (1993). Procedural differences across studies might be implicated, including whether DMTS success was manipulated within sessions (Critchfield & Perone, 1993) or across different experimental conditions (Critchfield, 1993). In other contexts, stimulus control has been found to depend partly on how frequently stimulus characteristics are manipulated. For example, Edhouse and White (1988) found that the discriminability of pigeons' memory for visual stimuli was higher when sample stimuli on successive trials were identical than when the stimuli were different. Similarly, Jones and White (1992), in comparing their data with other published studies (Harnett, McCarthy, & Davison, 1984; McCarthy & Davison, 1991), proposed that the relationship of reinforcement sensitivity to retention interval in DMTS depends on whether retention interval is manipulated within or between sessions.

The purpose of Experiment 2 was to compare the effects on self-report bias of two methods of manipulating DMTS difficulty. For each subject, self-reports about DMTS success were assessed twice across a range of DMTS success rates—once with DMTS trial types intermingled within each session, and once with DMTS trial types occurring only in different experimental conditions.

METHOD

Subjects and Apparatus

One male (Subject S15) and 2 female (Subjects S16 and S17) undergraduate students participated. Details of recruitment, compensation, and apparatus were identical to those in Experiment 1.

Procedure

Trial format, stimuli, instructions, and preliminary training were identical to those used in Experiment 1. During each 2-hr visit to the

Table 3

Experiment 2: Sessions in which the two experimental phases were completed by each subject. See text of Experiment 1 for explanation of stimulus condition names listed under the probe phase.

Phase	Subject		
	S15	S16	S17
Mixed			
DMTS-only	1-20	1-22	1-23
self-report training	21-29	23-31	24-32
self-report test	30-58	47-75	39-54, 64-73
Probe			
13	68-70	41-43	33-35
22	59-61	35-37	
23			55-57
24		44-46	
33	65-67	32-34	
34			36-38
42			61-63
43	62-64		
44	71-73	38-40	58-60

laboratory, subjects completed eight sessions, each lasting 99 trials (8 to 12 min). At the end of each session, a message on the subject's screen displayed the number of points accumulated during that session, as in Experiment 1. Each subject completed two phases, one in which trial types were mixed within a session, and one in which different trial types occurred in different sessions. Table 3 shows the sequence in which each subject completed the phases.

In the mixed phase, nine different stimulus conditions (22, 23, 24, 32, 33, 34, 42, 43, 44) alternated unpredictably across trials within each 99-trial session. Each occurred with a probability of .11, producing an average of about 11 trials of each stimulus condition per session. The first portion of this phase (DMTS-only training), designed to provide experience with the DMTS task, consisted of DMTS trials with no self-reports. Outcome feedback followed each DMTS response. DMTS-only sessions continued until the difference in mean percentage of successful DMTS trials for adjacent four-session blocks (summed across trial types) was less than 1% of the eight-session grand mean.

The second portion of this phase (self-report training) was designed to enhance the correspondence between self-reports and DMTS outcomes. Subjects completed eight sessions including both DMTS choices and self-reports.

Table 4

Experiment 2: Raw delayed matching-to-sample and self-report data for the test portion of the mixed-stimulus phase. See text of Experiment 1 for explanation of stimulus condition names.

Subject	Phase	Stimulus condition	% success-ful DMTS responses	Trials	Self-reports			
					Hits	Misses	False alarms	Correct rejections
S15	mixed	22	79	319	248	3	44	24
		23	59	319	213	6	43	57
		24	71	319	184	4	37	94
		32 ^a	72	638	446	15	122	55
		33	54	319	162	11	74	72
		42 ^a	67	638	422	7	150	49
	probe	43	48	319	147	7	105	60
		13	99	198	196	1	1	0
		22	83	198	160	4	23	11
		33	56	181 ^b	92	10	32	45
		42	63	198	120	4	58	16
		43	57	198	104	8	39	47
		44	39	198	64	13	28	93
S16	mixed	22	75	330	228	20	28	54
		23	61	311	174	16	25	96
		24	50	312	154	3	42	113
		32	65	312	185	19	42	66
		33	49	311	128	24	30	129
		34	34	312	95	10	31	176
		42	60	311	158	28	43	82
		43	38	311	96	23	31	162
	probe	44	30	310	75	19	27	189
		13	93	198	184	1	4	9
		22	78	198	143	11	9	35
		24	58	198	102	12	14	70
		33	53	198	76	28	13	81
		44	31	198	41	21	23	113
S17	mixed ^c	22	84	280	230	5	32	13
		23	64	281	168	12	45	56
		24	48	281	130	6	62	83
		32	71	281	190	9	67	15
		33	53	281	144	6	60	71
		34	38	281	98	10	70	103
		42	66	281	174	10	74	23
		43	48	280	114	19	92	55
	probe	44	38	280	99	8	88	85
		13	90	198	177	1	8	12
		23	64	198	124	2	18	54
		34	41	198	70	11	69	48
		42	63	198	120	4	58	16
		44	40	198	58	21	34	85

^a Due to a programming error, Type 34 trials were run as Type 32 trials, and Type 44 trials were run as Type 42 trials. Thus, for this subject there were no Type 34 or Type 44 trials, and the number of Type 32 and 42 trials was double that intended.

^b Subject left session prematurely, causing a loss of 17 trials.

^c Data from half of one session in the mixed phase (48 trials) were lost due to disk failure.

Each DMTS choice was followed in sequence by a self-report, feedback about the success of the DMTS choice, and feedback about the accuracy of the self-report. Successful DMTS choices earned 1 point. Each accurate self-report produced a 1-point gain, and each inaccurate self-report produced a 1-point loss.

The final portion of the mixed phase (self-report test) provided the data for this phase. Subjects completed 26 (S16) or 29 (S15 and S17) sessions of DMTS choices with self-reports. No feedback messages described DMTS success, and self-report consequences occurred on an RR 3 schedule.

Fig. 4. Experiment 2: Rates of self-report errors, false alarms, and misses, as a function of DMTS success.

Probe phase. Each subject completed at least four three-session probes designed to measure performance under a single stimulus condition occurring in isolation. The first session in each probe consisted of DMTS trials, with no self-reports, each followed by feedback describing the success of the DMTS choice (as in the first portion of the mixed phase). The remaining two sessions provided the data, and consisted of DMTS trials, followed by self-reports, with no DMTS outcome feedback. Self-reports produced point consequences, as in the final portion of the mixed phase. Table 3 lists the stimulus conditions completed by each subject in the probe phase and the sequence in which they were completed. These conditions were selected, based on data from Critchfield (1993), to produce a broad range of DMTS success rates. Because of time constraints on subject participation, no attempt was made to achieve a one-to-one correspondence between conditions in the mixed and probe phases. Rather, the intent was to determine whether bias scores from the two phases would fall along different functions.

RESULTS AND DISCUSSION

Table 4 shows the percentage of successful DMTS responses (those that were both correct and faster than the time limit) in each of the stimulus conditions in both phases. The stimulus conditions produced a broad range of DMTS success rates for each subject. Figure 4 shows rates of false alarms and misses plotted as a function of DMTS success rate in each stimulus condition. Table 4 shows the raw self-report data from which these rates were cal-

culated. In general, miss rates tended to increase, and false alarm rates tended to decrease, as DMTS success became less frequent. The top row of panels shows a slight tendency for miss rates to be higher in the probe phase, when each DMTS stimulus condition occurred in isolation. There was no systematic difference in false alarm rates as a function of phase.

Figure 5 shows the bias ($\log b$) and discriminability ($\log d$) of self-reports, calculated from the data in Table 4 and plotted as a function of DMTS success rate in each stimulus condition. The bottom row of panels shows that, in general, bias for reporting DMTS success was most pronounced when DMTS success was frequent and became less extreme as DMTS success became less frequent. Visual inspection suggests that bias was not profoundly different across phases, although the slopes of functions determined by least squares linear regression were slightly steeper for the probe phase, in which trial types varied across conditions ($-.031$, $-.018$, and $-.017$ for S15, S16, and S17, respectively), than for the mixed phase, in which trial types varied within sessions ($-.013$, $-.011$, and $-.012$, respectively). Only for S16, however, were bias scores consistently different, at all levels of DMTS success, in the two phases. In this sense, the present results may be inconsistent with those of Craig (1981), who found more pronounced bias in human vigilance performance when different target stimuli were mixed within sessions than when they occurred in separate conditions.

The top row of panels in Figure 5 shows

Fig. 5. Experiment 2: Self-report discriminability and bias as a function of DMTS success. In the bottom row of panels a horizontal line indicates zero bias. Data could be plotted only for those conditions producing at least one self-report in each of the categories in Figure 1 (see Table 4).

that the relationship of self-report discriminability to DMTS success rate tended to depend on the phase. For S16 and S17, discriminability was positively related to DMTS success rate in the probe phase, when trial types were manipulated across conditions, but not in the mixed phase, when trial types were manipulated within the session. For all 3 subjects, discriminability scores were more variable in the mixed phase. Comparison of these findings with those of previous signal-detection investigations is not a simple matter. Some studies suggest that discriminability is impaired when signal type is varied within sessions (e.g., Craig, 1979, 1981; Megaw, Alexander, & Richardson, 1979), but others do not (e.g., Fox & Haslegrave, 1969; Sheehan & Drury, 1971). Consistent with the data of S16 and S17, however, there is evidence that within-session manipulation of trial types disrupts the correlation between signal probability and discriminability (Craig, 1979, 1981).

The results of Experiment 2 replicate those of previous studies showing self-report bias patterns consistent with a signal probability interpretation. In addition, they may help to explain why bias scores in the study by Critchfield (1993) were more closely related to DMTS success rates than in previous studies (e.g., Critchfield & Perone, 1993). In the present study, bias scores (normally negative, indicating a bias for reporting DMTS success)

were somewhat less resistant to change in actual DMTS success rates when different trial types occurred in different probe conditions than when different trial types occurred mixed within a session. The Critchfield (1993) study manipulated trial difficulty across conditions, as in the present probe phase, whereas previous studies manipulated trial difficulty within sessions, as in the mixed phase (Critchfield & Perone, 1990a, 1990b, 1993).

GENERAL DISCUSSION

The results of the present research are consistent with a signal probability interpretation of self-report bias. Signal probability effects occurred independently of whether self-reports produced point consequences or whether signal probability (DMTS success rate) was manipulated within or between sessions. These findings bolster the analogy drawn by Critchfield (1993) between verbal self-reports and psychophysical judgments, and support the conclusion that signal probability effects can occur in self-reports much as they do in reports about external stimuli in psychophysical experiments.

The analogy between self-reports and psychophysical judgments has both theoretical and methodological implications. In the former case, data showing psychophysical effects in self-evaluation appear to run counter to proposals

that special principles apply when individuals observe, remember, and describe their own actions or characteristics (e.g., Ericsson & Simon, 1984; Keenan & Baillet, 1980; Nisbett & Wilson, 1977). The data suggest instead that self-observation shares properties with other forms of observation. Although behavior analysts typically endorse this point of view (Skinner, 1957), they have been slow to develop methods for investigating the relevant "self-evaluative" phenomena, especially verbal ones. In the latter case, therefore, it seems likely that studies based on psychophysical procedures can provide methodological underpinnings for operant investigations of verbal behavior. There is ample precedent for the use of signal-detection indices to improve the measurement of correspondence between self-reports and their presumed referents (e.g., Critchfield & Perone, 1993; Gardner, Martinez, & Espinoza, 1987; Hosseini & Ferrell, 1982), but a psychophysical approach can also suggest controlling variables to explore in the experimental analysis of self-awareness, such as signal probability effects and "asymmetrical" reinforcement contingencies (Davison & McCarthy, 1980; McCarthy & Davison, 1979, 1981a, 1981b; Nevin, 1970; Nevin et al., 1975, 1982).

Research that can address the complex relations that operant theorists believe comprise self-reports and other verbal behavior (e.g., Skinner, 1957) may also provide information of value to scientists whose explanations of behavior focus on the correspondence between acts of self-evaluation and actual performance, as in self-regulation theories (e.g., Bandura, 1986; Carver & Sheier, 1981; Duval & Wicklund, 1972). For example, some theories of self-regulation view degraded (e.g., biased) self-evaluation as necessarily either a precipitator or a symptom of pathology (Bandura, 1986), whereas others, paradoxically, assume "self-enhancing" biases to be integral to good mental health (Bjorkland & Green, 1992; Sackeim, 1983; Taylor & Brown, 1988). In extreme versions, neither of these perspectives seems to be well equipped to incorporate factors (such as signal probability effects) that are capable of shifting bias momentarily from "self-enhancing" to "self-effacing" and thus are likely to be present in the self-evaluation of normal individuals. Moreover, it is unclear whether these views can accommodate possible cultural differences in self-evaluation biases (Fahr, Dobbins, & Cheng, 1991; Stevenson, Chen, &

Lee, 1993). It seems likely that any viable theory of behavioral self-regulation would have to take into account the results of research on the situational determinants of self-evaluation, as would occur within a psychophysical framework.

Despite its possible heuristic value, the analogy between psychophysics and self-reporting can be expected to break down under close scrutiny on at least two levels. One limitation is essentially an occupational hazard of conducting basic research with human subjects: The behavior-environment relations harnessed when humans' self-reports are measured are likely to overlap with existing response classes (Baron, Perone, & Galizio, 1991; Branch, 1991), meaning that predispositions not typically encountered in basic psychophysical studies can be expected. The tendency toward "self-enhancing" biases observed in the present line of research (Critchfield, 1993; Critchfield & Perone, 1993) may provide one example, although it should be noted that signal-detection analyses have allowed this predisposition to be characterized in the first place.

A second and more fundamental problem in pursuing a true psychophysical approach to the experimental analysis of self-awareness lies in the definition of the referent event. A response that serves as referent, although superficially a unitary phenomenon from the experimenter's perspective, may have multiple stimulus features from the perspective of the self-observer (e.g., see Critchfield & Perone, 1993). Given this ambiguity about the critical features of the referent response, it may prove difficult to manipulate parametrically the physical characteristics of the behavioral "stimulus" as precisely as psychophysical researchers manipulate the physical characteristics of external stimuli such as lights or tones. Nevertheless, with operant research on verbal behavior in relative infancy, much remains to be learned about the psychophysical properties of verbal self-reports before research is substantially constrained by this concern.

REFERENCES

- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Baron, A., & Menich, S. R. (1985). Reaction times of older and younger men: Effects of compound samples and a prechoice signal on delayed matching-to-sample

- performances. *Journal of the Experimental Analysis of Behavior*, 44, 1-14.
- Baron, A., Perone, M., & Galizio, M. (1991). Analyzing the reinforcement process at the human level: Can application and behavioristic interpretation replace laboratory research? *The Behavior Analyst*, 14, 95-106.
- Bjorkland, D. F., & Green, B. L. (1992). The adaptive nature of cognitive immaturity. *The American Psychologist*, 47, 46-54.
- Boring, E. G. (1953). A history of introspection. *Psychological Bulletin*, 50, 169-189.
- Branch, M. N. (1991). On the difficulty of studying "basic" behavioral processes in humans. *The Behavior Analyst*, 14, 107-110.
- Carver, C. S., & Scheier, M. F. (1981). *Attention and self-regulation: A control-theory approach to human behavior*. New York: Springer-Verlag.
- Craig, A. (1978). Is the vigilance decrement simply a response adjustment towards probability matching? *Human Factors*, 20, 441-446.
- Craig, A. (1979). Vigilance for two kinds of signal with unequal probabilities of occurrence. *Human Factors*, 21, 647-653.
- Craig, A. (1980). Effect of prior knowledge of signal probabilities on vigilance performance at a two-signal task. *Human Factors*, 22, 361-371.
- Craig, A. (1981). Monitoring for one kind of signal in the presence of another: The effects of signal mix on detectability. *Human Factors*, 23, 191-197.
- Critchfield, T. S. (1993). Signal-detection properties of verbal self-reports. *Journal of the Experimental Analysis of Behavior*, 60, 495-514.
- Critchfield, T. S., & Perone, M. (1990a). Verbal self-reports as a function of speed, accuracy and reinforcement of the reported performance. *Psychological Record*, 40, 54-554.
- Critchfield, T. S., & Perone, M. (1990b). Verbal self-reports of delayed matching to sample by humans. *Journal of the Experimental Analysis of Behavior*, 53, 321-344.
- Critchfield, T. S., & Perone, M. (1993). Verbal self-reports about matching-to-sample: Effects of the number of elements in a compound sample stimulus. *Journal of the Experimental Analysis of Behavior*, 59, 193-214.
- Danziger, K. (1980). The history of introspection reconsidered. *Journal of the History of the Behavioral Sciences*, 16, 241-262.
- Davison, M., & McCarthy, D. (1980). Reinforcement for errors in a signal-detection procedure. *Journal of the Experimental Analysis of Behavior*, 34, 35-47.
- Duval, S., & Wicklund, R. A. (1972). *A theory of objective self-awareness*. New York: Academic Press.
- Edhouse, W. V., & White, K. G. (1988). Sources of proactive interference in animal memory. *Journal of Experimental Psychology: Animal Behavior Processes*, 14, 56-70.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Fahr, J., Dobbins, G. H., & Cheng, B. (1991). Cultural relativity in action: A comparison of self-ratings made by Chinese and U.S. workers. *Personnel Psychology*, 44, 129-147.
- Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row, Peterson.
- Flavell, J. H., & Wellman, H. M. (1977). Metamemory. In R. V. Kail & J. W. Hagen (Eds.), *Perspectives on the development of memory and cognition* (pp. 3-33). New York: Erlbaum.
- Fox, J. G., & Haslegrave, C. M. (1969). Industrial inspection efficiency and the probability of a defect occurring. *Ergonomics*, 12, 713-721.
- Gardner, R. M., Martinez, R., & Espinoza, T. (1987). Psychophysical measurement of body image of self and others in obese subjects. *Journal of Social Behavior and Personality*, 2, 205-217.
- Gescheider, G. A. (1985). *Psychophysics: Method, theory, and application*. Hillsdale, NJ: Erlbaum.
- Goldiamond, I. (1964). Response bias in perceptual communication. In D. M. Rioch & E. A. Weinstein (Eds.), *Disorders of communication* (pp. 334-363). Baltimore: Williams and Wilkins.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Harnett, P., McCarthy, D., & Davison, M. (1984). Delayed signal detection, differential reinforcement, and short-term memory in the pigeon. *Journal of the Experimental Analysis of Behavior*, 42, 87-111.
- Hobson, S. L. (1978). Discriminability of fixed-ratio schedules for pigeons: Effects of payoff values. *Journal of the Experimental Analysis of Behavior*, 30, 69-81.
- Hosseini, J., & Ferrell, W. R. (1982). Detectability of correctness: A measure of knowing that one knows. *Instructional Science*, 11, 113-127.
- Hume, A. L. (1974a). Auditory detection and optimal response biases. *Perception & Psychophysics*, 15, 425-433.
- Hume, A. L. (1974b). Optimal response biases and the slope of ROC curves as a function of signal intensity, signal probability, and relative payoff. *Perception & Psychophysics*, 16, 377-384.
- Jones, B. M., & White, K. G. (1992). Sample-stimulus discriminability and sensitivity to reinforcement in delayed matching to sample. *Journal of the Experimental Analysis of Behavior*, 58, 159-172.
- Keenan, J. M., & Baillet, S. D. (1980). Memory for personally and socially significant events. In R. S. Nickerson (Ed.), *Attention and performance VII* (pp. 651-669). Hillsdale, NJ: Erlbaum.
- McCarthy, D., & Davison, M. (1979). Signal probability, reinforcement, and signal detection. *Journal of the Experimental Analysis of Behavior*, 32, 373-386.
- McCarthy, D., & Davison, M. (1981a). Matching and signal detection. In M. L. Commons & J. A. Nevin (Eds.), *Quantitative analysis of behavior: Vol. 1. Discriminative properties of reinforcement schedules* (pp. 393-417). Cambridge, MA: Ballinger.
- McCarthy, D., & Davison, M. (1981b). Towards a behavioral theory of bias in signal detection. *Perception & Psychophysics*, 29, 371-382.
- McCarthy, D., & Davison, M. (1991). The interaction between stimulus and reinforcer control on remembering. *Journal of the Experimental Analysis of Behavior*, 56, 51-66.
- Megaw, E. D., Alexander, C. J., & Richardson, J. (1979). Fault mix and inspection performance. *International Journal of Production Research*, 17, 123-133.
- Nevin, J. A. (1970). On differential stimulation and differential reinforcement. In W. C. Stebbins (Ed.), *Animal psychophysics: The design and conduct of sensory experiments* (pp. 410-423). New York: Appleton-Century-Crofts.
- Nevin, J. A., Jenkins, P., Whittaker, S., & Yarensky, P.

- (1982). Reinforcement contingencies and signal detection. *Journal of the Experimental Analysis of Behavior*, 37, 65-79.
- Nevin, J. A., Olson, K., Mandell, C., & Yarensky, P. (1975). Differential reinforcement and signal detection. *Journal of the Experimental Analysis of Behavior*, 24, 355-367.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Oah, S., & Dickinson, A. M. (1989). A review of empirical studies of verbal behavior. *Analysis of Verbal Behavior*, 7, 53-68.
- Sackeim, H. A. (1983). Self-deception, self-esteem, and depression: The adaptive value of lying to oneself. In J. Masling (Ed.), *Empirical studies of psychoanalytic theories* (Vol. 1, pp. 101-157). Hillsdale, NJ: Analytic Press.
- Schusterman, R. J., & Johnson, B. W. (1975). Signal probability and response bias in California sea lions. *Psychological Record*, 25, 39-45.
- Sheehan, J. J., & Drury, C. G. (1971). The analysis of industrial inspection. *Applied Ergonomics*, 2, 74-78.
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Stevenson, H. W., Chen, C., & Lee, S. (1993). Mathematics achievement of Chinese, Japanese, and American children: Ten years later. *Science*, 259, 53-58.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin*, 103, 193-210.

Received July 6, 1993

Final acceptance April 29, 1994

APPENDIX

Experiment 1: Point consequences delivered per minute contingent on self-reports.

Subject	Phase	Condition	Rates of consequences for self-reports			
			“I succeeded” self-reports		“I failed” self-reports	
			Point gain	Point loss	Point gain	Point loss
S11	gain-loss	14	4.61	0.07	0.37	0.00
		23	3.09	0.17	0.93	0.29
		33	2.43	0.61	1.41	0.33
		44	0.73	0.18	3.05	0.60
	gain only	14	3.90	0.00	0.73	0.00
		23	2.10	0.00	1.45	0.00
		33	2.46	0.00	1.94	0.00
		44	1.18	0.00	2.85	0.00
S12	gain-loss	13	4.24	0.09	0.18	0.06
		22	4.06	0.24	0.54	0.09
		23	2.36	0.40	1.20	0.16
		43	1.44	0.82	1.53	0.44
	gain only	44	1.11	0.55	2.71	0.28
		13	4.28	0.00	0.28	0.00
		22	3.79	0.00	0.42	0.00
		23	1.71	0.00	1.43	0.00
S13	gain-loss	43	1.53	0.00	1.53	0.00
		44	0.89	0.00	2.75	0.00
		13	3.70	0.03	0.40	0.00
		22	2.88	0.44	0.32	0.10
	S14	gain-loss	23	2.20	0.38	1.03
43			0.80	0.58	1.69	0.21
13, 22, 23, 43			0.00	0.00	0.00	0.00
13			4.23	0.03	0.33	0.12
no consequences		22	3.82	0.09	0.52	0.03
	23	2.51	0.20	1.48	0.14	
	33	1.72	0.46	0.89	0.29	
	43	1.51	0.44	1.84	0.37	
	13, 22, 23, 33, 43	0.00	0.00	0.00	0.00	