EXPERIMENTS IN BALL-TOSSING: THE SIGNIFICANCE OF LEARNING CURVES

BY JOSEPH PETERSON

University of Minnesota

Swift's curves of learning to keep "two balls going with one hand, catching and throwing one while the other is in the air," do not show the rapid initial decline characteristic of most motor learning curves. On the contrary they are in general "concave toward the vertical axis, which means, of course, that the progress was at first slow and then more rapid."¹ In Swift's curves vertical distances represent the average number of successful catches in ten trials, a day's practice; and the horizontal distances indicate the successive days, or the number of practice periods. By this method, as Thorndike has pointed out,² equal distances on the abscissa line do not mean equal amounts of practice. A rearrangement of the data of one of Swift's subjects so that the horizontal distances will stand for equal number of tosses results in a curve which up to 4,500 tosses may be roughly represented by a straight line.³

In animal and other experiments in which errors are significant, as is true also of ball-tossing, curves of errors have been given along, usually, with curves of other data. Just what data to take as the most significant in any learning process, is a question that is not settled, and one that must obviously be answered differently for different learning processes. In most cases, even when time, errors, and performance are plotted in separate curves, the data are liable to be misleading if taken to represent the entire learning that goes on. If, for instance, one of the curves shows a plateau

¹ Swift, E. J., 'Mind in the Making,' 1909, p. 177. The interpretation expressed in this quotation is erroneous, though it is very frequently expressed in psychological literature.

² Thorndike, E. L., 'The Psychology of Learning,' 1913, p. 121.

³ Thorndike, *op. cit.*, p. 122.

we are too apt to conclude that learning was abated or wholly suspended at this point. This tendency is especially noticeable in cases of learning represented by a single curve.

THE EXPERIMENT

The experiment here reported was devised to secure extensive data on ball-tossing in terms of errors, so that the results might be compared with those of other kinds of experimentation. Though other data—time and ‘the feeling of efficiency’ before practice—were obtained, the conditions of the experiment were not such as to make these data very reliable. They will therefore be considered only in a general way in the discussion.¹ The subjects in the experiments were sophomore, junior, and senior students taking their second half year’s course in psychology. The experiment was performed in the spring of 1916. Each subject received written instructions as follows:

“Get two solid rubber balls, each of about 42 mm. diameter and weighing approximately 123 grams. Select for practice each day a period as free from interruption as possible, preferably in the morning before breakfast (or two periods daily, one in the morning and one in the afternoon). The place of practice, as well as the time, should be constant. A room with high ceiling and light walls is best; varying conditions of the sky would interfere with outside practice. Arrange something to catch, and thus to prevent from rolling away, the balls you miss. Practice the act of keeping two balls going with one hand, catching and throwing the one while the other is in the air. (This was illustrated to all subjects by the writer.) Count only the catches, beginning with that of the first ball tossed and ending with that preceding the first miss. This is one series. Each practice period is to continue until the total number of successful catches reaches 200, when it must stop. The two-hundredth, if stopped voluntarily, is not a miss. Rest between series whenever tired, but only for one minute. Continue, practicing regularly every day, until for five days you have made no misses.

“Keep a careful record of the number of catches in each series, the number and place in the practice period of rests, cause of miscatches when easily noticeable, time of day, conditions of light (never artificial), your own condition (and whether you feel that you are going to make a good record), chief difficulties and how overcome, total time of each practice period, and any other observations. Follow the directions closely, noting any departure from them. Be constantly on the alert for factors or conditions enabling you to explain any irregularities.”

The instructions are slightly changed as published here, but only in unessentials, such as the rearrangement of sen-

¹ In general it may be said here that time curves seem to parallel rather closely error curves and that one’s ‘feeling of efficiency’ before practice is, on the whole, not a very reliable index of efficiency as measured by actual results.

tences, and the making more explicit of detailed matters explained orally to all subjects together. All who found it possible were urged to take two periods of practice daily, of 200 catches each, so that we might have data on the distribution of practice. In such cases one practice was to take place in the morning and one in the afternoon, never both in the same half day. There were, of course, slight deviations in the sizes of the balls used by different individuals and slight irregularities in other details of practice. Occasionally, for instance, a day's practice would be omitted because of illness or absence, some of the students were obliged to practice throughout the experiment in artificial light, and occasionally one would forget in one's enthusiasm and run over the 200 limit of catches. All papers showing irregularities or carelessness judged sufficient to interfere with the reliability of results were excluded before the data for the writing of this report were collected. That is to say, the papers used in the report were *selected* only on the basis of regularity and carefulness of work, not on that of the kind of results secured! In all individual records used the details of the experimental procedure and results were sufficient, in connection with individual conferences and reports during and after the experiments, to enable the writer, the director of the experiments, to judge of their reliability. No records are included of individuals who had had previous practice in the performances, even though the subject had apparently lost the effects of the practice.

In all, the records of twenty-six individuals are included in the reports, though the data from three of these records are excluded from the averaged curves because the practice in each case was discontinued before a high degree of skill was acquired. Two of these were taken from the once-a-day group and one from the twice-a-day group. For reasons already explained the records of eight subjects were excluded from any consideration in this paper. The total number of persons performing the experiment was therefore thirty-four.¹

¹ It should be stated also that the experimenter himself practiced at ball-tossing daily for three months in the spring of 1914. In the following year he obtained results

RESULTS

The individual results of these experiments are given in the following tables and figures, so that the reader may see the extent of individual variations and that he may judge for himself on important points mentioned in the following discussion. It is important, moreover, to preserve these individual records, which in each case represent probably from ten to fifty hours' experimentation, so that they may be available for other investigators who may be doing work of a similar kind. Mere averages, we must also keep in mind, not infrequently cover up significant aspects of individual results, such as plateaus, initial 'spurts' and changes near the physiological limits of learning. In results such as these, deviations from central tendencies are not particularly illuminating, and have therefore not been given. They may be calculated by any one who desires them. It is more important for our purpose to give the individual results and the averages.

Explanatory Note.—In Table I. and Table II. a horizontal line between any two numbers in a column indicates one practice period missed; if more periods than one were missed the exact number is indicated by an italic figure at the right of the line. The columns bracketed at the beginning and end are not included in the averages given in Table III.

Table I. gives the results of the fifteen subjects who had one practice period each day. Table II. is the corresponding record of the group of thirteen subjects who had a practice period in the morning and one in the afternoon, or two in twenty-four hours. The first column, P. P., gives the number of the practice period. For example, 20 in the P. P. column in Table I. is the twentieth day of the practice and also the twentieth practice period, whereas in Table II. it represents the afternoon practice of the tenth day. The other columns give the number of errors—miscatches—made by the respective individuals in the practice periods of 200 from fifteen subjects, which agree substantially with those here reported. These results are not incorporated in the data of this report for the reason that the same uniformity of conditions of experimentation was not secured. The method used in the present grew out of the earlier work.

TABLE I
INDIVIDUAL RECORDS OF ONCE-A-DAY GROUP. ERRORS IN 200 CATCHES

P. P.	Subjects														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	209	104	190	111	79	155	157	120	153	64	38	72	122	105	116
2	195	55	185	42	53	128	166	69	113	37	42	25	102	68	78
3	181	45	187	44	34	107	168	72	120	42	19	21	100	63	65
4	165	40	175	42	27	76	150	52	83	19	23	23	70	60	55
5	152	37	172	42	24	56	145	62	88	22	13	10	86	49	42
6	143	36	169	21	13	49	147	53	81	16	12	11	75	49	35
7	154	34	158	19	11	37	139	48	89	15	9	11	65	37	22
8	129	28	155	30	10	35	132	45	85	12	10	6	62	32	26
9	116	26	143	26	10	30	126	44	75	13	11	8	63	26	28
10	119	14	141	24	9	26	111	44	78	11	15	7	57	19	16
11	118	13	145	18	8	20	103	45	80	15	12	2	53	14	18
12	107	8	138	16	10	19	93	45	81	11	6	12	39	12	15
13	109	9	139	17	4	26	135	42	77	9	13	4	40	11	14
14	105	8	135	15	9	20	77	37	77	12	11	3	37	11	14
15	112	9	121	16	8	20	71	34	77	10	12	2	38	9	17
16	103	5	112	9	8	15	55	36	80	11	5	6	44	10	15
17	108	12	110	9	6	14	55	36	70	10	4	8	34	8	13
18	121	6	102	9	4	23	57	36	69	7	7	3	28	8	16
19	117	10	96	8	9	16	55	35	65	5	5	2	29	6	12
20	107	10	96	7	9	14	54	37	79	7	3	3	20	7	7
21	103	10	96	6	5	15	55	37	68	7	4	3	27	7	6
22	104	10	85	6	5	12	53	32	69	5	6	3	20	6	5
23	106	6	79	9	5	11	53	25	66	6	4	1	26	5	5
24	109	9	65	9	5	11	50	24	63	5	3	3	33	4	6
25	103	10	63	10	6	20	55	24	64	5	2	5	21	4	5
26	101	7	67	7	8	18	51	27	66	5	4	3	23	4	4
27	106	10	57	6	4	14	50	26	53	4	3	2	18	4	3
28	109	8	59	7	4	15	45	24	58	4	3	3	14	4	8
29	102	8	45	9	5	13	77	24	45	4	2	3	14	4	4
30	100	5	43	2	4	13	55	23	46	4	2	10	18	4	5
31	105	5	*	8	5	10	47	22	44	3	0	5	13	11	3
32	108	5		14	6	9	39	23	40	3	0	4	11	11	5
33	104	5		7	5	19	46	18	36	3	2	3	12	8	2
34	99	2		9	5	16	35	15	44	4	2	2	11	7	3

P. P.	Subjects														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
35	107	7		3	6	10	36	18	33	4	2	3	8	7	2
36	102	2		4	4	9	33	12	34	2	0	4	7	7	3
37	106	3		7	4	9	31	13	31	2	0	3	8	7	0
38	97	1		5	4	9	30	14	37	3	0	3	16	6	1
39	99	*		7	5	8	27	13	28	2	0	4	7	6	1
40	101			7	7	8	33	12	28	2	0	2	12	6	1
41	100			5	4	8	30	12	32	2	*	3	6	5	1
42	104			5	5	10	30	7	29	3		4	10	5	0
43	102			6	3	9	39	8	26	2		3	8	4	1
44	*			7	5	9	34	4	28	2		5	12	3	1
45				3	4	8	29	8	24	2		2	8	3	0
46				5	4	10	24	5	25	2		4	7	3	2
47				5	2	10	23	6	24	2		10	9	3	2
48				9	2	9	22	6	22	2		5	7	3	1
49				5	3	7	27	7	20	2		2	8	2	1
50				4	4	5	35	*	18	2		2	5	2	0
51				4	2	*	30		24	2		2	8	2	0
52				4	3		26		17	2		3	5	2	0
53				5	3		21		20	0		3	11	2	*
54				2	0		18		18	0		0	6	2	
55				2	1		12		2	2		2	9	2	
56				3	1		9		2	5		1	4	1	
57				3	1		11		2	0		2	4	1	
58				1	1		7		0	0		1	5	1	
59				1	1		4		0	0		2	8	1	
60				3	2		3		0	0		4	4	*	
61				1	2		*		0	0		4	3		
62				4	2				*	0		1	4		
63				1	3							2	5		
64				1	1							2	6		
65				3	*							2	4		

(D's record for successive days continues as follows: 3, 2, 2, 0, 1, 1, 0, *; L's 0, 2, 2, 2, 2, 1, *; M's 5, 4, 3, 6, 6, 5, 4, 8, 4, 2, 3, 0, 4, 5, *.)

successful catches each. Horizontal lines between any two numbers in a given column indicate an interruption at that point in the practice, due to one of several conditions, such as temporary illness, absence from the place of practice, etc. A small number at the right of one of these lines shows the number of periods omitted. When no number is given with the line only one period was missed. Thus the record of 'B' shows two periods missed between the ninth and the tenth, one period between the sixteenth and the seventeenth, and so on. This student has the poorest record so far as regu-

TABLE II
INDIVIDUAL RECORDS OF TWICE-A-DAY GROUP. ERRORS IN 200 CATCHES

P. P.	Subjects												
	P	Q	R	S	T	U	V	W	X	Y	Z	AA	BB
1	103	161	98	47	91	148	200	66	72	117	87	174	213
2	79	143	95	34	74	113	178	52	62	58	80	146	205
3	66	112	93	20	69	110	137	46	26	57	38	109	208
4	62	111	91	9	46	111	121	38	25	76	27	96	200
5	52	78	90	7	33	97	123	33	17	24	30	70	203
6	48	68	87	7	26	79	121	55	16	23	29	61	200
7	48	66	85	4	18	86	109	28	19	19	15	44	195
8	37	55	84	5	26	108	109	21	12	13	17	40	200
9	34	57	83	5	17	79	121	18	9	12	15	37	180
10	34	59	80	2	17	88	91	17	13	10	18	33	187
11	28	36	78	4	17	65	102	15	8	8	12	28	185
12	31	29	75	3	12	73	77	13	8	10	13	28	165
13	37	29	74	3	13	71	102	12	9	9	9	27	163
14	34	42	73	3	11	53	71	12	6	6	11	25	170
15	32	30	71	1	10	50	109	9	6	5	11	26	173
16	26	32	69	1	12	40	82	8	8	6	11	27	168
17	27	22	65	3	11	42	81	9	5	5	12	27	150
18	31	28	60	1	7	39	71	7	10	4	11	30	153
19	20	23	58	1	12	27	74	6	6	4	10	41	149
20	22	26	55	2	12	25	80	9	8	8	9	32	154
21	22	21	51	1	8	33	69	7	4	4	8	30	152
22	21	25	49	1	11	35	65	6	6	4	8	25	135
23	16	18	48	1	5	27	48	5	6	4	5	24	151
24	25	15	43	1	6	25	44	5	11	7	6	20	140
25	21	21	40	1	6	33	36	3	6	3	4	20	147
26	14	18	37	1	6	29	36	3	7	6	4	19	138
27	14	20	30	0	6	25	50	2	3	3	5	20	120
28	13	15	26	0	7	28	64	2	4	6	5	20	127
29	14	15	25	0	4	22	53	2	3	2	5	18	100
30	16	18	21	0	7	22	50	1	3	4	4	15	105
31	9	12	35	0	8	23	55	1	3	3	5	17	109
32	10	11	29	*	7	21	40	3	3	3	5	16	105
33	8	17	24		7	24	54	1	4	3	2	18	95
34	8	18	21		4	22	37	2	3	2	5	21	92
35	10	18	19		7	21	41	2	2	5	4	12	99
36	5	18	17		3	23	63	2	3	3	7	9	85
37	15	13	15		6	21	46	2	4	3	5	8	94
38	7	15	19		1	25	33	3	2	3	6	7	110
39	10	13	16		4	27	37	3	2	1	2	9	80
40	3	12	14		4	20	37	2	2	2	2	16	84

P. P.	Subjects												
	P	Q	R	S	T	U	V	W	X	Y	Z	AA	BB
41	7	12	10		5	14	53	I	2	2	0	17	70
42	5	12	9		3	19	39	I	I	2	4	9	72
43	4	13	12		5	21	41	I	2	I	3	8	76
44	3	11	23		5	17	42	I	0	2	I	5	74
45	4	12	20		6	13	31	I	3	I	0	5	78
46	4	10	17		2	14	33	I	I	I	I	6	65
47	8	9	14		3	8	37	I	I	I	0	10	72
48	3	11	13		3	10	34	3	I	2	0	4	74
49	6	12	11		5	10	27	I	I	2	2	4	65
50	6	14	9		6	13	38	3	0	I	I	6	60
51	2	10	9		5	14	33	I	0	2	3	4	67
52	3	10	8		4	10	30	I	0	I	0	3	69
53	3	7	12		4	12	32	2	I	2	0	2	65
54	5	4	10		1	12	28	I	0	I	2	3	74
55	5	9	9		4	10	35	I	0	2	0	5	72
56	4	6	8		2	7	31	2	0	I	I	3	55
57	3	9	7		0	5	35	I	0	I	I	2	55
58	2	7	7		4	10	34	0	0	I	0	0	57
59	6	5	8		2	8	21	I	*	I	0	3	65
60	3	8	7		2	8	25	I		I	0	2	75
61	5	9	7		2	12	31	I		I	0	0	87
62	4	4	7		4	12	15	0		I	0	0	70
63	4	6	6		3	11	23	0	0	I	*	2	72
64	3	6	6		4	5	35	0	0	I		*	66
65	3	5	6		3	7	23	I		0			60
66	3	4	6		4	7	25	I		0			71
67	5	4	7		2	2	24	2		0			57
68	2	2	6		1	6	19	0		0			55
69	5	5	7		5	9	24	0		*			64
70	3	1	6		3	8	28	I					67
71	4	3	4		2	12	16	*					62
72	3	5	5		0	9	22						62
73	4	2	5		1	9	15						72
74	2	5	5		3	5	16						57
75	2	5	5		2	11	19						65
76	3	5	6		2	6	23						55
77	2	2	6		0	8	21						68
78	5	4	4		0	6	18						65
79	5	4	4		0	12	16						70

P. P.	Subjects												
	P	Q	R	S	T	U	V	W	X	Y	Z	AA	BB
80	2	2	4		0	8	16						64
81	5	6	5		0	13	16						60
82	2	7	4		*	4	23						62
83	5	3	4			4	30						58
84	5 -2	4 -3	3 -			1	15						*
85	4	3	4			5	15						
86	3	4	3			8	10						
87	3	5	3			3	27						
88	1	5	4			6	12						
89	1	1	4			4	24						
90	4	7	4			3	17						
91	2	6	3			4	13						
92	5	4	3			3	20						
93	3	3	3			5	20						
94	3	4	3			4	12						
95	1	3	2			5	10						
96	1	1	3			5	9						
97	2	3	2			6	11						
98	2	1	2			7	10						
99	1	2	*			1	15						
100	3	0				6	11						

(The four records incomplete at this point go as follows: *P*, 1, 0, 1, 0, 1, 0, 0, 0, 0, 0, 0, *; *Q*, 0, 0, 5, 1, 3, 2, 2, 0, 0, 0, 6, 2, 3, 1, 2, 0, 3, 1, 2, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, *; *U*, 3, 3, 4, 6, 2, 3, 2, 2, 3, 2, 3, 4, 6, 2, 7, 3, 4, 3, 4, 6, 6, 3, 4, 4, 3, 4, 4, 1, 3, 3, 2, 1, 1, 1, *; and *V*, 13, 15, 8, 8, 9, 10, 13, 4, 8, 15, 5, 10, 8, 11, 5, 7, 6, 13, 5, 6, 6, 8, 6, 7, 3, 6, 7, 6, 6, 11, 4, 12, 5, 4, *. Two of these, it will be noted, continue to the 134th period of practice, and are discontinued still unlearned, or below the standard set.)

larity of practice is concerned. The results of these irregularities are not obvious, though they no doubt affect the validity of any conclusions that may be drawn from the averages of the two groups as to the relative values of one practice a day and two practices a day.

Figs. 1 and 2 show the individual learning curves. Each curve is to be identified by the letter corresponding to the column in the tables which it represents. In these curves vertical distances represent the number of errors, or mis-catches, and the horizontal distances are successive practice periods of two hundred catches each. It is important to

note the frequent absence of plateaus in those curves which are carried to, or nearly to, a point of zero errors in 200 catches. Too much must not be made of this fact, however, as we shall see when we come to the interpretation of these, and of other, learning curves. The curves are marked by rapid initial declines, having the general form of the so-called normal curve of motor learning. *Up to the present time this rapid initial decline in errors or in time has been interpreted as showing a corresponding rapid initial learning—an unfortunate error*, due largely, it appears, to the ease with which 'errors' and time in seconds are measured, and to the lack of thorough-going criticisms of the significance of any such arbitrary measurements of learning. Various ingenious reasons why there should be rapid initial learning corresponding to the rapid initial decline of the error, or time, curves have been 'thought out.' We shall later on consider this matter more fully.

Cross lines on the curves, and the numbers that occasionally occur with them, indicate practice periods omitted, as has already been explained in connection with the tables.

Many interesting individual differences are noticeable in Figs. 1 and 2. *A* was a student of high rank in the class work. Her curve seems to indicate that some peculiar, obstructing habit developed early in the experiment, bringing the curve almost to a plateau. It is impossible to say how the curve would have developed if the experimentation had continued to the point of completion. In such a case a little help by the experimenter at the point of discontinuance would doubtless have resulted in a rapid drop at this point. Without such help *A* seemed to have approached a sort of pseudo-physiological limit, the learning likely being limited or obstructed by the formation of habits for certain constant errors. Curve *I* is a very peculiar one; it is also the curve of a woman. This student was older than most of the others, was above the normal in class standing, but was evidently less adaptable in this kind of physical performance than most of the other students. *C* and *G* are also women making poor ball-tossing records, *C* making also only a passing record

FIG. 1. Individual error curves of fifteen subjects practicing once a day. The small cross-lines and figures indicate practice periods missed. *Ja* is the curve of *J*'s average number of catches per practice period. The learning of any other subject may be given in a similar way. Letters underscored indicate women.

FIG. 2. Individual error curves of thirteen subjects practicing twice daily. Omissions of practice periods are indicated as in Fig. 1. *Ra* represents *R*'s learning in terms of average number of catches per practice period. Letters underscored indicate women.

in the psychology class. That women do not make the poorest records invariably is evident from the curves of *B*, *E*, *J* and *K*, all women. In Fig. 2 the three students making the poorest progress in the experiment are *V*, *BB*, and *R*, the first two being women and the last a young man who received a D grade in the psychology course. Here again *V* was one of the very best students in two sections of general psychology taught by the writer. *Y* is a woman of C grade in her class, and *S*, making the best record of all in the experiment, is a man who has had considerable experience playing ball. From these records it would be hard to give any reliable information for prediction as to one's success at the performance here practiced. Early interest in ball playing and practice at throwing and catching balls as children do in ball games would seem to be one of the best indicators of success in this experiment, but even here we could not say whether the practice had helped or whether the particular nerve and muscle conformation inherited had predisposed the subject both to interest in ball playing and to ability in ball juggling. The latter would seem to be the more probable condition.

Curves *Ja* and *Ra* should be studied carefully in relation to their reciprocals *J* and *R*.¹ It will be noted that, in conformity with the view suggested in this paper, the error curve changes most rapidly in the early part of the learning where the average attainment is small, or where the series are short. Each series is necessarily ended by an error. In the initial learning, then, there are a great number of series, or different trials, and hence a very small change in each, on the average, makes a big change in the total errors. The error curve is therefore liable to fluctuate more in the initial phases with the average curve advancing more constantly, while in the final part of the learning the error curve will assume a more constant or smooth appearance while the average curve will fluctuate between wide limits. If one to

¹ Note that the term 'reciprocal' is used here not in the strict mathematical sense of values whose product is equal to unity. Their product, with a slight qualification developed in subsequent pages, is equal to a number representing *K*, the constant of practice.

two hundred catches can be made successfully in a series, the total errors for two hundred catches will not change more than one or two units in successive practice periods, while the average number of catches may vary from less than fifty up to two hundred absolute units.

This reciprocal relation obtains between errors and average number of catches because the former directly condition the latter and *vice versa*.¹ This state of affairs is not so obviously present in some types of learning processes, but as we shall see it is well worth enquiring into the relation of errors to attainment in most such experiments.

Table III. gives the averages of the individual errors, column A. E. For reasons already explained these averages do not include the records of subjects *A*, *C*, and *BB*, bracketed at the top and the bottom of the columns. In determining these averages it has been necessary to take note of the fact that some of the columns terminate before reaching zero, no errors. In such cases the averages have been weighted as follows. If, for example, ten columns were being averaged, and one of these discontinued with the 50th practice period (P. P.), it would leave only nine columns from this point. Suppose the last five records of the column terminating here were 8, 5, 6, 6, 7. Suppose also that the sum of the remaining nine records for the fifty-first practice period was 27, then it would seem fair to add to the average, 3, of the nine columns about .3 (*i. e.*, $\frac{27 + 6}{10} = 3.3$). The amount added would

be gradually decreased to represent the decrease that the discontinued record would probably have had if it had been continued. This weighting of averages is to prevent the appearance in the curve of a sudden change that might be misinterpreted. None of the averages in the twice-a-day group is weighted. Only the last seven averages in the once-a-day group are weighted; these are of course less

¹ An error in ball-tossing ends in fact one *trial* (termed in this paper a series); and it is obvious that the number of trials *times* the average number of catches per trial *equals* the total number of catches. Hence dividing the number of trials ('errors') into the total number of catches we get the average number of catches, 'average attainment.' See next page.

TABLE III
AVERAGES OF ERRORS IN 200 CATCHES AND AVERAGE NUMBER OF SUCCESSIVE
CATCHES¹

Once-a-day Group						Twice-a-day Group					
P. P.	A. E.	A. C.	P. P.	A. E.	A. C.	P. P.	A. E.	A. C.	P. P.	A. E.	A. C.
1	107	1.8	34	12	16.7	1	97	2.1	34	12	16.7
2	75	2.6	35	11	18.2	2	93	2.2	35	12	16.7
3	69	2.9	36	9.3	21.5	3	74	2.7	36	13	15.5
4	55	3.6	37	9.1	21.9	4	68	2.9	37	12	16.7
5	52	3.9	38	10	20.0	5	55	3.6	38	10	20.0
6	46	4.3	39	8.3	24.1	6	50	4.0	39	10	20.0
7	41	4.9	40	9.1	21.5	7	45	4.4	40	10	20.0
8	39	5.1	41	8.3	24.1	8	44	4.5	41	10	20.0
9	37	5.4	42	8.3	24.1	9	41	4.9	42	9.7	20.9
10	33	6.1	43	8.4	23.8	10	38	5.2	43	9.3	21.5
11	31	6.5	44	8.5	23.5	11	33	6.1	44	9.2	21.8
12	28	7.2	45	7.0	28.6	12	31	6.5	45	8.0	25.0
13	31	6.5	46	7.2	27.8	13	33	6.1	46	7.5	26.7
14	26	7.7	47	7.5	26.7	14	29	6.9	47	7.5	26.7
15	25	8.0	48	6.8	29.4	15	30	6.7	48	7.0	28.6
16	23	8.7	49	6.4 ²	32.8	16	27	7.6	49	6.8	29.4
17	21	9.5	50	6.4	32.8	17	26	7.7	50	8.1	24.7
18	21	9.5	51	6.7	29.9	18	25	8.0	51	6.9	28.9
19	19	10.5	52	5.9	33.9	19	24	8.3	52	6.6	30.3
20	20	10.0	53	5.6	35.7	20	24	8.3	53	6.4	32.8
21	19	10.5	54	4.2	47.6	21	22	9.1	54	5.6	35.7
22	18	11.1	55	4.2	47.6	22	21	9.5	55	6.7	29.9
23	17	11.8	56	3.3	60.6	23	17	11.8	56	5.4	37.0
24	17	11.8				24	17	11.8	57	5.3	37.7
25	18	11.1				25	16	12.5	58	5.4	37.0
26	17	11.8				26	15	13.3	59	4.6	43.5
27	15	13.5				27	15	13.3	60	4.8	41.7
28	15	13.5				28	16	12.5	61	5.7	35.1
29	16	12.5				29	14	14.3	62	3.9	51.3
30	15	13.5				30	13	15.5	63	4.8	41.7
31	14	14.3				31	14	14.3	64	4.9	40.6
32	13	15.5				32	12	16.7	65	4.0	50.0
33	13	15.5				33	14	14.3			

reliable than the others. The averages of the errors in this group are necessarily discontinued earlier than those of the twice a day group; for a given number of days' practice will make the column only half as long as with two periods of practice daily. A number of the subjects did not get started early enough in the semester to complete the learning by the close of the academic year. While this is regrettable, it is to be noted that the records are complete enough for reliable data on the most important matters raised in the following

¹ P. P. = practice period; A. E. = average number of errors in 200 catches; A. C. = average number of successive catches in the P. P.

² Averages of this column are 'weighted' below this point, as explained in the text.

discussion of the significance of the curves. Our experiment is only to be regarded as preliminary, so far as affording data for finer points on learning curves is concerned. More data are desirable particularly toward the end of the learning process in the neighborhood of the 'physiological limit.' It will be noted that so far as our averages go in the once-a-day practices, they show more rapid learning for equal practice than do those of the other group with more frequent periods. This agrees with most other experiments on the distribution of practice. Individual variations, however, are so great and the number of subjects so small as to invalidate any conclusion on this point from our data. An unequal proportion of each sex in the two groups is another factor that may be responsible in a measure for this result, though the probability seems to be that a more equal sex distribution would have made the once-a-day results even comparatively better than they were.

In Table III. are also given the average number of successive catches, A. C., in each practice period, the average number of catches without an error. These averages were obtained simply by dividing the average number of errors, A. E., into 200 catches. In order that the successive practice periods might be as nearly equal as possible, the subjects were asked always to stop any series voluntarily, if it did not end there by a miscatch, when the total number of catches of the practice period reached 200.

But this voluntary discontinuance of the final series at 200 catches makes an error in our results which seemed unavoidable so far as the experimental procedure is concerned. This error is fortunately negligible in the first part of the learning process, that with which we are most concerned in this study, since no subject could have exceeded very far the 200 limit by continuing the series voluntarily stopped. As the average number of catches increases this error in our results becomes larger. More specifically, if a subject averaging four catches per series ends by a miss, say, at 198 catches and 50 miscatches or errors, the chances are that he could on the next series reach 202 catches (with 51 errors). Volun-

tarily stopping at the 200th catch therefore gives him fewer errors than he should have, and correspondingly increases the average number of catches. In this case one half error should have been added to those actually made. Instead of having an average of 4 catches to the series (*i. e.*, $200 \div 50$) he should really have 3.96 (*i. e.*, $200 \div 50.5$). If, on the other hand, one averaging nearly 50 catches to the series misses, say, at a total of 175 catches with 4 errors, and then makes up the 200 total and stops voluntarily, he should have not an average of 50 catches to the series ($200 \div 4$), but only one of 44.44 catches ($200 \div 4.5$). In the former case the voluntary discontinuance at 200 made a negligible error in results (1 per cent.), while in the latter the miscatches, 'errors,' should be 12.5 per cent. more than would appear from our method of procedure, and the average number of catches would be 12.5 per cent. less.

Fig. 3 presents in a graphic form the data of Table III., without any correction for the error just pointed out. Curves *E*, I., and *E*, II., represent the number of errors (miscatches) in the 200 catches of the one-practice-a-day group and the two-practices-a-day group, respectively. One practice a day seems to be a slightly more effective distribution than two a day. It is important to note—and this is one of the main contributions of this paper—that these curves are the so-called *normal curves of motor learning*, showing a rapid initial decline and a more gradual decline later, as is the rule with such curves. That this is not evidence of rapid initial learning is obvious from the fact that the other curves in this figure, *A*, I., and *A*, II., representing the average number of successive catches without a miss, have a different form altogether though they represent the same learning process as the *E* curves. That is to say, the progress of the one-practice-a-day group is represented by the curve *E*, I., and also by the curve *A*, I. If we consider only the former we are in danger of concluding—as has been the practice among practically all psychologists for error curves up to this writing—that the learning begins at a rapid rate and gradually decreases in absolute amount in the successive practice

FIG. 3. *E, I.* and *A, I.* are the error curve and the curve of average number of catches per practice period of thirteen subjects practicing once a day. *E, II.* and *A, II.* are the corresponding curves of twelve subjects practicing twice daily.

periods, being very slow toward the end of the experiment. On the other hand, if we consider only the latter curve, just the reverse of this condition will seem to be true. The only fair conclusion, taking both curves into consideration, would seem to be, so far as the data considered are concerned, *that we do not know anything reliable at all concerning the kind of learning in question, as to whether it is initially more rapid or more slow than at certain later stages.* It would seem safest to assume that in the case of ball-tossing, at least, learning goes on throughout, up to a point approaching a physiological limit, at about a constant rate. As we shall see later, there is good reason to suspect that the same thing is true of other learning processes. There is likely no real difference in this regard between the so-called motor learning and the sensory discrimination learning in animals. The distinction usually made between these 'kinds of learning' has long seemed artificial and erroneous to the writer, but only recently, in working up the results of the present experiment, did he secure a good basis in fact for the view that the distinction is incorrect. Every case of motor learning must involve some sort of 'sensory discrimination,' and conversely, all sensory stimuli must come to have some sort of motor significance.

One remarkable result of our experiments is the almost complete absence, in the individual as well as in the averaged results, of any important plateaus. As we shall see later, this may not mean that there were no actual plateaus in the learning, but it may indicate that certain kinds of curves tend to conceal plateaus. The very fact, however, that some kinds of curves tend to show plateaus where there are no plateaus by a different statement of results, or at least no marked plateaus,¹ may tend to cast doubt on the whole doctrine of plateaus in learning.

¹ A degree of evidence for this statement comes out in subsequent pages. Plateaus are shown best in the error or the time curves if they occur early in the learning process; if they occur in the latter part of the learning they show best in the average attainment curves, but are practically concealed in the error curves. See Fig. 10, for example. Some methods of plotting show them much better than others. If, moreover, the average attainment curve (which in ball-tossing seems to be concave to the y-axis) has a rapid initial rise because of actual rapid initial *learning* a plateau may seem to

An experiment of the kind here reported, taking a considerable amount of time daily from each of a large number of subjects, makes it impossible, under the conditions of full-time teaching by the experimenter, to keep as close check on each student as is desirable. Even if the experimenter were entirely free from other duties it would be very difficult to get the time of a great number of subjects regularly for the number of months that this experiment requires. Closer personal supervision than was possible to the experimenter in the present case might do something toward reducing individual differences, and especially toward explaining the causes of certain marked failures to learn.

In the introspections, observations, and explanations of the irregularities many interesting points respecting difficulties and the methods of their solution, or failure of solution, came out. These are perhaps not different from those already reported in previous studies, and need not detain us here. One of the very best psychology students of the subjects failed utterly to get a fair start, even with a little personal assistance—unfortunately somewhat too late—from the experimenter. She had had almost no experience in throwing and catching balls, and could not, after the day she came for special help, break up in the time available a confusion of impulses which blocked one another much as in the attempts of a stutterer to talk. Frequently she could not even throw up the first ball and hold the other one, but would throw both together in a spasmodic manner making them go in different directions. After a few weeks of practice she was seldom able even to catch the first ball tossed. The effort was then discontinued, and her results are not included in the data of our tables.

Another subject, *A*, whose results are not included in the averages because they did not reach a point of even fair efficiency, advanced rather slowly after a short initial gain, and at the point of discontinuance the process of learning occur between this rapid rise in the curve and the one that occurs in the latter part of the curve, due purely to the *mathematical relations* of errors to average attainment. (See p. 199.) Such plateaus are not *plateaus in learning* at all. Just how marked a plateau is, is not easy to determine absolutely.

came to a condition approaching a plateau. This may have been due to the development of some such confusing tendencies as referred to in an extreme form in the preceding paragraph. This subject did not get time to complete the process up to a point that would indicate the real nature of this difficulty. The curve in question is given in Fig. 1. It should be noted that the learning had not come to a complete standstill. *A* was also a student making a good grade in the psychology course; she had had but little experience at ball playing in childhood. Subject *V* is another woman of whom the same things may be said, but her record was carried out to the 134th practice period (26,800 catches!), at which point she still made four errors in the 200 catches. All those who made extremely bad records were women, though some of the women subjects did exceptionally well. On the other hand, some of the men who had had extensive practice at playing ball advanced at an exceedingly rapid rate. Subject *S* is the best example in our groups. On the whole it appears that men are more rapid learners in ball-tossing of the kind practiced in the present experiment than are women, though our data are not sufficient, in view of the degree of individual variation found, to justify a detailed comparison or any definite conclusion on the matter. For the aid of any one who may desire to test this aspect of the problem out further and who may find use of our own data it may be added that the following subjects were women: *A, B, C, F, G, H, I, J, K, N, P, V, X, Y, BB*. Since the results averaged in the once-a-day group are of five men and eight women, and those averaged in the twice-a-day group are of eight men and four women, it is possible that the superiority of the former group would be greater with a more equal distribution of men and women. Since the time for practice had to be determined by the convenience of the subjects a lack of balance of this kind was not easy to avoid. It may be suggested here, too, that the very poor records of some of the women may not be due simply to the lack of practice at ball playing in childhood; on the contrary the failure to take up with this popular kind of sport may have been due

to an innate unfitness for it which likewise unfits the subject for ball-tossing.

The reports of the subjects show various individual differences in procedure, such as differences in methods by which the balls are tossed, in the methods of rotating them in the air to prevent collisions, in the distribution and frequency of 'bad days,' in the degree of correlation between feeling of efficiency and actual efficiency as shown by results (this correlation on the whole being very low), and in various other respects. A complete analysis of these differences with the data at hand would contribute but little to previous work, and is somewhat aside from our present purpose.

Statements like the following from *D*, which statement, by the way, came spontaneously from a number of the subjects, show the value of extensive experiments of this kind for students being initiated into scientific psychology. *D* writes: "I now feel as if I should like to continue something of this kind and gradually become proficient in several things. This experiment has surely taken the place of several books so far as learning is concerned, and I am very glad that I have had the opportunity to spend whatever time I have spent on it." The apparatus is simple, the directions are not hard to follow, and the working up of his own results is a most instructive exercise to the student.

We recur now to the error in results pointed out above.¹ Correcting for the effects of voluntarily stopping at the 200th catch would make the average attainment curves—those representing the average number of catches—appreciably less concave to the *y*-axis than is shown in Fig. 3; *i. e.*, more nearly approaching a straight line. In the averaged results of all the subjects it would make a difference of .5 per cent. in the first practice period; of nearly 1.5 per cent. in the tenth; of about 2 per cent. in the twentieth; of about 3.5 per cent. in the thirtieth; of about 5 per cent. in the fortieth; of about 7 per cent. in the fiftieth; and of nearly 10 per cent. in the fifty-sixth. In general it would seem fair to assume that the last series, which is stopped voluntarily each time,

¹ *Supra*, p. 197.

varies at random all the way from zero catches to the number of successive catches possible at any given stage of learning in question. We ought therefore to count, on the average, one half error for this series voluntarily stopped. Adding $\frac{1}{2}$ to the average of the recorded miscatches (errors) and dividing the resulting number into 200, we get the corrected average number of catches per series. A corrected table of the averages of the two groups of twenty-five subjects given separately in Table III. is given below (Table IV.). The formula used for determining the values A.C. in this table is

$$\frac{K}{e + \frac{1}{2}} = a.$$

This is a better empirical formula for conversion than that given above.

TABLE IV

A TABLE OF CORRECTED VALUES OF a AND e ; RECORDS OF ALL SUBJECTS OF BOTH GROUPS ARE AVERAGED TOGETHER (28 SUBJECTS)

P. P.	E.	A. C.	P. P.	E.	A. C.
1	102.5	2.0	29	15.5	13
2	84.5	2.3	30	14.5	14
3	72.0	2.8	31	14.5	14
4	62.0	3.2	32	13.0	15
5	54.0	3.7	33	14.0	14
6	48.5	4.1	34	12.5	16
7	43.5	4.6	35	12.0	17
8	42.0	5.0	36	11.6	17
9	39.5	5.1	37	11.0	18
10	36.0	5.6	38	10.5	19
11	32.5	6.2	39	9.6	21
12	30.0	6.7	40	10.0	20
13	32.5	6.2	41	9.7	21
14	28.0	7.2	42	9.5	21
15	28.0	7.2	43	9.3	21
16	25.5	7.8	44	9.3	21
17	24.0	8.4	45	8.0	25
18	23.5	8.5	46	7.9	25
19	22.0	9.1	47	8.0	25
20	22.5	8.9	48	7.4	27
21	21.0	9.6	49	7.1	29
22	20.0	10	50	7.8	26
23	17.5	11	51	7.3	28
24	17.5	11	52	6.8	30
25	17.5	11	53	6.5	31
26	16.5	12	54	5.4	38
27	15.5	13	55	6.0	34
28	16.0	13	56	4.9	41

FIG. 4. Corrected error and average-number-of-catches curves of twenty-five subjects.

Fig. 4, giving graphically the data of Table IV., shows that the concavity of the a -curve to the y -axis is not removed, though it is somewhat reduced, by the correction described. If the averages could be reliably carried further than to the fifty-sixth practice period—which they cannot because of some subjects ‘falling out’ with incomplete records—the a -curve would rise comparatively much more rapidly as the e values become very small; it would reach 400 when the actually recorded errors equalled zero. From this point the fraction, $\frac{1}{2}$, added to e should gradually decrease to zero, bringing the a -curve up to infinity. If fatigue were eliminated by the subjects’ resting occasionally, it is possible that this point on the a -curve would be reached (theoretically) early enough to give the curve a decidedly concave aspect to the vertical axis. In the figure in question the a -curve agrees with this view by taking rather suddenly a marked upward trend. It would be interesting, with a large number of subjects practicing up to a high degree of perfection, to trace the trend of this curve out further, and to construct a rational mathematical curve to represent it as the ‘normal curve’ of average attainment in ball tossing for the value of $K = 200$.

Now since $K = ae + \frac{1}{2}a$, it is obvious that the constant practice in successive periods can be measured accurately neither in the number of catches alone nor in that of throws alone irrespective of catches. Since e begins with a large value and decreases gradually towards, but never reaches, zero, more practice is obtained in early than in later periods if a practice period is defined as 200 catches (as is done in this experiment). There is reason to believe, however, that the additional practice in early over later periods is small, if not actually negligible in amount.¹ Throwing a ball so

¹ It should be noted that less practice in the early practice periods than was given in the present experiment would have a tendency to make the average attainment curve rise more slowly and the error curve drop less abruptly, making the former curve therefore more concave to the vertical axis than it is. A good way to test the effect of the excess practice in early periods is to have a large number of subjects take practice periods of 200 tosses, irrespective of the catches, and then to arrange the results into successive groups of 200 catches each and compare the curve with those in Figs. 3 and 4, which represent averages by our method. Some procedure that is a mean between these two opposite extremes in method would doubtless be the best one.

that it cannot be caught likely gives but little practice in throwing one so that it can be and is caught. Moreover, the method here followed has been common in animal learning, and in all other experiments in which *something is to be done* once or more times for each practice. In all such experiments the constants, or 'equal amounts of practice,' set for successive periods are not strictly equal. Just how much this inequality affects results in any given case is not easy to determine. The difficulty arises from the fact that as practice advances errors and irrelevant and confounding movements are constantly on the decrease while the successful or correct movements are increasing in frequency and in rapidity of performance. Equal time periods of practice has been suggested and extensively used.¹ While this practice has many advantages it also shares the disadvantages here pointed out. A skilled juggler may toss and catch the balls two dozen or more times while the beginner fumbles about with possibly only one successful catch. Are such practice periods equal?

But the greater error comes not from the inequality of practice as such, but rather in the statements of results. This is a most deceiving thing. Total work performed, or average amount of work without (an arbitrarily defined) error, or with a given number of efforts, or within a fixed amount of time, becomes the unit and is represented usually by the ordinates. Abscissa units usually stand for successive practice periods however defined. Such definitions of ordi-

It is, of course, obvious that if our method errs in the direction of excess practice in early practice periods, that of equal tosses errs the other way in as much as wild tosses doubtless give less practice in throwing-and-catching than do tosses which result in catches. Our own method, as has been pointed out in the body of this paper, is more in harmony with the usual practice, and it also gives equal results—200 catches—in the several practice periods. My friend and colleague, Professor W. S. Miller, of the department of education, has kindly helped me to convert the data from one of his representative subjects by the equal-number-of-tosses method into average attainment for comparison with my own curves in Fig. 3. Only about the first half of the data from this subject were thus converted. So far as we went no noticeable difference from our own averaged results appeared. His practice periods, however, were 300 tosses instead of our own 200 catches. So the results are not strictly comparable. It would seem that the inequality of different practice periods by either method will show little effect in results so far as the mere practice differences are concerned. But see the next paragraph in the text.

¹ E. g., Thorndike, *op. cit.*, 295 ff.

nate values presuppose a homogeneity of the data measured; but as has been hinted above no such homogeneity exists in the majority of cases. Practice in the early part of the experiment is vastly different from practice in the later part, or, stated differently, no account is taken of how the attainment measured is conditioned reciprocally by 'errors,' of whatever nature they may be. This difficulty is so great, it seems to me, as seriously to invalidate the conclusions arrived at as to the different absolute rates of learning in different stages of the learning process.

In ball-tossing one might define a practice period in terms of *throws*, say as 200, 300, or 400 throws, thereby avoiding the difficulty suggested above. But this does not really avoid it, for all throws are certainly not of equal practice-for-throwing-and-catching value. In the second place one is deceived by such a method when one comes to the definition of results. In one case, not yet published, of which the writer is aware 'the score,' where this method was followed, is the number of catches. Not only is it wrong to assume that a bad throw—such as the beginner is prone to make—gives as much opportunity for a catch as a good throw; but it is even more fallacious to count only the *absolute* number of catches as results, since the *ratio of catches* to the *possible number of catches* is constantly changing. Such results will always, it would seem, give, deceptively, a rapid initial rise in the curve and negative acceleration later. This is due to the fact that out of, say, 200 tosses only 200 catches are possible, and as the 200-catch limit is approached the proportion of catches to possible-catches approaches a limit of 1 : 1. The influence of errors (both of throwing and of catching, if they can be successfully distinguished) is therefore not only a variable quantity directly affecting the catching, but reciprocally and indirectly the number of errors determine the ratio of actual to possible catches. From the very nature of the conditions the limit (in terms of the score) is physically imposed and it must be approached asymptotically. We are bound under such conditions, because of this relation to the errors not being taken into consideration,

to get a curve of attainment which resembles the 'normal' learning curve! All of which is only another way of stating the fallacy of counting results, or the score, in absolute terms of attainment, or of errors, only.¹

It should be noted that *error* as defined in the present experiment—a miscatch ending a series of catches—is practically the same as *trial*. It is perhaps more clear that results should be equal to the number of trials *times* the average amount of attainment per trial, than that they equal the 'errors' *times* the average amount of attainment. The reason that the term *error* is used here is that it is in common use in similar relations in psychological literature and that each trial ends by an error, or miscatch. It is also desirable not to confound trial with single tosses, or throws. Counting results in terms of single throws is liable to result in curves of a deceptive nature, as explained in the preceding paragraph, curves which seem, wrongly, to indicate that *learning* is at first comparatively rapid and that it later approaches a limit very slowly.

Because of certain widely current statements about the absolute rates of learning in its various stages, it seems worth while to carry this question of the significance of certain learning curves, raised in the ball-tossing experiment in particular, over into the more general field. Here prudence suggests, in view of the time that has been available for this study, that we proceed very tentatively. The aim is therefore rather to open up the problem and to stimulate further experimentation than to give it any appearance of finality. If anything like dogmatic statements appear, this is to be explained on the ground of brevity and of the belief

¹ The results of Mr. Batson, published in a recent *Psychological Monograph* which appeared after the present paper was sent to press, confirm the general position here taken. Mr. Batson has plotted his results in the three different ways discussed in this paper, and gets for error curves the rapid initial decline usual with such curves; for what we have called average attainment he gets curves which are concave to the vertical axis. Plotting a curve of the number of catches in successive practice periods with a constant number of *throws* he gets a curve which in general shows a rapid initial rise and a final asymptotic approach to the "physiological limit," or better to the 1 to 1 ratio. He does not, however, enquire into the meaning of the different kinds of curves.

that they may be more effective than highly guarded suggestions in bringing about this result.

AN EXAMINATION OF THE NATURE OF LEARNING CURVES

Why do our curves in this experiment, representing errors on the one hand and the average amount of accomplishment on the other, seem so contradictory? Both of them represent kinds of data (results) frequently used as bases of inferences as to relative rates of learning in the different stages of the process. Moreover, error curves and attainment curves are frequently compared directly without any kind of transformations. As has been pointed out in the specific cases considered, errors and average attainment have reciprocal relations to each other as dependent variables. Each factor conditions the other in such a manner that it is entirely unsafe to define learning in terms of absolute units of only one of these processes. In most cases, as was found true in ball-tossing, time curves and error curves have the same general form, so that the former have the same relations to the curves of average attainment that the latter have to them. In certain problems, such as maze learning, excess distance curves show the same general nature as time and error curves.¹ Of course, in these various cases 'errors' are not the *same kind of things*, even though they may be safely comparable. Much depends upon how they are defined, as is well known.² On the whole more errors means more time taken for a certain result in terms of performance; the same thing is true in a rough way of the relations of excess distance to performance or attainment. For the general purpose of this enquiry, then, we may confine our-

¹ Cf. Hicks, Vinnie C., 'Relative Values of the Different Curves of Learning,' *J. of Animal Beh.*, 1911, 1, 138-156.

² This is true even within the same kind of experimentation. For example, in the case of animals learning the maze an error may be defined as merely putting the head into the *cul-de-sac*, as entering it with the full length of the body, or as going the full length of the *cul-de-sac*. A return in a single section without a bend may also be counted as an error along with any complete return, or the latter may be counted for as many errors as there were sections passed through in the return, or, a third possibility, no returns may be regarded as errors if the animal does not depart from the true path.

selves to error curves, keeping in mind that, with certain cautions, time or excess distance curves may be substituted for them without serious error.

The reciprocal relations of errors and average attainment, as we have defined these terms in ball tossing, are obvious. An error always ends a series. When the error count is large the average number of successive catches (the 'series') must be small, and conversely. Suppose that on the average only one catch can be made in a trial, or a series, at the beginning of the learning. Then the errors in a practice period of 200 catches must amount to 200. But an improvement or gain of one in each series (making an average of two successive catches per trial) will bring the errors down to 100, a change of 100 absolute units. An average of 3 catches each trial brings the errors to 66; of 4, to 50; of 5, to 40; and so on. If we plot these values on an absolute scale, representing errors by e and average number of catches by a , we get a 'representative learning curve' of the e -values when the a -values advance by equal absolute amounts. That is, the e -curve drops rapidly at first and then progressively less and less. This is of course on the assumption that we give these values the relations indicated in the equation

$$\frac{K}{e + \frac{1}{2}} = a,$$

and regard K as a constant. Very small changes in a will mean very large changes in e , and *vice versa*. Proportionately, of course, the changes are equal in both factors; *but learning curves are almost universally plotted in absolute amounts, despite this fact*. In the beginning of any learning process the values of a *must be small*; hence e *will always change rapidly on the initial part of the curve* thus plotted. When a gets rather large, e must change slowly, if the changes in a , absolutely, are constant. These changes in the error curve are obviously in no sense indications that the *learning* is at first rapid and that it then gradually decreases in rate.

Let us take a theoretical case of ball-tossing and assume that the increase in the average number of successive catches

—the a -values—is constant, adding one unit for each successive practice period. In such a case the a - and e -curves will have the forms shown in Fig. 5, when $K = 200$. When K is given a value of 40, the e -curve takes the form of e_1 in Fig. 6. The curves e_2 and e_3 correspond to gains in the

FIG. 5. Curve e is a theoretical error curve corresponding to an increase in a , average number of catches, of one catch in each succeeding period, when $K = 200$.

a -values of $\frac{1}{2}$ and $\frac{1}{4}$ units, respectively, in successive practice periods. These e -curves all have the general shape of the 'normal curve of motor learning,' falling off rapidly at first and more and more slowly as the learning reaches an advanced point. This type of curve is *normal* to motor learning

only because errors, and number of seconds (time), are the easiest data to get in such processes! This unfortunate circumstance, as it has proved to be, a mere contingency, seems to have led to a misapprehension of the real nature of learning. The relative rate of learning at its different stages has come to be judged, in the hands of not a few investigators, by the rate of decline of this curve. It should be noted that the smaller the values assigned to K the more gradual the decline of e . Relatively slow, constant gains of the a -values also affect in a similar way the e -curve. A rapid initial rise of a produces, of course, a more abrupt fall of the e -curve, as shown in Fig. 7, curve e .

FIG. 6.

FIG. 6. e_1 , e_2 , and e_3 are theoretical error curves corresponding respectively to equal absolute successive gains of one, one half, and one fourth in the average numbers of catches, a_1 , a_2 , and a_3 . $K = 40$.

FIG. 7.

FIG. 7. e and e' are theoretical error curves derived from the average attainment curve a , when $K = 40$ and 200 respectively.

The *form* of the a -curve is a third factor that affects the reciprocal e -curve. If, for instance, a begins with a value of 1 and increases in each succeeding period by 1/10 of its value in the next preceding period, we get curves like those shown in Fig. 8, when $K = 200$. These curves resemble very closely those actually obtained in ball-tossing by certain subjects. Whether this type of a -curve, representing *equal proportional gains*, comes nearer being a constant progress in

FIG. 8. Curve *a* represents equal proportionate gains of $1/10$ over each preceding period, the initial value being 1. *e* is the corresponding error curve when $K = 200$.

learning than equal absolute gains of the a -values is a question that may be answered differently by different authorities and by the same authority for different learning processes. At any rate when plotted in absolute units the curve of such a gain is concave to the vertical axis, as are the average attainment curves of ball-tossing. This correspondence might suggest that learning in ball-tossing advances by approximately equal proportionate gains; but to take this stand would be to get into a questionable position, to make the mistake with a slightly different application which has been made to date by certain psychologists in the interpretation of error and time curves. When the average attainment is great the errors are few, and because of the reciprocal relations of a and e , pointed out in the foregoing, *slight changes in errors must mean large changes in average attainment*, when these values are plotted in absolute units.

If the a -curve has an initial rapid increase, as shown in Fig. 7, making a convexity toward the vertical axis, the initial drop of the e -curve is increased in amount. If the K in our equation, by which the e -values were derived, is decreased in value the initial drop will decrease correspondingly. (Compare curves e and e' , Fig. 7.) It is of course obvious that any tendency to a rapid initial increase in average attainment, as defined in this paper, will correspondingly increase the decline of errors, or trials.

When we consider these relations of curves of different kinds of data obtained in learning it becomes evident that great care in the comparisons of different learning curves is advisable. All 'errors' and 'scores' are more or less arbitrarily defined, and it seems necessary for comparative purposes that some standardized methods of conversion of data be devised. At present we are forced to recognize the insecurity of any conclusion as to the absolute rates of different stages of learning, based on such curves of learning as are available. As a rule curves of learning are based on the total amount of accomplishments of the kind specified as 'results,' but in many cases they are constructed on the basis of what *can be done* in a stated number of minutes

after certain multiple units (weeks, hours, or minutes) of practice. In such cases the initial stages do not allow of as much practice *in the particular aspect of learning that is plotted* as do later stages, for more time absolutely is lost on various conditioning functions. The considerations of some of these matters may even result in the view that a relatively more rapid initial learning takes place in certain cases than is commonly assumed from the nature of the curves of the data obtained. On the whole, however, the evidence for the general statements current, that learning is initially rapid and gradually decreases in rate to a physiological limit, is far from satisfactory. Meanwhile it must be recognized that the practice of putting into graphic form the 'progress' in different learning processes studied, has had fruitful results toward a more quantitative psychology of learning and toward the establishment of a better knowledge and control of various aspects of learning in educational fields. It may be asked, May we not still continue to plot curves of various learning processes and to say certain definite things about them? Yes; provided we are careful as to *what* we say, and as to what comparisons we make. The curves are not the learning process and in most cases they represent it but poorly. For example, plateaus may not be plateaus *in learning* at all; we may be plotting some aspect of the process which for the time is neglected for some other aspect. On the other hand, a real plateau on the latter part of an error curve, where change is slow, may be practically obscured, while it would have been much more conspicuous in another type of curve.

SOME APPLICATIONS AND SUGGESTIONS

In every case of learning scientifically studied some definite units of practice must be decided upon—food must be obtained once or several times by an animal either escaping from or getting into a problem box, or by one finding its way past a certain number of blind alleys to the food box in a maze; a given number of catches of balls, of additions of numbers, of throws of javelins, of substitution of symbols,

or of distribution of cards must be made; or, a more comparable and constant procedure in many cases, a given number of minutes or hours must be devoted to a certain specified kind of practice. Whatever the unit of practice may be we may designate it, with certain cautions, by a constant, K .

The score, or results, at any stage of the learning may be determined in at least two, and probably only two, general ways. First, the number of errors (or trials)¹ made in performing the 'work' designated by the constant—such as errors in getting to the food, in making 200 catches of the balls, in striking within certain limits in javelin throwing, in adding certain groups of numbers, in tracing a star, etc.—may be recorded; or the number of seconds required for some such performance may be indicated;² or we may in the maze problem take the excess distance covered. In the second place, we may choose the reciprocal results, the amount of performance in certain units (determined by the nature of K) which can be reached on an average without an error or in a given number of seconds. The first of these classes of data on results will give what we have called ϵ -curves; the second, what we have in a general way spoken of as average attainment curves. In the following very tentative comparison of curves taken almost at random from various important experiments on learning by different investigators we shall try to keep distinctly in mind these different factors. In doing this we in no way claim for our classification and formula superiority over any others that may be proposed. We shall use K always to designate the constant of practice; ϵ , the errors, number of seconds (time), or excess distance in maze; a , the amount of performance, attainment, or average attainment. For the transformations of the data from various experiments we shall take some form of the

¹ See *supra*, p. 205.

² In experiments on typewriting, telegraphy, crossing out letters or numbers, etc., an error may be defined as failure to make two successive strokes within a given fraction of a section. This would not be greatly unlike our errors in ball-tossing, since too much attention to the case at hand would make an error in the next one. Here again time and error curves show close relationships.

equation,

$$\frac{K}{a} = e,$$

to suit the particular conditions of the case in question. Our purpose is to attempt schematically, under some one of the two classes of results defined, a comparison of data from different experiments. It is obvious that different values of K , whether stated numerically or otherwise, and different arbitrary definitions of e , particularly when e designates errors, will wholly determine the values of a and therefore the fruitfulness of our comparisons.¹ Our purposes are, however, chiefly to illustrate a procedure that is possible if the various experiments are originally planned for comparative purposes.

In most reports of experiments only two of our factors, K and either e or a , are given, and from these the other one—even if only a theoretical factor—may easily be derived by the simple arithmetical process indicated by our equation. We may thus in a general way compare different ‘kinds’ of learning and see whether there seems to be so much difference as has generally been supposed, or we may at least suggest a method for comparisons in future experimental work. In our illustrative cases considerable difficulty has been experienced because of the fact that data were not fully given in all cases. To facilitate comparisons we will make all e -curves represent decreasing values and a -curves increasing values. This will require in some cases simple transformations of some of the authors’ results before the application of our equation is made, but it will not change their values. In every case the curve represented by a continuous line (——) is the curve giving the original data of the author, and the curve in a discontinuous or broken line (— — —) is the one supplied by the present writer through dividing the original values (the a ’s or the e ’s) into the constant, K . The broken

¹That is, even for purposes of comparisons between different experiments of the rates of change in e - or in a -curves, we must, to make our conclusions safe, keep the other two factors of the equation, K and a (or e) approximately equal in both cases. We must also be sure that our e (or a) values compared are comparable.

line therefore represents only another statement of the results supplied by the original author, one, however, which may be more directly comparable with results of other experiments.

Since it is manifestly impossible to rearrange the data from most of the experiments referred to so as to get accurately comparable values of K , a , and e , we can expect only a general conformity of the curves compared. More quantitative comparisons must come from results of experiments deliberately planned for comparative purposes, thus giving the different factors of the transformation equation as nearly equal values as possible. The relative size of the abscissa and the ordinate units is another matter that greatly determines the form of the curve, as is well known; but with unequal values of the factors of our equation it is manifestly incorrect to make these units in the several cases all equal. This makes it clear that our comparisons must be regarded only as illustrative.

In Fig. 9 curve a represents the average number of catches without a miss in practice periods of 200 catches each, by Swift's subject A , and e is the error curve, corresponding to the error curves of our own subjects. The data for the average number of catches was taken from Swift's curve,¹ and it was rearranged into such a form that successive practice periods are equal so far as the number of catches are concerned. As is well known the actual practice periods in Swift's experiments were far from equal. The error curve, whose values were obtained by dividing the successive average number of catches into 200, corresponds very closely to many of our own error curves. The learning of Swift's subjects has usually been considered to differ from the normal type of learning. There is evidently little ground for this view when similar results are compared.

Bryan and Harter's classical receiving curve for learning telegraphy is represented by a in Fig. 10.² A corresponding time-curve is represented by e , the number of seconds required for 'receiving' five letters. In this case a difficulty

¹ *Amer. J. of Psychol.*, 1903, 14, 201-251.

² *Ibid.*, 1899, 6, p. 350. Subject W. J. R.

FIG. 9. Error curve, *e*, and average-number-of-catches curve, *a*, of Swift's subject *A*, rearranged into practice periods of 200 catches each.

is encountered. The original curve, *a*, of the authors represents the number of letters per minute *after the several successive weeks' practice*, and is not an average of the whole practice period. It seems probable that with a proper adjustment of ordinate and abscissa units, and with a slightly different tabulation of data—which is impossible for us to

FIG. 10. *a* is Bryan and Harter's receiving curve (number of letters per minute), telegraphy experiment, subject W. J. R. *e* is the corresponding time-curve, representing number of seconds for receiving five letters. Abscissa unit = one week of practice.

show because the complete data are neither given in the original report nor embodied in the curve—we should get the 'normal (time) curve.' The curve here given from such data as was available has an initial drop that is very sudden. Evidently, however, this is the curve that we must compare with other time and error curves, not that (*a* in the figure)

originally given by the authors. It is interesting to note that the plateau, so marked in the original, or average, curve, is not noticeable in the time curve, because it occurs where absolute changes are very small in this curve, on the principle already explained in this report. This may be a valuable suggestion as to the reason why some experiments on learning do not show, and others do show, plateaus. Speculation here without further data, however, is unprofitable.

In animal psychology it has become customary to regard 'motor learning' and 'sensory discrimination learning' as somehow different in *kind*. As is well known, both kinds of learning are frequently represented in terms of errors. But their error curves are markedly different. To quote from Yerkes: "The former [error curves of the labyrinth tests] fall very abruptly at first, then with decreasing rapidity, to the base line; the latter [those for the discrimination test], on the contrary, fall gradually throughout their course. Evidently the labyrinth habit is more readily acquired by the dancer than is the visual discrimination habit. Certain motor tendencies can be established quickly, it would seem, whereas others, and especially those which depend for their guidance upon visual stimuli, are acquired with extreme slowness."¹ Watson, in 'Behavior,' seems to take the same position in general, making different divisions for discussion of sensory and of motor habits.² While it is, of course, not to be denied that sensory discriminations may be made very difficult for an animal, so that learning to discriminate will be slow comparatively, it should not be overlooked that the same thing is true of 'motor learning' and that a difference based on sensory and motor habits cannot properly be made. The distinction has always seemed to the writer to be artificial and based on bad psychology. Any habit may be regarded as either sensory or motor, or better, both, according to the point of emphasis. This was long ago made clear by Dewey in his well known article on the reflex arc concept.

One of Yerkes's sensory discrimination curves,³ modified

¹ 'The Dancing Mouse,' p. 236.

² Chapter 6.

³ *Op. cit.*, p. 230.

to represent increase over 50 per cent. correct responses, as explained in the foregoing, instead of decrease from 50 per cent. error as stated by the author originally, is shown in Fig. 11. Here it is the usual average attainment curve, as might be expected from the data which it represents, even though it is called an error curve.¹ Now, on the basis of

FIG. 11. Curve *a* is rearranged from Yerkes's curve of brightness discrimination by the dancing mouse. *a* represents per cent. correct choices above 50 out of ten trials. *e* is the corresponding (theoretical) error curve derived with a value of $K = 50$.

our equation it is possible to get a value corresponding more nearly to the usual error curve of so-called motor learning. For this purpose we have used the equation

$$\frac{K}{a} - 1 = e,$$

so that the error curve when $a = K$ will reach the zero line, as would actually be the case. Taking $K = 50$, the highest per cent. correct above 50 that is possible, and getting our values of *a* by subtracting the values given in the last column

¹ This statement, it is frankly confessed, needs more study and analysis. It is not easy to see in the present case the values comparable to our *a*- and *e*-values; but the curve given by Yerkes is certainly not an *e*-curve, as here defined, and should not be directly compared with the error curve in the maze learning as is done. If Yerkes had compared it with some sort of attainment curves he could not have come to the conclusion that he expresses.

of Yerkes's table from 50, we obtain the ϵ -curve shown in Fig. 11. Whatever this ϵ -curve represents concretely in the present discrimination experiment, it is certainly interesting to note that it is almost identical with Yerkes's own error curve of the labyrinth test.¹ Of course, this may not be wholly convincing since the ordinate and the abscissa units are not respectively equal in the two cases, but it is very suggestive of a better mode of comparison of curves than has hitherto been used, one that should avoid some of the errors made in comparisons in the past. It is, of course, always to be admitted that in spite of any manipulation of data one learning process may be made very much more difficult than another. By the use in comparative psychology of some such conversion method as here suggested, however, we may avoid the errors of making misleading comparisons between curves which may chance to be called by the same name and thus conclude that there are different kinds of learning, and we may also avoid erroneous statements regarding the absolute rate of learning at different stages of the process.

Murphy, in a recent article,² shows the usual confusion between error (or time) curves and average attainment curves. His subjects practiced at throwing the javelin. Five (?) throws were made by each subject at each practice period, these periods being distributed differently for different groups. Results for each practice period were given in terms of the average distance in centimeters from the bull's eye. "In the above experiment," the author concludes, "the curves did not follow the normal curve of learning to any marked degree; but those in the alternate day group generally gave a better approximation to a regular learning curve."³ As a matter of fact, Murphy's curves fluctuate a good deal because of the comparatively few trials in each

¹ 'The Dancing Mouse,' p. 235.

² Murphy, H. H., 'Distribution of Practice Periods in Learning,' *J. of Educ. Psychol.*, 1916, 7, 150-162.

³ *Ibid.*, p. 161. Murphy plainly refers to the error, or time, curve as the 'normal curve of learning.' Such curves will be found on examination almost universally, if not quite so (even in spite of the fact that errors are differently defined) to correspond rather closely to our ϵ -curve in the foregoing. It is rather surprising that this fact has not received explicit attention before, as it does not seem to have done.

practice period. Some of them, however, show a fair approximation to the average attainment curves usually found. His own results were plainly average attainment. A bunching of his results into more practice for each period—a process impossible to us because he does not give individual results—would obviously give more uniformly declining curves. If, then, the results were stated for successive periods in terms of differences from the largest average distance from the bull's eye (*i. e.*, from an error of 82 cm.), we should have a gradually rising average attainment curve very similar to the usual *a*-curve as defined in this paper. It is unfortunate that Murphy's results are not given in enough detail so that the transformation suggested here could be made as illustrative. To compare curves of average attainment, as Murphy has done, with the 'normal curve' (of error or time) is to make an unfortunate error that is all too common in psychology and education. All curves of different aspects of learning have frequently been regarded as somehow representing the same thing—the learning process—however different the data may be which they represent in the various kinds of practice.

In the javelin-throwing experiment it is not at once obvious what corresponds to errors in an experiment such as ball-tossing or the learning of the maze. Let us suppose, however, that every throw missing by more than a given distance, say 10 cm., had been called an error, and that every hit within this limit had been counted a score. Under these conditions if each practice period consisted of some appropriate number of hits (hits within the arbitrary limit) it would seem probable that the results would be comparable in a satisfactory way with those of our ball tossing experiment, for instance. It is true that the correspondence of the data from two such experiments would yet be only approximate in as much as the balls tossed and not caught would give relatively less practice in tossing-and-catching than throwing a javelin slightly outside of the 10 cm. limit would give in throwing-the-javelin-within-the-limit. It must be noted that in our own experiment the successive practice periods are equal in amount of practice only in as far as

throws that are caught are counted as *practice*. Though there is some degree of justification in not counting as practice balls thrown wildly and missed, it is true that many balls are missed through the merest slip in technique, such as letting one ball strike the edge of another when a fraction of a millimeter to the side would have resulted in a catch. Certainly there is practice in such a throw and miscatch and one learns by it, among other things, to avoid such collisions.

In comparing results of different kinds of learning we must exercise considerable care not only to get curves of the same general type, as above illustrated, but also to see that practice is approximately equal in successive periods in each of the experiments compared and that the practice periods of each kind of learning are about equal. The writer hopes to find opportunity in the near future to obtain data of more nearly comparable values in two or more kinds of experiments than exist at present. It must be emphasized, however, that errors resulting from slight inequalities of practice periods are really less serious than those from errors in arrangement of data, some of which have been briefly and very imperfectly pointed out in this paper. It is unfortunate that important comparisons of incomparable data, and that classifications of different 'kinds' of learning have worked into psychological literature and thought with scarcely any attention to this matter of basic import to any reliable comparisons among learning processes in different kinds of performances.

A single result of taking time, or error, curves to signify changes in the learning process itself comparable to changes in the curves may be found in Hobhouse's early reaction¹ to certain statements of Thorndike. Thorndike showed time curves of various animal-learning processes and contended that learning resulted from gradual physiological changes in the nerve tracts. Hobhouse, it will be remembered, objected that the curves in question were in reality steeple-like, the sudden drops being indicative of the influence of something like ideas, or 'practical judgment.' Thorndike did not apparently make the mistake of assuming a corre-

¹ Cf. 'Mind in Evolution,' 1901, 145 ff.

spondence between the curves and the actual learning, and his general position was against such an assumption; but he did not explicitly show why the rapid initial drop occurred. He has more recently in the volume already referred to, for example, given various reasons why learning should be more rapid initially than later in the process. These reasons may, indeed, apply in certain cases; but certainly the whole matter needs further experimental investigation. Hobhouse's arguments, so far as they were based on the steeple-like features of the curves, were certainly groundless, though critics of his position have been slow to point out this fact, apparently accepting his own assumption. Numerous statements might be quoted, besides those already given, to show that this has been the usual understanding of the significance of learning curves.

SUMMARY OF RESULTS AND CONCLUSIONS

As a result of experiments on ball-tossing with twenty-eight subjects it has been found that the learning curves of ball-tossing are not different from learning curves of other processes, provided that similar kinds of data be compared. The error curve of ball-tossing, for example, has a rapid initial decline as have error curves—the usual 'normal curve'—of various other types of learning. The curve of the average number of catches per trial, on the other hand, rises gradually at first, making almost a straight line, and more rapidly later. This gives it a concave aspect to the y-axis, as has been noted by Swift. Neither of these curves, however, is to be taken as exactly indicative of the relative rates of learning at different stages of the process. Other forms of plotting the results are possible, some of which would seem to suggest that learning is initially rapid and becomes slower as the process advances. Serious errors have resulted in the interpretation of learning from giving too much significance to the forms of any of these curves. On the whole little evidence exists, so far as ball-tossing is concerned, for the current view that learning is at first much more rapid than later when considerable skill has been attained, this despite the form of the error curves. Error curves are

found to be mathematically so related to curves of average attainment that the former must drop comparatively very rapidly at first while the latter change slowly, and, conversely, they must drop very slowly later while the average attainment curves rise rapidly. This is in ball-tossing. The changes described do not mean corresponding changes in *the learning*. There is good ground for the view that the same thing may be true of all learning curves. In general, it appears, curves of errors and of time (and of excess distance in the maze) will, because of this mathematical relation to attainment curves, always have a sudden initial drop. In cases showing a rapid initial rise of the average attainment curves it is probable that learning is initially more rapid, as has commonly been supposed, but the difference in the initial and the later rates is likely not as great as it has been thought to be. Even in such cases, and ball-tossing does not seem to be one of them, the evidence is questionable—the rapid initial progress may in many cases not be one of learning at all, but one resulting from the method of stating the results. The whole subject needs a careful experimental overhauling and methods of making data from different experiments more comparable need to be worked out. A tentative procedure is suggested in the present paper, and a few applications are given as illustrative, but only roughly, due to the incompleteness of reported data by various authors.

The basis on which our present views of plateaus and of the approaches to physiological zero rest is insecure. It is found that whether or not a plateau appears and the degree of its occurrence, depend much on the type of curve chosen to represent it. If, for instance, it occurs late in the learning process it may be almost wholly obscured in the error or the time curve. This may be one reason why error and time curves so seldom show plateaus. When the proper conversion of data is made before comparisons there seems to be little ground for the distinction frequently made between sensory discrimination learning and motor learning in animal behavior. The difference is likely mainly one of the method of calling out the reaction and of stating the score; the learning in both cases can be so stated as to make the two processes quite comparable.