Journal of Applied Psychology
1992, Vol. 77, No. 5, 615-622

0021-9010/92/$3.00

Effect of Overlearning on Retention

James E. Driskell
Florida Maxima Corporation
Winter Park, Florida

Ruth P. Willis
Rollins College

Carolyn Copper
Syracuse University

The effectiveness of overlearning in enhancing performance has been acknowledged by researchers within the training community for years. In spite of this general consensus, the empirical basis for this claim is often not clear. This article presents a meta-analysis of the effects of overlearning on retention. Results indicate that overlearning produces a significant effect on retention of moderate overall magnitude and that the effect of overlearning on retention is moderated by the degree of overlearning, type of task, and length of retention period.

The term *overlearning* refers to the deliberate overtraining of a task past a set criterion. In a typical overlearning study, a task criterion may be set at one errorless trial. Subjects in the control condition practice the task until performance reaches the criterion level. Subjects in the treatment condition practice the task until they reach this level and then receive additional practice trials. For example, if reaching the criterion level took 10 trials, the overlearning manipulation may constitute an additional 5 trials of practice (50% overlearning), an additional 10 trials (100% overlearning), or other degrees of overlearning. Retention is then assessed at some interval after the training session.

Although initial work examining overlearning can be traced to Ebbinghaus (1885/1913), some of the earliest empirical studies to establish the effectiveness of overlearning were performed by Krueger (1929, 1930). Krueger (1930) had subjects perform a maze tracing task to a 100% learning, 50% overlearning, 100% overlearning, or 200% overlearning criterion. Tests for retention given at set intervals after training indicated that the greater the degree of overlearning, the greater the retention.

More recently, Schendel and Hagman (1982) also demonstrated that overlearning is an effective means for enhancing task performance. Schendel and Hagman examined the effects of 100% overlearning on retention of a military procedural task (disassembly and assembly of an M60 machine gun). In this procedure, if a trainee required five trials to achieve the criterion (one errorless disassembly and assembly), overtraining consisted of five further trials. Schendel and Hagman found that the overtrained group made 65% fewer errors than a control group when retested after 8 weeks.

Most researchers have concluded that overlearning is an effective training technique (see Cascio, 1991; Goldstein, 1986; Hagman & Rose, 1983; Wexley & Latham, 1981). Fitts (1965) noted that "The importance of continuing practice beyond the point in time where some . . . criterion is reached cannot be overemphasized" (p. 195). However, the empirical basis for this claim is often not clear. Many texts cite the early work of Kreuger as empirical support for the effectiveness of overlearning; however, Krueger's (1930) data yield an average effect size of $r = .22$, a relatively small effect. Furthermore, it is not clear what type of overlearning is effective: Kreuger (1930) reported a significant enhancement of retention with 50% overlearning, Schendel and Hagman (1982) reported significant results with 100% overlearning, and so on. Therefore, although the general consensus of researchers is that overlearning is effective in enhancing retention, it is difficult if not impossible to establish from a narrative review of this literature the precise nature of this effect (i.e., how strong the effect is, what degree of overlearning is required to enhance performance, etc.). One reason for this is that different studies have examined different overlearning manipulations (50%, 100%, 150%), have used different retention periods, and have reported different study statistics.

We conducted a meta-analysis of the research literature on overlearning and retention. A meta-analytic integration of the overlearning literature accomplishes two primary objectives. First, it provides a very specific and precise summary of the overall effects within this research domain. Basic meta-analytic combinations of significance levels and effect sizes provide a gauge of the overall combined probability and strength of the effect of overlearning on retention. Therefore, one goal of our analysis was to establish the overall magnitude of the effect of overlearning on retention.

A second goal of the analysis was to examine the extent to which the effect of overlearning on retention increases or decreases as a function of some theoretically relevant and practically important moderators. This strategy allowed us to examine factors that may moderate the effect of overlearning on retention, such as the degree of the overlearning manipulation. By examining these relationships at the meta-analytic level, we hoped to address several questions of considerable practical interest, such as how much overlearning is required to produce a significant effect on retention, and how long the beneficial

Support for this research was provided by the Office of Basic Research, Army Research Institute.

We thank Brian Mullen for his valuable collaboration.

Correspondence concerning this article should be addressed to James E. Driskell, Florida Maxima Corporation, 147 E. Lyman Ave., Winter Park, Florida 32789.

effects of overlearning are retained over time. In the following section, we describe three factors that may moderate the relationship between overlearning and retention: the degree of overlearning, the retention interval, and the type of task.

Effects of Moderators

Degree of Overlearning

The magnitude of the overlearning manipulation was expected to predict the magnitude of the effect of overlearning on retention. Common sense dictates that more overlearning should certainly be better than less overlearning. However, as Judd and Glaser (1969) observed: "While it is known that overlearning increases retention, the amount of overlearning to be provided has always been a relatively arbitrary decision" (p. 29). Ausubel (1968) noted that overlearning requires that an "adequate number" of repetitions be presented (p. 159). Therefore, the interesting question is, At what point does overlearning begin to enhance retention? For example, is 50%, 100%, or 150% overlearning required before this procedure becomes effective?

To address this question, we first operationalized the degree of overlearning (*DOV*) in the following way:

$$DOV = \text{percentage learning in higher condition} \div (\text{percentage learning in higher condition} + \text{percentage learning in lower condition}).$$

Consider a study that compares the effects of 0% overlearning with 100% overlearning on retention. Zero percent overlearning is actually 100% learning (i.e., a 0% overlearning condition is one in which subjects learn no further beyond the criterion level; thus, this is simply 100% learning of the task). Similarly, 100% overlearning is equivalent to 200% learning. Therefore, the index of overlearning for this example is as follows: $DOV = 200/(200 + 100)$, or .667. *DOV* approaches a lower limit of .500 when there is little difference in magnitude between the manipulations of overlearning in a particular hypothesis test. For example, a study comparing 0% overlearning with 5% overlearning would produce a *DOV* value of .512 (105/[105 + 100]). *DOV* approaches 1 when the difference between conditions becomes infinitely large. By quantifying the degree of overlearning for each hypothesis test included in the analysis, we were able to examine the effect of the degree of overlearning on retention.

Retention Interval

A second predictor to be examined was the retention interval: the number of days between the overlearning manipulation and the test for retention. Again, Ebbinghaus (1885/1913) was one of the first to examine systematically the effects of retention over time. Generally, the longer the delay between acquisition and retrieval, the greater the opportunity for interference and forgetting. In general, long retention intervals should certainly produce worse performance than short retention intervals. The interesting question is, At what point are the benefits of overlearning reduced to zero by the length of the retention interval? In other words, do the effects of overlearning dissipate after a few days, several weeks, or several months? By extracting the retention interval for each of the hypothesis tests included in the meta-analysis, we were able to examine the degree to which the effectiveness of overlearning is moderated by the retention interval—or, in more concrete terms, over what period overlearning is effective.

Type of Task

Studies have examined the effects of overlearning on tasks as dissimilar as balancing on a stabilometer (Melnick, 1971) and remembering verbal information (Gilbert, 1957). Although researchers have suggested that the type of task may moderate the effect of overlearning (Hagman & Rose, 1983; Melnick, 1971), the nature of this relationship is not clear. For example, Melnick noted that, whereas the retention of verbal material seems to be enhanced by overlearning, evidence is inconclusive regarding motor skills. To examine the effect of task type on overlearning, we asked two independent raters to judge each hypothesis test as to whether the task was primarily physical or cognitive in nature. These two judges' ratings reached perfect agreement.

Method

In accordance with the procedures specified by Cooper (1982) and Mullen and Rosenthal (1985), we conducted an exhaustive search of the literature to locate relevant studies, using the ancestry approach, the descendency approach, the invisible college approach, and key word searches (for example, overlearning) of computerized databases such as the PsycINFO Database, Dissertation Abstracts International, and the National Technical Information Service (NTIS). We also manually searched the reference lists of relevant articles and books, and we searched through major psychological journals and association proceedings. Studies were selected for inclusion in this meta-analysis if they reported (or allowed the retrieval of) a test of the effect of overlearning on retention.

Studies that did not meet this criterion and were omitted included those in which basic statistical information was not retrievable (Wasim, 1974), those in which there was no difference in the degree of overlearning (e.g., Melnick, Lersten, & Lockhart [1972] compared the performance of fast and slow learners who both received 100% overlearning), and those in which there was no set learning criterion beyond which overlearning was implemented (e.g., Castaneda & Palermo [1955] examined the amount or degree of training on performance). With these criteria, we located a total of 15 studies with 88 separate hypothesis tests, representing the behavior of 3,771 subjects.

In addition to the basic statistical information (statistical test of the hypothesis, corresponding degrees of freedom, sample size, and direction of effect), each hypothesis test was coded for the predictors mentioned earlier: *DOV* (i.e., the percentage of learning in a higher condition divided by the sum of the percentage of learning in the higher and lower conditions), retention interval (i.e., number of days between the last learning or practice trial and the test for recall), and task type (physical = 1, cognitive = 0). The hypothesis tests included in this meta-analysis are presented in Table 1.

In the analyses, hypothesis tests were subjected to standard meta-analytic procedures (see Rosenthal, 1991). Combination of significance levels and combination of effect sizes were used to gauge the combined probability and strength (respectively) of the effect of overlearning on retention. Diffuse comparisons of significance levels and effect sizes were used to gauge the heterogeneity of study outcomes. Focused comparisons of effect sizes were used to determine whether effects varied in a predictable way as a function of theoretically relevant predictors.

Table 1
Studies Included in the Overlearning Meta-Analysis

Study	Statistic	N	Effect size[a] (d)	Type of task[b]	DOV	Retention interval (no. of days)
Ausubel & Youssef (1965)	$t(85) = 3.85$	87	0.84	0	0.67	2
Ausubel, Stager, & Gaite (1968)	$F(1, 152) = 49.9$	156	1.15	0	0.67	2
Chambers (1969)	$F(1, 48) = 4.56$	56	0.62	0	0.55	7
Craig, Sternthal, & Olshan (1972)	$t(168) = 0.17$	30	−0.02	0	0.67	0
	$t(168) = 0.45$	30	−0.07	0	0.60	6
	$t(168) = 0.62$	30	−0.09	0	0.75	0
	$t(168) = 3.68$	30	0.57	0	0.67	1
	$t(168) = 1.84$	30	−0.28	0	0.60	1
	$t(168) = 1.84$	30	0.28	0	0.75	1
	$t(168) = 0.97$	30	0.15	0	0.67	7
	$t(168) = 1.66$	30	0.26	0	0.60	7
	$t(168) = 2.63$	30	0.40	0	0.75	7
	$t(168) = 2.55$	30	0.39	0	0.67	28
	$t(168) = 1.31$	30	−0.12	0	0.60	28
	$t(168) = 1.24$	30	0.19	0	0.75	28
Earhard, Fried, & Carlson (1972)	$F(1, 204) = 105.6$	216	1.44	0	0.74	0
Gilbert (1957)	$t(18) = 6.32$	18	2.98	0	0.67	0.01
	$t(18) = 10.22$	18	4.82	0	0.75	0.01
	$t(18) = 3.90$	18	1.84	0	0.60	0.01
	$t(18) = 2.92$	18	1.38	0	0.67	1
	$t(18) = 8.28$	18	3.90	0	0.75	1
	$t(18) = 5.36$	18	2.53	0	0.60	1
	$t(18) = 1.94$	18	0.91	0	0.67	2
	$t(18) = 5.94$	18	2.80	0	0.75	2
	$t(18) = 4.00$	18	1.88	0	0.60	2
Juola & Hergenhahn (1967)	$F(1, 54) = 13.43$	60	0.99	0	0.67	0
Krueger (1929)	$t(57) = 1.84$	40	0.49	0	0.60	1
	$t(57) = 3.34$	40	0.88	0	0.67	1
	$t(57) = 1.5$	40	0.40	0	0.57	1
	$t(57) = 2.73$	40	0.72	0	0.60	2
	$t(57) = 4.32$	40	1.14	0	0.67	2
	$t(57) = 1.59$	40	0.42	0	0.57	2
	$t(57) = 2.74$	40	0.72	0	0.60	4
	$t(57) = 4.94$	40	1.31	0	0.67	4
	$t(57) = 2.21$	40	0.58	0	0.57	4
	$t(57) = 3.05$	40	0.81	0	0.60	7
	$t(57) = 4.02$	40	1.06	0	0.67	7
	$t(57) = 0.97$	40	0.26	0	0.57	7
	$t(57) = 2.53$	40	0.67	0	0.60	14
	$t(57) = 3.80$	40	1.01	0	0.67	14
	$t(57) = 1.27$	40	0.34	0	0.57	14
	$t(57) = 2.04$	40	0.54	0	0.60	28
	$t(57) = 3.27$	40	0.87	0	0.67	28
	$t(57) = 1.22$	40	0.32	0	0.57	28
Krueger (1930)	$t(93) = 2.27$	64	0.47	0	0.60	1
	$t(93) = 3.73$	64	0.77	1	0.67	1
	$t(93) = 1.47$	64	0.30	1	0.57	1
	$t(93) = 1.98$	64	0.41	1	0.60	2
	$t(93) = 3.40$	64	0.70	1	0.67	2
	$t(93) = 1.41$	64	0.29	1	0.57	2
	$t(93) = 1.34$	64	0.28	1	0.60	3
	$t(93) = 3.38$	64	0.70	1	0.67	3
	$t(93) = 2.04$	64	0.42	1	0.57	3
	$t(93) = 1.20$	64	0.25	1	0.60	4
	$t(93) = 3.38$	64	0.70	1	0.67	4
	$t(93) = 2.18$	64	0.45	1	0.57	4
	$t(93) = 1.05$	64	0.22	1	0.60	7
	$t(93) = 2.97$	64	0.62	1	0.67	7
	$t(93) = 1.92$	64	0.40	1	0.57	7
	$t(93) = 0.62$	64	0.13	1	0.60	14

Table 1 (*continued*)

Study	Statistic	N	Effect size[a] (d)	Type of task[b]	DOV	Retention interval (no. of days)
Krueger (1930)	$t(93) = 2.34$	64	0.48	1	0.67	14
(*continued*)	$t(93) = 1.72$	64	0.36	1	0.57	14
Langer & Imber (1979)	$t(117) = 2.73$	28	0.50	0	0.75	0
Melnick (1971)	$t(72) = 4.45$	20	1.05	1	0.60	7
	$t(72) = 3.47$	20	0.82	1	0.67	7
	$t(72) = 5.96$	20	1.40	1	0.75	7
	$t(72) = 0.97$	20	−0.23	1	0.57	7
	$t(72) = 1.51$	20	0.36	1	0.67	7
	$t(72) = 2.49$	20	0.59	1	0.60	7
	$t(72) = 5.51$	20	1.30	1	0.60	28
	$t(72) = 6.37$	20	1.50	1	0.67	28
	$t(72) = 8.65$	20	2.04	1	0.75	28
	$t(72) = 0.86$	20	0.20	1	0.57	28
	$t(72) = 3.14$	20	0.74	1	0.67	28
	$t(72) = 2.29$	20	0.54	1	0.60	28
Molander & Garvill (1979)	$t(27) = 1.23$	20	−0.47	1	0.55	0
	$t(27) = 0.73$	20	0.28	1	0.54	0
	$t(27) = 0.55$	20	−0.21	1	0.60	0
	$t(27) = 2.06$	20	−0.79	1	0.55	0
	$t(27) = 0.00$	20	0.00	1	0.54	0
	$t(27) = 2.05$	20	−0.79	1	0.60	0
Postman (1962)	$t(90) = 0.52$	64	0.11	0	0.60	7
	$t(90) = 2.71$	64	0.57	0	0.67	7
	$t(90) = 2.19$	64	0.46	0	0.57	7
Richardson (1973)	$t(276) = 2.45$	66	0.29	0	0.64	0
	$t(276) = 2.19$	66	0.26	0	0.61	0
	$t(276) = 4.64$	64	0.56	0	0.70	0
Schendel & Hagman (1982)	$t(33) = 2.20$	24	0.76	1	0.67	56

Note. DOV = index of overlearning.
[a] Positive effect sizes indicate a positive effect of overlearning on retention. [b] Physical = 1, cognitive = 0.

Formulae and computational procedures for these meta-analytic techniques were presented by Mullen (1989), Mullen and Rosenthal (1985), and Rosenthal (1991).

Results

General Effects

Table 2 presents the results of the combinations and diffuse comparisons of significance levels and effect sizes for the 88 hypothesis tests included in the meta-analytic database, in which each study was weighted by its sample size. The combined effects of these hypothesis tests were of moderate magnitude ($Z = 0.307$, $r = .298$, $d = .625$) and significant ($z = 21.782$, $p < .0001$). Thus, overlearning is shown to produce an overall moderate improvement in retention.

Note that most articles listed in Table 1 contributed multiple effect sizes (e.g., Craig, Sternthal, & Olshan, 1972, provided 12 hypothesis tests). We treated each hypothesis test as an independent observation—an assumption of independence that is patently false and inflates the significance levels of the combined probability tests. One alternative was to pool the results within each study into a single hypothesis test; however, this approach would have sacrificed valuable information regarding the effect of differences in the degree of overlearning or in the retention interval. Although this type of violation has no effect on the mean r or mean d indices of effect size, appropriate caution should be applied in interpreting combined probability and chi-square values.

Type of Task

Separate analyses were conducted to examine the effect of overlearning on physical tasks and on cognitive tasks. For the 37 hypothesis tests involving physical tasks, the effect of overlearning on retention was of small to moderate magnitude and significant ($Z = .22$, $r = .216$, $d = .443$, $p < .0001$). For the 51 hypothesis tests involving cognitive tasks, the effect of overlearning on retention was somewhat stronger ($Z = .368$, $r = .352$, $d = .753$, $p < .0001$). The corresponding focused comparison indicated that this difference between physical and cognitive tasks was significant ($z = 4.089$, $p < .0001$). Overlearning seems to be effective for both physical and cognitive tasks, although the effect of overlearning was somewhat stronger for cognitive tasks.

Degree of Overlearning

In this and the following analyses, we used Fisher's r-to-z transformation to represent the effect size. For these compu-

Table 2
Results of Combinations and Diffuse Comparisons of Effect Sizes and Significance Levels

Statistic	Value
Combinations	
Effect sizes	
Z	.308
r	.298
d	.625
Significance levels	
z	21.78*
Diffuse comparisons	
Effect sizes	
χ^2	230.631*
df	87
Significance levels	
χ^2	338.490*
df	87

Note. $k = 88$, where k denotes the number of hypothesis tests.
* $p < .0001$.

tations, we believed that Fisher's Z was preferable to r because r becomes nonlinear at its extreme values (see James, Demaree, & Mulaik, 1986, though Hunter & Schmidt, 1990, pp. 213–217, challenged these conclusions), and we preferred Z to d because d may take different values depending on the procedure chosen to derive d, among other reasons (see Rosenthal, 1984). Of course, the same relationship may be expressed by Z, r, or d, and each may be algebraically transformed into the other.

We found a significant relationship between the degree of overlearning and the effect size ($r = .477$, $z = 7.537$, $p < .0001$). Thus, the greater the degree of overlearning, the greater the effect of overlearning on retention. Using the regression formula $Z = -1.409 + 2.719(DOV)$, we gauged the magnitude of the effect of overlearning on retention that would result from varying degrees of overlearning. These results are shown in Table 3.

An overlearning manipulation of 50% produced a small overall effect on retention ($Z = .222$), using Cohen's (1977) guidelines for effect size. This suggests that a 50% overlearning manipulation should be considered as a minimum practical operationalization of overlearning and that small improvements in performance can be expected from this level of training. Table 3 also indicates that an overlearning manipulation of 100% resulted in a moderate effect on retention ($Z = .413$) and that an overlearning manipulation of 150% produced a large overall effect ($Z = .520$).

Further analyses were conducted to assess whether the predictive power of the magnitude of overlearning was moderated by the type of task; however, the interaction between the degree of overlearning and task type did not attain significance, $z = 1.223$, $p > .1$.

Retention Interval

There was no significant overall relationship between the retention interval and the effect size ($r = -.0021$, $z = 0.162$, $p > .1$). Thus, surprisingly, we found no overall relationship between the length of the retention interval and the strength of the overlearning effect. However, the interaction between retention interval and task type was significant ($z = 4.619$, $p < .0001$).

Because the impact of the retention period on overlearning was moderated by the type of task, we conducted separate analyses for physical and cognitive tasks. For the 37 hypothesis tests involving physical tasks, there was a significant positive relationship between the length of the retention period and the effect size ($r = .465$, $z = 3.372$, $p < .001$). Thus, for physical tasks, as the time interval between the overlearning practice sessions and subsequent performance increased, retention was enhanced. It is likely that this result can be attributed to some odd factor related to the overlearning manipulation for physical tasks. Subjects were typically given a set number of overlearning trials and then were told not to practice the target task between the overlearning session and the subsequent testing period. However, once primed, it may have been difficult for subjects not to practice a task, especially a relatively simple physical task, given an upcoming testing session. Furthermore, it may have been easier for subjects to "cheat" and to practice these behaviors with physical tasks than with cognitive tasks. For example, after initially practicing a balancing task, it may have been easier for subjects to inadvertently or otherwise practice balancing between the last practice session and the time they were going to be tested; conversely, it may have been more difficult for subjects to practice a cognitive task, such as verbal recall of unfamiliar consonants. If the subjects did practice physical tasks between the practice and test periods, a positive relationship would be expected between the retention interval and overlearning. That this relationship was indeed positive suggests that subjects learning physical tasks may indeed have been getting in this extra practice.

For the 51 hypothesis tests involving cognitive tasks, there was a significant negative relationship between the length of the retention period and the effect size ($r = -.279$, $z = 4.045$, $p < .0001$). For cognitive tasks, longer retention intervals weakened the effects of overlearning. Therefore, although there was an overall positive effect of overlearning on cognitive tasks ($Z = .368$), this effect became weaker as the interval between practice and performance lengthened. Using the regression formula $Z = .462 - .012(RI)$, where RI stands for retention interval, we gauged the precise degree to which overlearning was degraded by a given retention interval. The results are shown in Table 4.

Table 3
Average Effect Sizes Corresponding to 50%, 100%, 150%, and 200% Overlearning

DOV	% overlearning	*Z*
.500	0	.000
.600	50	.222
.670	100	.413
.710	150	.520
.750	200	.630

Note. *DOV* = index of overlearning.

As expected, the strongest effect of overlearning ($Z = .462$) was obtained with the shortest interval period (0 days, or when performance was tested immediately after overlearning). The initial beneficial effect of overlearning was decreased by half (.462/2 = .23) if the retention period was extended to 19 days. We refer to this as the half-life of the overlearning effect, representing the point at which the increase in retention due to the overlearning treatment is reduced by half. This represents a practical guideline for refresher training: If overlearning is to be used as a training procedure to ensure maximum performance effectiveness, refresher training should be implemented at approximate 3-week intervals.

Finally, note the retention period that predicts a zero effect of overlearning (38 days). This indicates that the benefits of overlearning do have an identifiable lifespan: After approximately 38 days, the enhanced retention due to overlearning had dissipated to zero. Therefore, these studies indicate that any benefit provided by overlearning is likely to disappear after 5 to 6 weeks.

Discussion

One question that organizations and training practitioners face is the question of how much training is enough. One approach is to provide the level of training that is estimated to meet the requirement of the average trainee. That is, if the average trainee can achieve proficiency in 5 sessions, then this becomes the level of training provided to all. A second approach is to provide training to meet some set criterion, such as one errorless performance of the task. In this case, each individual is trained until he or she reaches the criterion level. A third approach, overlearning, requires that training continue for a period past this initial mastery level. The results of this meta-analysis document the effectiveness of overlearning and show that retention is enhanced when learning proceeds beyond initial mastery.

Given the efficacy of this approach, what does overlearning do? In other words, why is overlearning effective? First, it provides more training than that required for initial proficiency. So, in one sense, the increased repetition represents a greater degree of learning. However, and perhaps more important, continued practice past initial proficiency allows further feedback to be received on the correctness of response. If a training criterion is set at one errorless trial, the trainee is likely to make progressively fewer errors until this criterion is reached. In this training-to-proficiency approach, training ceases when initial mastery is attained. However, training beyond initial proficiency (i.e., overlearning) allows the trainee to repeat the performance to establish and confirm the correctness of response. In fact, Fitts (1965) noted the importance of training beyond that point at which the correct pattern of behavior is fixed. Therefore, a well-designed overlearning intervention may incorporate the basic principles of learning—repetition, feedback, and contiguity—and, more important, it allows practice of the established behavior beyond initial mastery.

Furthermore, overlearning may provide one approach to the problem of individual differences in ability among trainees. Vineberg (1975) found that differences in the ability level of trainees affected the level of initial skill acquisition but had little effect on the rate of retention. Although high-ability trainees learned more, decay in skill retention occurred at a similar rate for both high- and low-ability trainees. Therefore, the strategy of training each individual to proficiency and then providing overtraining may ensure that all trainees acquire initial skill mastery and retain this information for an optimal period. This approach to managing individual differences in ability may be particularly important for team training, in that integrated team performance requires that each team member retain a high level of proficiency to support overall team performance.

One unexpected finding of this analysis was a positive correlation for physical tasks between the length of the retention interval and the effect of overlearning on retention. Surprisingly, the longer the interval between overlearning and testing, the better the retention. We argue that, for physical tasks, subjects may have been cheating by practicing these skills after the training session. From a practical standpoint, this cheating has positive effects: It results in enhanced performance. Thus, one question becomes "How do we promote this practice?" This cheating, or self-directed practice, may stem from the fact that, after the training session, subjects knew that within a relatively short period (0–28 days), their performance would be evaluated. If this knowledge was indeed a factor, this finding suggests that training designers should consider scheduling post-training assessment tests as a means of promoting practice. Furthermore, practice may have occurred for physical and not cognitive tasks because the cognitive tasks may have been more difficult to practice independently. If this was the case, this further suggests that, whenever possible, training should emphasize the actual physical performance of a skill. For example, training for a cognitive skill (such as effective communication) should include both concept learning and active practice. Although this is simply good instructional strategy, the current results suggest that it may also allow the skill to be more easily practiced after initial skill acquisition.

There are, of course, limitations to our analysis and to the implications for training that can be drawn from it. First, several different meta-analytic approaches can be found in the scientific literature. The meta-analytic approach taken in this study (see Rosenthal, 1991) differs in a number of ways from that offered by Hunter and Schmidt (1990) or by Hedges and Olkin (1985). For example, Hunter and Schmidt (1990) proposed an elaborate set of adjustments for measurement error that Rosenthal (1991) in general did not advocate. In fact, Rosenthal argued that, in many cases, the goal of a meta-analysis is

Table 4
Effect of Retention Interval on Magnitude of Effect for Cognitive Tasks

Retention interval (no. of days)	Magnitude of effect (Z)
0	.462
2	.438
8	.368[a]
11	.333
19	.231
38	.000

[a] Mean effect size.

to establish the relationship between variables as these patterns exist in the extant literature, rather than to estimate the relationship that might be found if the studies had been done with perfect measurement. In this particular analysis, the included studies did not provide the information needed to make corrections for measurement error. Thus, to the extent that measurement error exists in the dependent variable, estimates of the effect of overlearning may be conservative. In brief, there are advantages and disadvantages to each of the generally accepted meta-analytic approaches. Furthermore, relationships found at the meta-analytic level should always be examined at the primary level of analysis, at which the influence of unwanted sources of variation can be controlled.

This analysis allowed us to test several hypotheses regarding moderators of overlearning's effect on retention. These moderators (degree of overlearning, retention interval, and task type) were chosen because they were theoretically interesting (i.e., past research suggested their relationship to overlearning) and because the available empirical literature allowed their examination (i.e., the information presented in the studies allowed this variable to be coded or rated). However, there are other potentially informative factors that we were not able to examine. For example, motivation certainly plays a role in training effectiveness: It is likely that subjects will be more motivated to learn in studies that use relevant real-world tasks (i.e., tasks that are related to their jobs) than in studies that use laboratory tasks. If this is the case, then the effect of overlearning on retention ought to be greater for real-world tasks than for laboratory tasks. However, only one study in this database was performed in a real-world setting with a real-world task, so we were not able to make any meaningful comparison among studies on this factor. This limitation suggests two points. First, if the unsubstantiated hypothesis regarding the effects of motivation on overlearning is tentatively accepted, then the fact that overlearning was indeed effective in this pool of largely laboratory studies suggests its overall robustness. Furthermore, a meta-analysis often serves to point out what researchers don't know: The fact that we were not able to examine at the meta-analytic level the effect of such factors as motivation on overlearning suggests areas that require further study.

Given the overall positive effects of overlearning on retention, what are the drawbacks to this technique? A potential disadvantage is the cost of overlearning. Although our analysis documents the effectiveness of this approach, training beyond initial proficiency requires increased resources. This increased cost, however, buys enhanced retention over time. This benefit may be particularly important for those tasks, such as emergency procedures, that are not likely to be practiced daily and in which first-trial performance is critical. Furthermore, the increased cost of overlearning may be illusory: Schendel and Hagman (1982) found that, after an 8-week retention interval, overtrained subjects required 22% fewer trials to retrain to the criterion level than did controls. Therefore, the increased initial costs associated with overlearning may be partially offset by lowered costs for subsequent retraining or refresher training.

A second caveat relates to the fidelity of the training task. It may seem an obvious point, but it is critical that the task trained reflect the task that is required in the actual performance setting. Thus, Weitz and Adler (1973) cautioned against the use of overlearning in simulations in which the training environment differs significantly from the actual task. To the extent that the simulation contains features unique to that environment and inappropriate for the actual performance setting, overlearning the training task may actually degrade performance. This point is not as commonsense as it may seem: It suggests that training, and overlearning in particular, must consider the environmental parameters in which the actual performance will take place. In fact, Zakay and Wooler (1984) found that training conducted under normal conditions improved decision performance under normal conditions, but did not improve performance when subjects performed under time pressure. In summary, although overlearning will improve retention, the training designer must ensure that the task that is overlearned is the task called for in the actual performance setting.

Finally, why is there so little recent research on overlearning? The most recent study in our database was conducted in 1982. This is likely a consequence of the fact that, although there is near-unanimous agreement on the value of overlearning, there have been few specific, prescriptive guidelines for training designers to use to implement this approach. Therefore, one goal of this study was to provide some practical guidelines for implementing overlearning, which we summarize in the following paragraphs.

First, we found that overlearning is an effective means of enhancing retention, as evidenced by a significant overall effect of moderate magnitude. Second, overlearning is an effective training procedure for both physical and cognitive tasks; however, the effect of overlearning is stronger for cognitive tasks. Third, the greater the degree of overlearning, the greater the resulting retention. For practical purposes, both 100% overlearning and 150% overlearning produced moderate to strong effects on performance. Fourth, for cognitive tasks, the longer the delay between the overlearning manipulation and performance, the weaker the overlearning effect. The benefit in performance gained from overlearning was reduced by one-half after 19 days: Thus, maintaining optimal performance requires that additional training take place after approximately 3 weeks. Without further refresher training, the increase in retention due to overlearning is likely to dissipate to zero after 5 to 6 weeks. Results also indicate that this decay is less likely for physical tasks, and thus there may be considerable practical value in ensuring active, physical practice of skills during training.

References

Ausubel, D. P. (1968). *Educational psychology: A cognitive view.* New York: Holt.

Ausubel, D. P., Stager, M., & Gaite, J. H. (1968). Retroactive facilitation in meaningful verbal learning. *Journal of Educational Psychology, 59,* 250–255.

Ausubel, D. P., & Youssef, M. (1965). The effect of spaced repetition on meaningful retention. *Journal of General Psychology, 73,* 147–150.

Cascio, W. F. (1991). *Applied psychology in personnel management.* Englewood Cliffs, NJ: Prentice Hall.

Castaneda, A., & Palermo, D. S. (1955). Psychomotor performance as a function of amount of training and stress. *Journal of Experimental Psychology, 50,* 175–179.

Chambers, D. W. (1969). *A reevaluation of the discovery learning hypoth-*

esis in terms of overlearning. Unpublished doctoral dissertation, Stanford University.
Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
Cooper, H. M. (1982). Scientific guidelines for conducting integrative research reviews. *Review of Educational Research, 52,* 291–302.
Craig, C. S., Sternthal, B., & Olshan, K. (1972). The effect of overlearning on retention. *Journal of General Psychology, 87,* 86–94.
Earhard, B., Fried, C., & Carlson, G. (1972). Interference, overlearning, and anticipation time. *Journal of Educational Psychology, 94,* 345–347.
Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussenius, Trans.). New York: Columbia University, New York Teacher's College. (Original work published 1885)
Fitts, P. M. (1965). Factors in complex skill training. In R. Glaser (Ed.), *Training research and education* (pp. 177–197). New York: Wiley.
Gilbert, T. (1957). Overlearning and the retention of meaningful prose. *Journal of General Psychology, 56,* 281–289.
Goldstein, I. L. (1986). *Training in organizations: Needs assessment, development, and evaluation* (2nd ed). Monterey, CA: Brooks/Cole.
Hagman, J. D., & Rose, A. M. (1983). Retention of military tasks: A review. *Human Factors, 25,* 199–213.
Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* San Diego, CA: Academic Press.
Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis.* Newbury Park, CA: Sage.
James, L. R., Demaree, R. G., & Mulaik, S. A. (1986). A note on validity generalization procedures. *Journal of Applied Psychology, 71,* 440–450.
Judd, W. A., & Glaser, R. (1969). Response latency as a function of training method, information level, acquisition, and overlearning. *Journal of Educational Psychology, 60,* 1–30.
Juola, J. F., & Hergenhahn, B. R. (1967). Probability matching and the overlearning reversal effect. *Psychonomic Science, 8,* 309–310.
Krueger, W. C. F. (1929). The effect of overlearning on retention. *Journal of Experimental Psychology, 12,* 71–78.
Krueger, W. C. F. (1930). Further studies in overlearning. *Journal of Experimental Psychology, 13,* 152–163.
Langer, E. J., & Imber, L. G. (1979). When practice makes imperfect: Debilitating effects of overlearning. *Journal of Personality and Social Psychology, 37,* 2014–2024.
Melnick, M. (1971). Effects of overlearning on the retention of a gross motor skill. *Research Quarterly, 42,* 60–69.
Melnick, M. J., Lersten, K. C., & Lockhart, A. S. (1972). Retention of fast and slow learners following overlearning of a gross motor skill. *Journal of Motor Behavior, 4,* 187–193.
Molander, B., & Garvill, J. (1979). The invariance of asymmetric cross-modal transfer effects. *Scandinavian Journal of Psychology, 20,* 171–177.
Mullen, B. (1989). Advanced BASIC meta-analysis. Hillsdale, NJ: Erlbaum.
Mullen, B., & Rosenthal, R. (1985). *BASIC meta-analysis: Procedures and programs.* Hillsdale, NJ: Erlbaum.
Postman, L. (1962). Retention as a function of degree of overlearning. *Science, 135,* 666–667.
Richardson, J. (1973). Effect of speed of learning and degree of learning on cue selection. *Journal of Experimental Psychology, 98,* 396–403.
Rosenthal, R. (1984). *Meta-analytic procedures for social research.* Beverly Hills, CA: Sage.
Rosenthal, R. (1991). *Meta-analytic procedures for social research* (2nd ed.). Beverly Hills, CA.: Sage.
Schendel, J. D., & Hagman, J. D. (1982). On sustaining procedural skills over a prolonged retention interval. *Journal of Applied Psychology, 67,* 605–610.
Vineberg, R. (1975). *A study of the retention of skills and knowledge acquired in basic training* (Tech. Rep. No. 75–10). Alexandria, VA: Human Resources Research Organization.
Wasim, M. (1974). Effect of repeated learning and overlearning on recall. *Pakistan Journal of Psychology, 7,* 59–61.
Weitz, J., & Adler, S. (1973). The optimal use of simulation. *Journal of Applied Psychology, 58,* 219–224.
Wexley, K. N., & Latham, G. P. (1981). *Developing and training human resources in organizations.* Glenview, IL: Scott, Foresman.
Zakay, D., & Wooler, S. (1984). Time pressure, training, and decision effectiveness. *Ergonomics, 27,* 273–284.

Received October 21, 1991
Revision received March 9, 1992
Accepted March 10, 1992 ■