School Psychology Review,
2008, Volume 37, No. 3, pp. 309–314

COMMENTARY

Theoretical and Applied Implications of Precisely Measuring Learning Rates

Christopher H. Skinner
The University of Tennessee

Nist and Joseph (2008) have confirmed earlier research showing that adding and interspersing a large number of time-consuming learning trials targeting known items (e.g., incremental rehearsal [IR] or interspersal) retards student learning rates (Cates et al., 2003 Joseph & Nist, 2006; Schmidgall & Joseph, 2007). In addition, their current study has confirmed earlier research that adding and interspersing known items can improve the quality of learning trials, or the amount learned per unknown-word trials (Burns, Dean, & Foley, 2004; MacQuarrie-Klender, Tucker, Burns, & Hartman, 2002).

Learning Levels Versus Learning Rates

Learning is often described as a relatively permanent change in behavior or behavioral potential brought about by experience, with some experiences (e.g., an accident causing severe brain damage) being excluded. Across many applied and theoretical experiments, the experience is the instructional activity (e.g., intervention, remedial procedure). In the Nist and Joseph study, the learning experience was held constant across all conditions is the flashcard drill procedure. Across all conditions, each session included six trials, across six unknown words, repeated nine times per session. However, these flashcard trials were not the independent variable or the experience under investigation. Rather, the independent variable was the presence of additional flashcard trials targeting known words. These known-word trials were placed around the unknown-word trials using either IR (nine known-word trials placed among six unknown-word trials using a folding-in strategy) or interspersal (three known-word trials evenly distributed among six unknown-word trials). In the control condition, no known-word trials were added.

In many intervention studies, the behavior change (learning) often serves as the primary dependent variable. Nist and Joseph used unknown words read accurately in isolation (acquisition and maintenance) and in sentences (generalization) as their dependent variables and repeated assessments were used to assess behavior change (learning) as intervention procedures were applied over time and across sets of unknown words. Their study has strong applied value because they compared three strategies and assessed changes in learn-

Correspondence regarding this article should be addressed to Christopher H. Skinner, The University of Tennessee, BEC 525, Knoxville, TN 37996-3452; E-mail: cskinne1@utk.edu

ing *levels* by including acquisition, maintenance, and generalization measures. Changes in levels of behavior (e.g., number correct, number maintained, number of words read correctly in sentences) do not take into consideration instructional time required to bring about that change. Nist's and Joseph's study is superior to most because they also measured learning rates. Assessing learning rates requires measuring the change in behavior (level change) while also taking into account the time required for the learning experience. Learning rates = learning level/instructional time (Cates et al., 2003; Skinner, Belfiore, Mace, Williams, & Johns, 1997; Skinner, Belfiore, & Watson, 1995/2002; Skinner, Fletcher, & Henington, 1996).

Implications of Learning Levels Results

Over 100 years ago, using himself as a subject, Ebbinghaus (1885) showed that increased practice enhanced learning. Thus, in some instances researchers comparing learning strategies may want to hold practice, often referred to as opportunities to respond (Greenwood, Delquadri, & Hall, 1984) or learning trials (Albers & Greer, 1991), constant. Holding the type (flashcard) and number of unknown-word learning trials constant across the three different learning experiences allowed Nist and Joseph to evaluate the relative quality of the learning experiences when the additional known-word flashcard trials were added (Skinner et al., 1997). Confirming previous research, Nist and Joseph found greater increases in learning *levels* under the IR procedure (Burns et al., 2004; MacQuarrie-Klender et al., 2002).

Nist's and Joseph's findings have theoretical significance as they suggest that something involved with adding these known-word trials and folding them in enhances the quality of the unknown-word trials (amount learned per unknown-word trial). Identifying the mechanism that causes this increase is a clear direction for future researchers. Some plausible hypotheses suggested by previous researchers include distributed practice effects (Burns et al., 2004; MacQuarrie-Klender et al., 2002), increasing rates of reinforcement (Skinner, 2002) and/or enhancing students' sense of success (Neef, Iwata, & Page, 1997, 1980), and pace of responding and attention (Hawkins, Skinner, & Oliver, 2005; Robinson & Skinner, 2002).

Implications of the Learning Rates

Although Nist's and Joseph's data on learning *levels* showed that IR was superior, their learning rate data suggest that of the three procedures used, IR is the last procedure that educators should select if the goal is remediation. Before readers conclude that I am making an insignificant nuanced argument, consider the comfort I take in Nist's and Joseph's social validity data, which suggest that the students and teachers who participated agree with me.

Remedying Deficits

School psychologists attempt to prevent and remedy learning problems. In almost all cases, a learning problem is not a failure to learn, but a failure to learn specific skills or behaviors as *rapidly* as expected (Skinner et al., 1995/2002; Skinner et al., 1996). Thus, learning problems are more precisely defined as learning *rate* problems. In the Nist and Joseph study, the students who participated could read words in sentences. Thus, prior experiences had caused learning. Presumably, they were nominated to participate because their reading skills were not as developed as expected at this moment in time, a learning rate problem.

Remediation frequently occurs when our procedures enhance learning rates and/or when we allot more learning and/or instructional time to remediation targets (Skinner et al., 1996). Often we do both. For example, under many response-to-intervention (RTI) models, school time is reallotted from time devoted to other activities including recess and other classes (e.g., art, music, foreign language instruction, physical education, and health class). When I was in elementary through high school, a similar time reallotment strategy was used when students with

learning rate deficits were enrolled in extended school years (we called it summer school). Time reallotment causes resistance from those concerned that reducing time spent in other activities will decrease the quality of the students' life both now and later. Students may develop social skills during recess, the ability to produce and appreciate art during special classes, and a joy for physical activities and habits that enhances their health during physical education and health class. All may enhance the quality of their lives and their ability to contribute to society.

The counterargument is that some skills are more important than are others. For example, learning to read is critical to lifelong success, because, while we learn to read, often we read to learn. Having made both arguments, I feel there is no correct answer. However, a shift in how we evaluate remedial strategies may eventually allow us to remedy deficits more efficiently, thus avoiding some of these difficult time reallotment decisions. That shift is to evaluate our remedial procedures by precisely measuring learning rates.

Previous Research

One of my first studies published on learning rates (see Skinner et al., 1995/2002) was a reanalysis of earlier data (see Skinner, Johnson, Larkin, Lessley, & Glowacki, 1995). The Skinner et al. (1995) study was similar to the Nist and Joseph study; we compared sight-word reading interventions, one intervention took longer than the other, and our measure of learning levels was designed to evaluate theoretical questions. The reanalysis (Skinner et al., 1995/2002) showed that our learning levels and learning rate data conflicted.

Cates et al. (2003) provided a clear example of how adding known-item trials retarded learning rates. The Nist and Joseph study is an extension of Cates' dissertation; however, Cates targeted spelling, as opposed to sight-word reading. Regardless, the critical finding from Cates et al. (interspersing additional known words retarded learning rates) was supported by Joseph's earlier studies of sight-word reading (Joseph & Nist, 2006; Schmidgall & Joseph, 2007). The current study is an improvement because maintenance and generalization to sentence reading were assessed.

The importance of measuring learning rates seems lost on many researchers and those responsible for what is published in school psychology and other journals. One reason for this is the measure itself. Nist's and Joseph's rate data are converted to words learned per minute of instruction under the three different conditions. Although this measure has clear meaning to some (e.g., Cates et al., 2003), it is uncommon and may be difficult for others to understand as most instructional procedures are delivered in units of time that exceed 1 min. Furthermore, Nist and Joseph did not convert the data on words read correctly in sentences to learning rate measures. To highlight the importance of learning rates, I next describe an applied scenario, extrapolate Nist and Joseph data to measure learning rates for generalization, and interpret these data.

Interpreting an Applied Scenario

Scenario

There are six students in first-grade classes who are in need of remediation because their sight-word reading skills are behind their peers. The remedial instruction is going to be delivered at a 1–1 student-to-teacher ratio using a flashcard procedure during a 20-min period each day. During that time, the students will not be able to participate in recess (time reallocation). Remediation is defined as being able to read 330 unknown words in sentences (generalization). All students enjoy recess; their parents and teachers feel that recess is important for various reasons (e.g., incidental learning of social skills). Thus, the goal is to remedy their deficits as rapidly as possible so that they may return to recess and have more success with learning material that is currently being covered in their general education reading instruction. In addition, because 1–1 instruction requires much teacher time, more rapid remediation will allow these resources to be redirected to remedy other deficits across students.

Table 1
Number of Words Read Correctly, Number of Words Maintained, and Number of Words Read Correctly in Sentences per 100 Minutes of Instruction Time

	Traditional			Interspersing			Incremental Rehearsal		
Student	WC/ 100 M	WM/ 100 M	WS/ 100 M	WC/ 100 M	WM/ 100 M	WS/ 100 M	WC/ 100 M	WM/ 100 M	WS/ 100 M
Sarah	106	89	81	66	60	57	33	30	30
Lori	100	85	66	70	60	51	29	23	21
Marie	114	96	73	67	58	51	34	33	30
Dan	127	99	80	84	70	57	32	29	28
Chris	114	93	88	71	61	54	25	24	22
Beth	111	89	75	63	54	52	34	31	30
Group	113	93	77	71	61	54	31	28	27

Note. WC = number of words read correctly; M = minutes; WM = number of words maintained; WS = number of words read correctly in sentences; Group = the mean for the group of 6 children.

Extrapolated Data

Table 1 shows the extrapolated data from Nist's and Joseph's study including the number of words read correctly, the number of words maintained, and the number of words read correctly in a sentence per 100 min of instructional time (1 week) across the three conditions. Across all students and all dependent variables employed, the results are consistent. When time allotted to learning is held constant, drilling only unknown words (control condition) resulted in the most learning and IR the least, even for the generalization data.

Interpretation

Given our hypothetical remediation goal (330 unknown words read correctly in sentences), all students would be remedied in 5 weeks or less using the control condition. Using the IR condition, remedying the very strongest IR learners (Sara, Marie, and Beth) would require 11 weeks. Relative to IR, the control procedures would enhance students' time in recess (by a minimum of 6 weeks), perhaps allowing them to learn more in their general education class over these same 6 weeks (assuming that learning these words will help them in class) and save considerable resources (1–1 instruction, 2 hr per day, for 6 weeks would cost 60 teacher hr). Who would prefer the IR procedure over the control?

Direction for Future Researchers

As school psychologists become more involved in remedying students' problems (e.g., RTI), they should become keenly aware of the need for studies that compare remedial procedures. Unfortunately, very few researchers measure learning rates with any sort of precision. Thus, our current research base may provide misleading results for those interested in identifying the most effective remedial procedures (see Skinner et al., 1995/2002).

Although measuring learning rates is important, learning is a very complex behavior. When we measure learning rates precisely, I have no doubt that we will find learning is highly unstable within students being exposed to the same instructional procedure. For example, in my scenario the initial trials (e.g., first 5 min of a 20-min learning session) may yield higher learning rates than later trials (e.g., the last 5 min or the session). Thus, generalizing

data from one instructional time interval to other instructional time intervals (e.g., my extrapolation of the Nist and Joseph data) is poor science. This instability in learning rates could be controlled by holding time constant under each condition and letting the trials vary (see Carroll, Skinner, Turner, McCallum, & Masters, 2006; Poncy, Skinner, & Jaspers, 2007; Skinner et al., 1997). Such experiments may show that some strategies are superior to others when instructional time is 10 min per day, but opposite results are found when 20 min per day is available. Regardless, these questions cannot and will not be answered until researchers begin using precise measures of learning rates to evaluate and compare interventions.

Conclusions

Although I have stressed learning rates, the Nist and Joseph finding that adding and folding in unknown items enhanced learning levels is important. If we can determine why IR caused increases in learning levels (not rates), educators and researchers may be able to develop strategies that allow them to strengthen the effect (e.g., cause larger learning level increases) and/or make the application of the procedure more efficient (e.g., take less instructional time). Thus, these findings may lead to the development of strategies that can then be applied to enhance rather than retard learning rates.

The focus on evidence-based practice, empirically validated interventions, and remedying learning problems has brought applied learning research to the forefront. Concomitantly, procedures have been developed that improve the quality of this research (e.g., including measures of treatment integrity). It is my hope that Nist's and Joseph's study is representative of another change in intervention research that will improve the precision of our evaluations by precisely measuring learning rates.

Almost all studies designed to evaluate learning should include measures of learning rates. More important, all applied studies designed to compare remedial procedures should determine which procedure results in the greatest learning rates. Finally, editors and reviewers for journals should stop insisting that all researchers conducting intervention comparison studies hold opportunities to respond or learning trials constant, while allowing instructional time to vary. Because many of the questions we are attempting to answer are basic applied learning questions, as opposed to theoretical questions regarding the quality of response opportunities or learning trials (quality = learning level change per learning trial), these gatekeepers should allow researchers to let opportunities to respond vary, but hold instructional time constant across conditions. Such a change will take us toward the basic human condition (learning requires time) and make our studies more relevant to practitioners.

References

Albers, E. S., & Greer, R. D. (1991). Is the three term contingency trial a predictor of effective instruction? *Journal of Behavioral Education, 1,* 337–354.

Burns, M. K., Dean, J. D., & Foley, S. (2004). Preteaching unknown key words with incremental rehearsal to improve reading fluency comprehension with children identified as reading disabled. *Journal of School Psychology, 42,* 303–314.

Carroll, E., Skinner, C. H., Turner, H., McCallum, E., & Masters, S. (2006). Evaluating and comparing responsiveness to two interventions designed to enhance math-fact fluency. *School Psychology Forum: Research in Practice, 1,* 1–18.

Cates, G. L., Skinner, C. H., Watson, T. W., Meadows, T. J., Weaver, A., & Jackson, B. (2003). Instructional effectiveness and instructional efficiency as considerations for data-based decision making: An evaluation of interspersing procedures. *School Psychology Review, 32,* 601–616.

Ebbinghaus, H. (1885). *Uber das Dedachtinis.* Leipzig: Duncker & Humbolt (Memory; H. A. Ruger & C. E. Bussenius, Trans. New York: Teachers College, 1913; reprinted by Dover, 1964.)

Greenwood, C. R., Delquadri, J., & Hall, R. V. (1984). Opportunity to respond and student academic performance. In W. Heward, T. Heron, D. Hill, & J. Trap-Porter (Eds.), *Behavior analysis in education* (pp. 58–88). Columbus, OH: Charles E. Merrill.

Hawkins, J., Skinner, C. H., & Oliver, R. (2005). The effects of task demands and additive interspersal ratios on fifth-grade students' mathematics accuracy. *School Psychology Review, 34,* 543–555.

Joseph, L. M., & Nist, L. M. (2006). Comparing the effects of unknown-known ratios on word reading learning versus learning rates. *Journal of Behavioral Education, 15,* 69–79.

MacQuarrie-Klender, L. L., Tucker, J. A., Burns, M., & Hartman, B. (2002). Comparison of retention rates

using traditional, drill sandwich, and incremental rehearsal flash card methods. *School Psychology Review, 31,* 584–595.

Neef, N. A., Iwata, B. A., & Page, T. J. (1977). The effects of known-item interspersal on acquisition and retention of spelling and sight-reading words. *Journal of Applied Behavior Analysis, 10,* 738

Neef, N. A., Iwata, B. A., & Page, T. J. (1980). The effects of interspersal training versus high-density reinforcement on spelling acquisition and retention. *Journal of Applied Behavior Analysis, 13,* 153–158.

Nist, L., & Joseph, L. M. (2008). Effectiveness and efficiency of flashcard drill instructional methods on urban first-graders' word recognition, acquisition, maintenance, and generalization. *School Psychology Review,* 37, 294–308.

Poncy, B. C., Skinner, C. H., & Jaspers, K. E. (2007). Evaluating and comparing interventions designed to enhance math fact accuracy and fluency: Cover, copy, and compare versus taped problems. *Journal of Behavioral Education, 16,* 27–37.

Robinson, S. L., & Skinner, C. H. (2002). Interspersing additional briefer, simpler items to enhance mathematics performance on subtests requiring different levels of attention. *School Psychology Quarterly, 2,* 191–205.

Schmidgall, M., & Joseph, L. M. (2007). Comparison of phonic analysis and whole word reading on first graders' cumulative words read and cumulative reading rate: An extension in examining instructional effectiveness and efficiency. *Psychology in the Schools, 44,* 319–332.

Skinner, C. H. (2002). An empirical analysis of interspersal research: Evidence, implications and applications of the discrete task completion hypothesis. *Journal of School Psychology, 40,* 347–368.

Skinner, C. H., Belfiore, P. J., Mace, H. W., Williams, S., & Johns, G. A. (1997). Altering response topography to increase response efficiency and learning rates. *School Psychology Quarterly, 12,* 54–64.

Skinner, C. H., Belfiore, P. B., & Watson, T. S. (1995/2002). Assessing the relative effects of interventions in students with mild disabilities: Assessing instructional time. *Journal of Psychoeducational Assessment,* 20, 345–356 (Reprinted from *Assessment in Rehabilitation and Exceptionality, 2,* 207–220.)

Skinner, C. H., Fletcher, P. A., & Henington, C. (1996). Increasing learning trial rates by increasing student response rates. *School Psychology Quarterly, 11,* 313–325.

Skinner, C. H., Johnson, C. W., Larkin, M. J., Lessley, D. J.,& Glowacki, M. L. (1995). The influence of rate of presentation during taped-words interventions on reading performance. *Journal of Emotional and Behavioral Disorders, 3,* 214–223.

Date Received: May 30, 2008
Date Accepted: June 17, 2008
Action Editor: Thomas Power ■

Christopher H. Skinner, PhD, is Professor and Coordinator of School Psychology Programs at The University of Tennessee. His research interests include prevention and remediation of academic and social problems within educational settings, single-subject design research, and applied behavioral analysis. He received his PhD from Lehigh University in 1989 and served as editor of *Journal of Behavioral Education*.