In Response
A Refocus on Response-Rate Measurement: Comment on Doughty, Chase, and O’Sheilds (2004)

Carl Binder
Binder Riha Associates

The recent review by Doughty, Chase, and O’Sheilds (2004) of studies that have examined effects of response-rate building on learning outcomes is a valuable contribution to research on use of response-rate measures for instructional decision making. It raises important experimental issues that require greater attention in the future. In its conceptual explication and analysis of published articles that meet the authors’ criteria for inclusion, however, the review introduces some definitional slippage and at least one factual error that require clarification for the sake of future analysis and research.

The Great Falls Precision Teaching Project

To dispense quickly with the factual error, the authors describe Beck and Clement’s (1991) report of the Great Falls Precision Teaching Project as “anecdotal evidence” (p. 7). In fact, the Great Falls project was a comparison of the Iowa Test of Basic Skills achievement scores over a 3-year period between the Sacajawea School and other elementary schools in the same district. The 20 to 40 percentile-point improvement (depending on the subtest) achieved at the Sacajawea School was attributed to 20 to 30 min per day of 1-min practices that were charted by teachers and shared with students at Sacajawea, without any other differences in curriculum or teaching methods between Sacajawea and other schools. This, one could argue, was one of the most cost-effective educational improvement demonstrations in the literature. Although it did not control for each feature of the simplified precision teaching procedures it employed, it was a robust empirical demonstration and surely not “anecdotal.”

The Key Issue Is Measurement Sensitivity

At the core of the precision teaching methods and discoveries discussed by Doughty et al. (2004) is a question about measurement: Is rate of response a better predictor of important learning outcomes (e.g., retention, maintenance, application) than the more traditional percentage correct, a dimensionless quantity (Barrett, 2002; Barrett, Johnston, & Pennypacker, 1986; Johnston & Pennypacker, 1980)?

In education, where percentage correct is ubiquitous, we sometimes forget Skinner’s assertion that his most significant contributions were “the use of rate of responding as a basic datum and the so-called cumulative record” (Evans, 1968, p. 103). Vargas (1977) clarified the implications of this statement for instructional applications when she said,

Teaching . . . is not only producing new behavior, it is also changing the likelihood that a student will respond in a certain way. Since we cannot see likelihood, we look instead at how frequently a student does something. We see how fast he can add. The student who does problems

Address correspondence to Carl Binder, Binder Riha Associates, 4966 Wilshire Dr., Santa Rosa, California 95404 (e-mail: CarlBinder@aol.com). For downloadable versions of many publications cited in this article visit www.fluency.org.
correctly at a higher rate is said to know addition facts better than one who does them at a lower rate. (p. 62)

In the instructional context, a measure of response likelihood, strength, or probability refers to the ability of an individual to behave in a particular (skillful) way and not merely to that individual’s inclination to do so. Thus, use of response-rate measures is essential at every step in controlled research that seeks to verify clinical discoveries from precision teaching. A possibly more interesting way to frame rate-building research than as a comparison of the effects of controlled trials with the effects of self-paced practice on retention, application, and other learning outcomes is to ask if rate of freely emitted responding better predicts these learning outcomes, whether that rate is the product of controlled-trials practice, self-paced practice, or some combination. Does it tell us more than percentage correct?

Achieving Normal Competent Rates of Response—Not “High” Rates

In their review, Doughty et al. (2004) frequently use the term high when referring to response rates achieved with precision teaching. This modifier unnecessarily biases the discussion, because it suggests rates that are out of the ordinary, extreme, or unusual. As Barrett (1979) illustrated with behavioral education for developmentally delayed students aimed at helping them perform in so-called normal environments, response-rate measures enable us to detect clear differences in ability among competent adults, children, and disabled learners and lead to better instructional decisions than do percentage correct scores, which do not reveal such differences. Her data (reprinted in Binder, 2003; Johnson & Layng, 1992) show systematic differences in rates of correct responding among the three aforementioned populations that are not reflected in percentage correct scores. With these data in mind, the phrase “high response rates” should be changed in Doughty et al. to “normal ranges of response rate.” There is nothing unusual or extreme about using a more sensitive measure (i.e., response rate) to identify normal ranges of competence and then applying procedures that enable learners to achieve those normal ranges. To the contrary, this is what Barrett called “communication” or behavioral normalization.

Defining Application

When describing key learning outcomes cited in the precision teaching literature, Doughty et al. (2004) redefine application in a way that differs from how the term is used in that literature. When Starlin (1971), Haughton (1972), and their colleagues first determined that students who could correctly perform key component skills (e.g., adding numbers) at or above certain response rates (e.g., 40 to 50 per minute) were successful in more advanced composite skills (e.g., math story problems) while those with lower rates (e.g., 20 to 30 per minute) were not, it dawned on them that combination of smaller units of behavior (components) into larger units (composites) might depend on development of higher component response rates. Clinical measures and observations since the early 1970s have confirmed this result (Binder, 2003), as have a few published experimental studies (e.g., Van Houten, 1980). At one point referred to as compounding (Haughton, 1980), this was the origin of the term application later communicated with Haughton’s acronyms RAPS (retention/application performance standards) and REAPS; these are further discussed by Johnson and Layng (1992, 1996), Binder (1993, 1996), and others.

Despite this documented history, Doughty et al. (2004) redefine application as “generalization, or the occurrence of a skill in novel stimulus conditions” (p. 9). Although generalization might in some instances be a
result of achieving competent rates of responding, a narrower generalization gradient might just as well occur, depending on the specificity of stimulus control during rate building. In any case, Doughty et al.'s definition of application does not correspond to that of the precision teaching and fluency literature, and therefore does not reflect claims made or questions asked among precision teaching practitioners and researchers. Instead, precision teachers are looking for application when they check to see if learners readily combine components into composites (also called compounds) during the instruction and practice of behavior composites once they have achieved specified ranges of count-per-minute responding on those components (Fabrizio & Moors, 2003). Bucklin, Dickinson, and Brethower (2000), one of the studies reviewed by Doughty et al., demonstrated this application effect in controlled research with a stimulus equivalence task. When precision teachers refer to adduction (Binder, 1996; Johnson & Layng, 1996), they are pointing to a special case of application in which behavior components at certain response rates combine with no explicit training on the composites. Optimizing adduction is a key element of generative instruction (Johnson & Layng, 1992; Johnson & Street, 2004).

**REAPS and Rate of Response**

In addition to a different definition of application, Doughty et al. (2004) introduce an error regarding Haughton's acronym REAPS (Binder, 1996; Johnson & Layng, 1996). Printed incorrectly as REAPs (small s) in their review, the authors claim that the acronym represents possible learning outcomes listed as "retention, endurance, application, and performance standards" (p. 8). In fact, the acronym posed a challenge from Haughton to identify performance standards (as count-per-minute ranges) that optimize retention, endurance, and application. This was an empirical alternative to the methodology of collecting samples of response rates from groups judged to be competent by other means, including social validation (Barrett, 1979; Haughton, 1980; Wood, Burke, Kunzelman, & Koenig, 1978).

Like reading teachers and language specialists, precision teachers use the term fluency to point to the temporal dimension of competent performance. Reading teachers measure fluency as count of words read per minute. Thesaurus entries such as fluidity, smoothness, quickness, dexterity, and rapidity reflect both the qualitative and temporal features of speech or other behavior referenced in everyday use of the term. Defining the term fluency separate from the time dimension, which Doughty et al. (2004) seem to do in their review (i.e., by appealing to this author and to others who have emphasized retention, endurance, and application as outcomes associated with achieving competent response rates), is potentially confusing. To define fluency as a level of performance that predicts retention, endurance and application without specifying a frequency range within which those outcomes are optimized is not consistent with use of the term by precision teachers. It ignores the fact that percentage correct—the available alternative—cannot distinguish between levels of performance that do and do not predict those outcomes.

Again, the central question is not whether practice procedures carried out beyond the point of 100% accuracy in the absence of time-based measurement can result in better retention, application, and so on. Extensive overlearning research during the 20th century demonstrated that such procedures do result in better long-term outcomes (Binder, 1993). The more fundamental issue in evaluating precision teaching is whether using Skinner's response-rate measure rather than percentage correct during teaching and learning is a more sensitive way to distinguish competence from incompetence and to make instructional decisions. In fact,
precision teachers use a variety of teaching and practice procedures, some timed and others untimed, to produce normal rates of freely emitted behavior. What distinguishes precision teachers from conventional educators (including other behavioral educators) is their use of daily response-rate measures graphed on the standard celeration chart (Pennypacker, Gutierrez, & Lindsley, 2003) for making instructional decisions. It is how they measure the effects of their procedures, not the myriad of procedures themselves, that primarily differentiates them from more conventional educational practitioners. Although Doughty et al. (2004) do not explicitly address this more general issue, reframing the discussion with this in mind may help to prevent future misunderstandings about precision teaching and its foundation in Skinner's response-rate measure.

Defining Endurance and Distractibility

In precision teaching research and practice, the term *endurance* has referred from its earliest use to a combination of features that characterize truly masterful fluent behavior (Binder, 1996). Masterful performers are able to continue performing over comparatively extended periods of time without being unusually subject to errors or distraction from events that can potentially compete for stimulus control. Although it is possible to observe unstable behavior under a variety of conditions—for example, in the presence of conditioned aversive stimuli or with extinction procedures—the type of stability referred to by precision teachers (e.g., Johnson & Layng, 1996) is a feature of the overall “endurance effect” observed in classrooms (Binder, Haughton, & Van Eyk, 1990).

Early laboratory investigation of the endurance effect (Binder, 1996) used a free-operant analogue of the discrete-trials procedure reported by LaBerge and Samuels (1974) in a verbal learn-
rate-controlled format” (p. 20). Stated another way, this is a recommendation to compare the effects of self-paced practice procedures that include counter-minute goals (rate building) with rate-controlled trials procedures that provide the same number of response opportunities, total exposure time, or rate of reinforcement.

Typically, those who conduct such research control for amount of practice by providing the same number of practice trials as the number of responses in a yoked procedure in which learners move at their own pace. Controlling for total time of exposure yokes time during rate-controlled trials to total time in which learners respond at their own pace. Controlling for amount of reinforcement might likewise compare the effects of rate-controlled trials with the effects of practicing freely emitted behavior in one or more control conditions, although number of reinforcements (one per session or one per response?) and sources of reinforcement (extrinsic, or both extrinsic and intrinsic?) can be difficult to define unambiguously in procedures that allow continuous self-paced responding.

One could argue from a technology standpoint that if a given number of rate-controlled trials were to produce equivalent retention, application, or other learning outcomes as the same number of freely emitted responses during self-paced practice, the less time-consuming procedure (i.e., freely emitted responses) is more efficient and therefore practically a superior procedure. But for those who wish to control for amount of practice, number of reinforcers, or any other features using controlled trials in comparison with self-paced practice procedures, it is critical to probe freely emitted response rates after the controlled trials conditions and before tests for learning outcomes such as retention and application to compare the immediate and direct impact of each procedure on response strength, using Skinner’s measure.

An obscure finding in the overlearn-
Conclusion

Doughty et al. (2004) have served the field by raising a number of experimental design issues to a new level of visibility. This article is an attempt to clarify potential confusions in their review. It also seeks to introduce several additional considerations that should be integrated into any research plan on the use of response-rate measures to define competence or to make instructional decisions.

REFERENCES


