Spring 5-1-2021

A Review of Behavior Analysis in Education

Devon A. Warner
Rollins College, dwarner@rollins.edu

A Review of Behavior Analysis in Education

A Capstone
by
Devon Allyn Warner

Submitted to the Faculty of the Department of Health Professions
at Rollins College in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF ARTS IN APPLIED BEHAVIOR ANALYSIS AND CLINICAL SCIENCE

April 2021
Winter Park, FL

Acknowledgements

I would first like to thank my extremely dedicated parents for their continuous, loving support and encouragement throughout this long, but beautiful journey of growth and discovery. I would also like to thank my best friend, Tommy, for never failing to remind me that I could, in fact, complete this daunting task. Thank you to my supervisors Michele McGlamory, Joy Bennett, Melissa Lopez, and Frances Torres, as well as my entire Quest and Quest Kids Academy Team, for all the incredible knowledge and lessons you have taught me throughout this research process. A special thanks goes out to my writing tutors Nia Page and Amir Famil for their guidance and patience. Morgan, Tierney, Quinn, Austin, Phillip, and all my amazing Focusmate friends (you know who you are): this paper would not have been possible without your friendship, kindness, and support. To my project advisors, Dr. Stephanie Kincaid and Dr. Michele Williams: I am forever grateful for your patience, inspiration, and encouragement. Dr. Sarah Freeman: you will never know how much of a positive impact you had on me personally and on this research project when you told me, “You are so brilliant, just start believing it!” Finally, thank you to my students who continue to inspire me daily and have helped make my dream of integrating ABA into education a reality.

Table of Contents

Abstract

Most education research in applied behavior analysis (ABA) is specific to early learners, which presents significant opportunity to better determine the effectiveness of various instruction methods for college-aged learners (CALs). Within the context of pedagogy, or method of teaching, ABA is the scientific study of the effect of instructional techniques on student behaviors. The available research for college-aged learners is scattered, non-definitive, and has gaps over time. Despite declension over the past half-century, research examining evidence-based practices in education has identified multiple approaches to help instructors manage and improve individual student behaviors and academic performance for CALs. In this review, I evaluated the available research in education generally, identified methods with the most empirically supported evidence as a best practice for teachers, and suggested topics for future research to help fill in some of the current knowledge gaps in ABA within the confines of education. Additionally, I reviewed the available behavior-analytic research with CALs. These studies were evaluated by determining whether the instructional methods demonstrated clear results with significant improvements in student behaviors and student satisfaction. Analysis of the research identified personalized system of instruction (PSI) as the most effective teaching method for improving CAL outcomes. Given this finding, it is recommended that researchers in the field of ABA who are focused on CALs, develop more methods specific to PSI, encourage teachers apply them in a real-world setting, and determine how best to make this information widely available.

Keywords: behavioral analysis in education, evidence-based teaching methods, pedagogy

A Review of Behavior Analysis in Education

Applied behavior analysis (ABA) is the scientific approach to studying and understanding behavior and its relationship with the environment. ABA is a science in which procedures developed from behavior principles are systematically applied to improve socially significant behavior. It is also a data-driven science that aims to determine whether behavior improvements are meaningful, and the methods used are responsible for positive behavioral outcomes (Cooper et al., 2007). Over the past half-century, behavior analysts have designed, developed, implemented, and tested many procedures that have produced considerable benefits. These procedures have demonstrated effectiveness with all types of learners in many different settings, including education (Twyman, 2014a).

Based on the most recent review of publication trends in the *Journal of Applied Behavior Analysis* (JABA) from 1968 to 1992, there was an increase of studies on target behaviors for individuals with developmental disabilities and a significant decline of studies researching academic behavior, verbal behavior, and behavioral excesses in typically developing children (Northup et al., 1993). Along with the decrease in studies on academic behavior, Northup et al. (1993) identified a shift in more research being conducted within community and naturalistic settings. There are recent data showing roughly that 75% of behavior analysts work in the practice area of treating children with developmental disabilities (DD) (Carr, 2016). The other 25% of behavior analysts work in practice areas such as education, animal behavior, criminal behavior, organizational behavior management, sports performance, and mental health (Carr, 2016). Additionally, over 90% of board-certified behavior analysts (BCBAs) work in the United States (Carr, 2016). There is parallel between these findings, which suggests a call for diversity of practice. The distribution of practitioners is heavily concentrated in work with children with DDs and areas outside education.

A small proportion of ABA research and practice in education is with college-aged learners (CALs) in applying behavioral teaching methods to improve student performance (Twyman, 2014a). The purpose of this review is to: (a) provide a brief history of behavior analysis in education, (b) give a general overview of behavioral teaching methods for CALs, (c) critically analyze the research on behavior teaching techniques with CALs, (d) compare the instructional methods with greatest empirical support and identify the most effective method for this population, (e) provide recommendations based on research on the most effective method for teaching CALs, and (f) provide a guide for future considerations in this research area.

History of Behavior Analysis in Education

In the mid-1960's, behavioral psychologists and researchers began systematically applying operant science to the development of teaching methods (Twyman, 2014a). Early researchers that contributed to the development of a behavior analytic approach for addressing socially significant behavior in education included B.F. Skinner, Fred S. Keller, Ogden Lindsley, Sidney Bijou and others.

A theoretical account of how to use behavior analytic techniques to address issues in teaching and learning is found in B.F. Skinner's *Technology of Teaching* (1968/2003). The *Technology of Teaching* (Skinner, 1968/2003), provides detailed information on teaching methods, such as programmed instruction, behavioral mechanisms for improving learning, and techniques for limiting behavioral excesses in the classroom environment. For example, this book (Skinner, 1968/2003) has models for how concepts such as, "contingencies of reinforcement" and "stimulus control" in operant conditioning may operate in everyday learning environments (Skinner, 1968/2003).

Concurrent with the emergence of Skinner's *Technology of Teaching* (1968/2003), several other researchers made advances in the application of behavior analysis to many different areas of education (Twyman, 2014a). After leaving Columbia University in the mid-1960s, Fred S. Keller developed the Personalized System of Instruction (PSI), a teaching practice initially designed for and used with college students (Keller, 1968). In 1965, Ogden Lindsley began a line of applied research at the University of Kansas Medical Center in special education, teacher training, and field-based educational research (Lindsley, 1971). Roughly simultaneously, Sidney Bijou began his work with children at the University of Washington's Institute of Child Development, where he recognized education could benefit immensely from the principles of ABA (Bijou & Baer, 1961).

The Decline of ABA in Education

Recent research in behavior analysis in education (Twyman, 2014a) suggests a decline in the widespread use of empirically supported interventions in schools and classrooms. This finding is related to an earlier research discovery (Sulzer-Azaroff & Gillat, 1990), which found that behavior-analytic research in education publication trends in JABA had significantly declined. Most of the research conducted in educational settings almost exclusively focused on early learners, thereby largely neglecting students at the middle-school through the university level (Sulzer-Azaroff & Gillat, 1990). Similarly, there is declining trend in the percentage of publications in JABA dedicated to studying academic related variables (Northup et al., 1993).

There are several contributing factors to this decline (Twyman, 2014a). One reason may be that educators and practitioners are simply not aware of the available evidence-based interventions (Twyman, 2014a). There is a missing link between evidence of effective instructions and the methods being implemented (Carnine, 1997, 2000; Hall, 1991; Individuals with Disabilities Education Act (IDEA), 2004; No Child Left Behind (NCLB), 2001). The issue of translating knowledge to practice has been referred to as the "research-to-practice gap," and has been seen in many fields other than behavioral education (Kratochwill et al., 2004). Also, a lack of dissemination of evidence-based practices within behavioral education to areas outside of the field has led to many recommendations from researchers. These recommendations include strongly encouraging behavior analysts to collaborate with, clearly distribute ABA knowledge to, and work alongside other fields and practitioners outside ABA (Binder, 1994; Foxx, 1996; Pennypacker & Hench, 1997).

Current State of ABA in Education & Necessary Changes

Technology has had an impact on education overtime. Additionally, the educational system could benefit from further integration of behavior analysis and technology. Education technology refers to the hardware, software, and technological devices used within education (Twyman, 2014b; Layng & Twyman, 2013). A transformation in how we view and provide education is actively occurring in today's society. This is because education has become more interactive and personalized with focus in self-learning, independent of behavior analysis (Twyman, 2014b). This transformation originates from the evolution in digital technologies, which sparked a revolution in education across both K-12 and university classroom environments. The evolution of digital technologies, such with online learning and computer-based instructional programs, has increased instructor access to sufficient student data. This increase in data can be used to improve education and add to the research in learning and behavior change (Twyman, 2014b). Behavior analysis is a data-driven science that has a lot to offer to this education revolution.

"Education 3.0"

The transformation and revolution in education is referred to as "Education 3.0" (Twyman, 2014b). "Education 3.0" features universally flexible digital environments and experiences, with programs that possess the ability to self-learn and adjust to the individual learner (Twyman, 2014b). Instigated by technology, educational systems are realizing the potential of “Education 3.0,” including the capacity for increased personalization and improved academic outcomes (Twyman, 2014b).

The evolution of education overtime has been paralleled with that of the web (Lengal, 2013); for example, “Education 1.0” with the agricultural age, “Education 2.0” with the industrial age, and “Education 3.0” with the era in which we’re moving. The role of technology, teacher-student interaction, levels of teacher training, school settings, learning materials and outcomes, and the role of student and teacher have developed significantly across each variation (e.g., Education 1.0, 2.0, and 3.0; Twyman, 2014b). Characteristics of “Education 3.0” include: a ubiquitous technological influence; versatile teacher-student relationships; teachers with various experience, such as professional training, certifications, or students themselves; school locations found anywhere and everywhere, including remotely; instructional materials that derive from multiple sources; academic and social learning outcomes; teachers who simultaneously orchestrate, curate, and collaborate; and finally students whose role involves actively learning by taking ownership and responsibility (Keats & Schmidt, 2007; Moravec, 2008, as cited in Twyman, 2014b).

“Education 3.0” aligns with ABA in several respects including: instructional-objective mastery depends on the individual pace of each student; objectives feature rapid, frequent responding during self-paced instructional sessions; and finally, instructional methods adjust accordingly in response to measures of the individual’s learning and performance (Barrett et al., 1991). Components such as “personalized learning” and “competency-based education,” (Twyman, 2014c; US Department of Education, n.d.; Sota, 2016a) are consistent with a behavior analytic approach to education and contribute significantly to the concept of “Education 3.0” (Cooper et al., 2007; Twyman, 2014b).

A symbiotic relationship must occur between teaching methodologies and education technology to maximize and improve teaching and learning outcomes. Subsequently, education may benefit from a fusion of behavior analysis, learning science, and technology (Layng & Twyman, 2013). For example, a teacher could do an in-real time fluency check of previously learned material by asking questions aloud to the class while simultaneously presenting questions and possible answers via a digital media tool (i.e., Kahoot!©, Poll Everywhere, etc.), all while student responses and subsequent response feedback could be provided using an electronic device (i.e., computer, phone, tablet). The above example emphasizes a technological application of several behavior teaching methods (e.g., precision teaching, response cards). Providing insightful, targeted, and informative data is highly suggested (Layng & Twyman, 2013) Additionally, selecting educational technology tools that provide data can allow for continuous, systematic evaluation of the instruction or learning environment's effectiveness and create a foundation for maximizing their efficiency (Layng & Twyman, 2013).

"Education 3.0" could be actualized in educational systems throughout the U.S. and other countries in the world through the integration of adaptable behavioral principles, behavioral teaching methods, and education technologies. Before integrating education and technology across educational systems, it might be beneficial to start small and incorporate this fusion within classrooms and evaluate the effects on student performance. For example, researchers within ABA, education, and government education policy could collaborate to systematically test the integration of technology with different evidence-based practices within classrooms, schools, and school districts, respectively. This revolution in education for all populations, including CALs, may be accomplished by following the recommendations made by earlier researchers (Axelrod, 1991, 1992; Binder, 1994; Foxx, 1996; Pennypacker & Hench, 1997; Twyman & Heward, 2018); behavior analysts must actively disseminate evidence-based practices and collaborate with others outside of the field on research and in practice.

While there is a clear declining trend of research being conducted in behavioral teaching methods, there is also an issue with the translation of evidence-based instructions to practice in numerous educational settings. Therefore, it is important to identify the specific basic principles of ABA that evidence-based behavioral teaching methods rely on.

Principles of Learning in ABA

Programming principles for learning were first found in the early writings of Comenius in the seventeenth century and even as early as in the Ancient Greek era in Socrates' writings (Skinner, 1968/2003). For example, "the student should not be asked to take a step he cannot take" (Comenius paraphrased in Skinner, 1968/2003, p. 75). It is easy to see how this sentiment is reflected by subsequent researchers in behavior analysis in education, including "the child knows best" (Lindsley, 1990, p. 12), and "the student is always right" (Keller, 1968, p. 88). Taken together, these quotes express a fundamental tenet of behaviorist educational philosophy: If a student is not learning, it is not the fault of the student, but rather the system of instruction (Engelmann & Carnine, 1991). Behavioral research in education aims to improve practice, and subsequently the outcomes and lives of students. A behavior analytic approach to educational research can achieve this by focusing on improving teaching and learning in ways that are direct, observable, measurable, and socially valid (Twyman, 2014a).

Foundational ABA Concepts in Education

The principles necessary for implementing an effective behavior analytic approach to teaching rely on fundamental concepts; these concepts are foundational for some common procedures used in college teaching, such as contingent reinforcement and response cost (Austin, 2000). A behavior analytic approach must include clear, operationally defined behavioral units for behavior change that can be verified experimentally (Embry & Biglan, 2008). Basically, student behaviors targeted for change are clearly defined and easily measurable for instructors to check for improvement. This approach must also be relatively simple to implement with easily detectable outcomes across germane behaviors that are socially significant (Embry & Biglan, 2008). Teachers are more likely to implement interventions if they are easy to implement and test, have observable effects, allow for significant improvements to important classroom problems, and can easily be integrated to current routines (Rogers, 1995).

Learning Principles & Behavioral Outcomes Necessary for Education

Equivalence Relations and Stimulus Equivalence. An equivalence relation is the emergence of new, predictable units of behavior from units of behavior indicated previously in the behavioral repertoire (Sidman, 1997). Another way to explain this principle is learning builds more opportunities for learning. Therefore, a previously mastered skill might be a prerequisite for learning additional, related skills. For example, a child who has learned to identify the color 'red' when presented with a toy fire truck is then able to identify an apple as the color 'red'; this is an equivalence relation.

The notion of equivalence relations and stimulus equivalence has been applied to the development of instructional designs in higher education, such as equivalence-based instruction (EBI) (Brodsky & Fienup, 2018). Procedures like EBI provide instructors with a framework for breaking down a skill into all the prerequisite components required to acquire specific target skills. Based on how discrete the instruction is in design, the responsibility of learning is less on the learner, and instead is more attributed to the instruction itself (Sidman, 1997; Brodsky & Fienup, 2018).

Concept Learning. One behavioral outcome that is important to education is concept learning. A concept refers to a set of shared features found in each example of the concept, and all examples of the concept will share certain "must have" and certain "can have" features in order to be included in the concept (Layng, 2013). The "must have" features are the definite features required for inclusion in the concept, whereas the “can have” features refer to the features that examples might or might not have for inclusion. The “can have” features might change from example to example, but do not necessarily define the example as an instance of that concept (Layng, 2013).

There are several necessary requirements for a student to correctly demonstrate a learned concept (Layng, 2013). The learner must be able to distinguish examples of the concept from similar nonexamples that are missing one or few defining “must have” features; the learner must also be able to recognize examples of the concept among a wide variety of non-defining “can have” features; and the learner must be able to generalize the understanding of the concept beyond instruction when presented with examples and nonexamples in the natural environment (Layng, 2013).

***Concept Learning in Education and Relevant Research*.** Concept learning, also known as concept formation, is typically referred to as the process of classifying information into meaningful categories. At the basic level of concept formation, the learner comes into contact with positive and negative (i.e., examples and non-examples) occurrences of the concept (Owens, 2016). Before teaching concepts, teachers must consider these positive and negative examples of each and the conditions under which they occur. Teachers must also instruct students to identify the critical features that define what an example of a concept is and is not.

Behavior analysts rely on another method of teaching, shaping, in which small, successive approximations to a target skill are taught in sequential order (Cooper et al., 2007). This method is also incorporated in behavioral teaching methods used with CALs (Mayer et al., 2014). When instructors solely rely on shaping to teach students across a variety of education settings including in higher education, they might overlook basic methods for teaching concepts, such as rule governance, stimulus control techniques, discrimination training, etc. (Owens, 2016).

Some of these methods for concept formation were initially discovered and tested in basic behavioral research studies in experiments with pigeons (Sidman, 1960). For example, in a three-experiment study (Herrnstein et al., 1976), researchers used discrimination training to successfully teach pigeons to only press a key when presented with correct examples of several different stimuli found in nature (e.g., trees, water, person). Discrimination of stimuli in concept formation is when stimulus control of behaviors has been achieved (Sidman, 1960). Stimulus control describes situations in which a behavior is triggered by the presence or absence of a stimulus or signal (Cooper et al., 2007). Animal research, like the study mentioned above, helped researchers understand concept formation from a behavioral perspective. The term ‘memory’ processes, which is typically a more mentalistic ideology of concept formation and more commonly recognized in education literature, can be stated in behavioral terms simply as the “acquisition and loss of stimulus control” (Branch, 1977, p. 178). Finally, another effective method for teaching concepts is the use of task analyses, which is a procedure that allows the learner to learn each prerequisite component in a sequence that helps to build the repertoire for understanding complex and abstract concepts (Owens, 2016).

In summary, concept learning, as a behavioral outcome, can be achieved using procedures from basic behavioral research discussed above (Herrnstein et al., 1976), which have been incorporated in several evidence-based practices used for teaching CALs. Many behavioral principles functionally depend on each other. Task analyses, which are used to teach concepts, require a technique called shaping.

Shaping. Shaping is a procedure in which differential reinforcement is provided for successive approximations (e.g., baby steps) towards a specific behavioral goal (Mayer et al., 2014). Successive approximation is a term used by behavior analysts which refers to when a particular response has changed in dimension (i.e., rate, quality, or intensity) towards the direction of more closely resembling the target behavior (Mayer et al., 2014). Shaping has been used in many practice areas of ABA: such as teaching skill acquisition, appropriate social skills, and decreasing problem behaviors in individuals with DDs (Harris et al., 1967; Cameron & Cappello, 1993; Carr et al., 1980). Shaping has also been used effectively to teach CALs (Sulzar-Azaroff, 2004).

Shaping in Education. Shaping has been a key feature in the development of procedures used within many different practice areas in the field of ABA. Shaping has also been incorporated in many evidence-based methods for teaching CALs, including programmed instruction (PI), direct instruction (DI), precision teaching (PT), a form of PT called "say-all-fast-minute-everyday-shuffled" (SAFMEDS), PSI, and computer-aided personalized system of instruction (CAPSI) (Mayer et al., 2014). Shaping has proven an effective behavioral technique for teaching successive approximations to various preferred targets, or end-goal responses. For example, DI instructors use reinforcement (e.g., praise) and extinction (i.e., the absence of reinforcement) paired systematically and simultaneously with verbal, auditory, visual, and gestural prompts, to teach correct individual and/or group responding to academic goals.

There are several different steps for using shaping effectively (Mayer et al., 2014). Research suggests instructors become familiar with and assess their students' behavioral repertoires prior to implementing or incorporating shaping within the teaching methods they plan to use. By assessing current repertoires, teachers will be able to identify the best place to start, as this might vary significantly across individual students (Mayer et al., 2014). It is also suggested that teachers select and apply a reliable, valid ongoing measurement system which will make assessment of progress much easier.

As mentioned briefly in the section prior for teaching concept formation, task analyses or breaking up a task into subcomponents, can be used to better identify each sequential step for shaping successive approximations toward the desired end goal (Mayer et al., 2014). This research also suggested several other procedures instructors can use in combination with shaping, include discriminative stimuli and prompt fading (Mayer et al., 2014). Prompts and descriptive feedback can function as discriminative stimuli to help teach the baby steps in a sequence, but these prompts need to be faded out eventually to allow students to perform the desired task more independently.

"Goal-setting" is another way shaping is used in behavioral literature with adults, in which the completion of each previous goal and the subsequent access to the following goal, acts as a reinforcer for that individual (Sulzar-Azaroff & Harsharger, 1995).

Evidence-Based Practices & Behavioral Teaching Methods

Overall, research in ABA and education has led to the development of many teaching tactics used to identify and change alterable classroom variables (Twyman, 2014a; Embry & Biglan, 2008; Heward et al., 1994/2005; Twyman & Heward, 2018). All evidence-based practices that yield proficient student outcomes share this basic structure: a solid instructional design, high rates of relevant learner responses with contingent feedback, and an ongoing instructional decision-making based on direct and frequent measures of student performance (Markle, 1983/1990; Twyman et al., 2005; Hattie & Timperly, 2007; Heward, 1994; Bushell & Baer, 1994; Twyman & Heward, 2018).

There are several evidence-based practices and models found to improve student learning and produce effective outcomes across a variety of populations within education. These include, but are not limited to: programmed instruction (PI), direct instruction (DI), precision teaching (PT), personalized system of instruction (PSI), response cards, functional behavior analysis and assessment (FBA), class-wide peer tutoring (CWPT), The Competent

Learner Model (CLM), The Comprehensive Application of Behavior Analysis to Schooling (CABAS), The Pyramid Approach and the Picture Exchange System (PECS), School wide Positive Behavioral Supports (SWPBS), equivalence-based instruction (EBI) (Brodsky & Fienup, 2018), interteaching (Querol et al., 2015), flipped learning (Sota, 2016b), and many others (Twyman, 2014a). See Twyman's review (2014a) for a more detailed description of the literature on the instructional methods effective in K-12 education. For the benefit of this literature review, detailed discussion will be provided only on the evidence-based practices found to be effective and empirically supported with CALs, as this is the population of learners that has been neglected the most in the research (Twyman, 2014a; Sulzer-Azaroff & Gillat, 1990; Northup et al., 1993).

Programmed Instruction

Programmed instruction (PI) was the first instructional design to fuse a science of learning with technology and is also an instructional application of shaping procedures and reinforcement contingencies (Skinner, 1958; Mayer et al., 2014). In 1958, Skinner created the teaching machine, a device for delivering PI created from laboratory derived methods.

PI is a specialized method of teaching that presents material in structured, logical, and systematic sequences that feature behavioral objectives, such as reinforcement, high rates of relevant activities, shaping successive approximations, and mastery progression (Twyman, 2014a). After experimental use of PI began in the 1920s and 1930s, B.F. Skinner and J.G. Holland implemented PI in behavioral psychology courses at Harvard University in the late 1950s (Holland & Skinner, 1961). Adaptations of PI have helped teachers to deliver strategic instruction, which continues to be integrated in teaching curricula today (Twyman, 2014a). PI is an early instructional design based on basic behavioral principles that has been integral in the development of other behavior-analytic approaches to education with CALs that exist today (i.e., PT, PSI, PI, Interteaching, etc.) (Querol et al., 2015). PI was the first behavioral teaching method to heavily rely on technology. Many other behavioral teaching methods developed have followed suit and incorporated technology to increase the efficiency of design.

Direct Instruction (DI)

Direct Instruction (DI) involves carefully planned lessons broken into small learning sections specifically designed for individuals (Kinder & Carnine, 1991). Using scientific methodology to inform the development of procedures effective for teaching concepts and ideas, Zig Engelmann and colleagues Carl Bereiter and Wes Becker constructed what became known as "DI" in the early 1960s. This method is a carefully designed instruction model that relies heavily on prewritten scripted sequences (Twyman, 2014a).

The instructional design and delivery principles of DI differs from other behavioral education approaches in its degree of manipulation of antecedent stimuli, such as teacher wording and delivery of instruction (Kinder & Carnie, 1991). For example, teachers might use key phrases, or add vocal inflections and other auditory-visual prompts, to help cue students to how and when to respond. DI relies on multiple design and delivery principles that allow for effective instruction.

The instructional design principles employed in DI-based curricula include explicit teaching of rules and strategies, example selection, example sequencing, and covertization (Kinder & Carnine, 1991). The first design principle, the explicit teaching of rules and strategies, involves teaching steps of a target response or sequence of responses overtly. Example selection requires the instructor to implement a guided practice of the previous learning material. During guided practice, the instructor presents a variety of exemplars and sequences that include examples and matching non-examples, which promote generalization of the learning material. The third design principle, example sequencing, involves the instruction of examples and matched non-examples be delivered in succession to foster the detection of small changes and teach discrimination. Lastly, covertization, requires steps initially taught overtly, be systematically faded in subsequent instructions. This systematic fading procedure will promote students' ability to apply strategies covertly (i.e., privately to themselves) and independently.

The delivery principles or specific teaching procedures, which are essential features to DI, include brisk pacing of questions and specific correction procedures (Kinder & Carnine, 1991). The delivery principle involving the quick pacing of questions, is the fast pace delivery of items and subsequent answers, which helps to maintain the interest of individuals in groups and create behavioral momentum for high-rate responding. This delivery principle also increases the potential amount of instructional material delivered in a fixed time. The other delivery principle which involves implementing specific correction procedures, is vital for providing consistent, immediate academic feedback, while the particular instructor feedback is provided based on the type of error committed. This delivery principle involves adult direction, repetitious practice, and choral responding, which allows for more opportunities to respond (Kinder & Carnine, 1991).

There is extensive research supporting the effectiveness of DI procedures across a variety of ages, and ability levels in typical classroom and in special education settings. Overall, findings indicate DI produces higher academic gains when compared with other forms of instruction (Kinder & Carnine, 1991; Becker, 1977; Becker & Carnine, 1980; Becker & Gersten, 1982; Gerstan, 1985; Horner & Albin, 1988; White, 1988; Ganschow & Sparks, 1995; Patching et al., 1983). The prior research literature has primarily examined effects of DI with elementary, middle, and high school students (Becker, 1977; Becker & Carnine, 1980; Becker & Gersten, 1982; Ganschow & Sparks, 1995; Patching et al., 1983), and there are very few studies that examine effectiveness of DI with CALs or adult learners.

DI Research with CALs

One study (Kitz & Thorpe, 1995) used DI with 26 recent high-school graduates with learning disabilities to assess the acquisition of algebra skills. Students enrolled in an 8-week summer course in college algebra were randomly assigned to either the control or experimental group. The control group received a traditional textbook, while the experimental group received a videodisc with an algebra instructional guide based on DI principles. Results showed that students in the videodisc group scored higher in both post-tests and quiz scores from their first college algebra course (Kitz & Thorpe, 1995), relative to the students who only received textbooks.

While somewhat less behavioral in design, the following studies were the only other DI research studies found that addressed the performance of CALs. For example, in Casazza (1993) used a model of DI to teach summary writing in a college reading class. While this study presented no experimental data on the students writing improvements, there were some social validity measures used and high student satisfaction was reported. Another study (Cummings et al., 2010), used a quasi-experimental design to evaluate the effects of DI on undergraduate education majors' moral reasoning and compare the pre- and post-test scores of students in the experimental group (course section that received DI) and control groups (course sections that received traditional lecture format). Results showed a statistical difference between the pre- and post-test scores for students in the DI course and between the students' post-test scores in the DI course and the students in traditional lecture courses. These results indicated that students in the DI course made significant improvements from pre- to post-test scores in moral reasoning tests compared to that of the students in the traditional lecture courses.

There are several limitations to the available DI research, some of which are indicative of the limitations of the behavior-analytic educational research literature, generally. The majority of DI research utilizes group designs (Kinder & Carnine, 1991), without comparison to single-subject data. There are several limitations with this approach (Barlow et al., 2009) including the ethical issues with control groups not receiving interventions, and that inter-subject variability can be lost in aggregated data. Also, the dependent variables (i.e., measures of academic behavior change) in some of the available DI research with CALs used were standardized tests (Cummings et al., 2010), which are only as effective in measurement as they are in their design (Larsson, 2014). Furthermore, Casazza (1993) was lacking empirical support for the effectiveness of DI for improving college writing, but based on the reports of high social validity, could be a useful model for future research in writing courses for CALs.

Precision Teaching (PT)

Precision Teaching (PT) is a precise and systematic method for evaluating instructional tactics and the components of school curricula that involve making effective educational decisions by closely monitoring the changes in performance patterns along standard celeration charts (Lindsley, 1992b). PT follows four guiding principles that include: an emphasis on directly observable behavior, frequency as a measure for performance, relying on standard celeration charts (SCC) as a simplified graphic display tool for conducting data analysis and predicting future performance (Lindsley, 1991; Aniano et al., 2015), and the concept of individualized, student-centered learning (Lindsley, 1992b). The teacher's role involves arranging the learning environment, including the structure of the materials and methods used. The goal is to create an environment that teaches the students to teach themselves using strategies such as self-counting, timing, charting, one-on-one direction, and support (Lindsley, 1992b; Twyman, 2014a). PT has also been referred to as generative learning (Mayer et al., 2014; Johnson & Layng, 1992) because fluency training encourages behavioral cusps to occur, which strengthens the comprehension of new material as a result of previous learning.

PT has been incorporated in many educational program settings to improve various academic skills with learners of all ages (Layng et al., 2003, 2004; Johnson & Layng, 1992, 1994; Merbitz et al., 2004). PT was founded by Lindsley in 1965 and initially implemented in special education classrooms in the Children’ Rehabilitation Unit of the University of Kansas Medical Center (Linsley, 1991). Initial success with self-monitored behaviors specific to children's handicap deficits led special education teachers within the unit to implement PT to improve their primary and elementary academic skills (Koenig, 1967; Slezak, 1969; Edwards, 1969; Fink, 1968). Through the work directed by Eric Haughton and fellow students from the University of Oregon, they integrated PT within the elementary classrooms of typically developing children (Haughton, 1972; Johnson, 1971; Starlin, A., 1971, 1972; Starlin, C. M., 1970). A short time later, district- and state-wide projects gained federal funding to train numerous teachers to use PT, which instigated class-wide implementation with middle and high schoolers (Beck, 1976; Sokolove, 1978; Willis, 1974; Beck, 1981).

Lindsley emphasized PT can be incorporated or used in combination with any teaching method or approach. For example, there have been many effective applications of PT when combined with DI (e.g., Kubina et al., 2009) and as a component in contemporary sophistication of PI educational designs (e.g., The Headsprout Basic Reading Program) (Layng et al., 2003, 2004). Additionally, many collegiate programs (i.e., Rollins College, University of Florida, Florida Institute of Technology, Jacksonville State University, etc.) incorporate several forms of PT commonly used with CALs (e.g., SAFMEDS and computer-based PT).

Precision Teaching with CALs

The say-all-fast-minute-everyday-shuffled (SAFMEDS) strategy is used for instruction and assessment (Quigley et al., 2017) . This method emerged from Lindsley’s between the late 1970s and early 1980s, and was first used with his graduate students (Calkin, 2003; Graf & Auman, 2005; Potts et al., 1993). SAFMEDS is a high-frequency oriented method of using flashcards, in which students ‘see’ a stimulus and ‘say’ a response. There are other variations of the stimulus and response in SAFMEDS, which can be altered depending on the instructional material and learner’s preference (e.g., see/write, hear/say, see/sort) (Quigley et al., 2017). The standard SAFMEDS procedure requires the student: 1) hold the full deck and shuffle cards, 2) start one-minute timer, 3) 'see' information on the front and 'say' what is on the back, 4) turn card over to check for accuracy, 5) sort correct and incorrect responses into piles, 6) count correct and incorrect answers when the timer ends, 7) take data on progress and review difficult material, and finally, 8) repeat daily (Lindsley, 1996; Quigley et al., 2017).

A recent comprehensive review of SAFMEDS literature (Quigley et al., 2017) found 16 of 27 evidence-based studies including participants who were either CALs or adult learners in general; the remaining studies were with adolescent participants. While all learners gained improvements in various disciplines, all peer-reviewed studies reportedly used adaptations of the standard SAFMEDS procedure (Lindsley, 1996) to meet individual needs (Quigley et al., 2017).

Another study (McDade et al., 1985) evaluated the differences in the fluency performances of undergraduate students receiving both SAFMEDS and computer-based PT formats separately. Assessment of individual student preferences for either PT strategy occurred following the investigation. The study found the majority of students (14 of the 18 students in one section and 13 of the 15 students in the other) demonstrated more significant fluency improvements with SAFMEDS compared with the computer-based PT program. However, a higher majority of the students preferred computer-based instruction over SAFMEDS, as there were considerable student reports of anxiety with SAFMEDS, mostly related to the procedure's self-monitoring component (McDade et al., 1985). While, some students reported the computer-based PT was "less personal."

Several projects and programs have implemented PT with success in improving academic performance for CALs who have benefited little from the traditional lecture format. These include the Malcolm X College in Chicago (Johnson & Layng, 1992, 1994), which has been operating since the early 1990s, and Jacksonville State University (JSU) in Alabama, which was founded in 1978 (McDade & Brown, 2001). Additionally, an adaptation of PT was implemented successfully with job seekers and pre-college learners seeking enrollment in community colleges (Johnson & Layng, 1992), and demonstrated academic and pre-vocational skill improvements for more than half of the students.

There is an abundance of empirical support for the application of PT with CALs and many successful variations used with this population. The available PT research with this population has used empirically valid experimental designs and data evaluation techniques. For example, more than half of the SAFMEDS articles reviewed in Quigley et al. (2017) met their data-based and peer-reviewed criteria. Also, McDade et al. (1985) used visual analysis to evaluate individual student data when comparing their fluency improvement rates with SAFMEDS and the computer-based PT program.

There are also some disadvantages to the available research on SAFMEDS and general PT research. While individualization might be beneficial to some learners, regarding the SAFMEDS procedure design specifically, too many variations in procedure across the research makes it challenging to provide a generalized analysis of the effectiveness of SAFMEDS (Quigley et al., 2017). Future research might benefit from the use of a more conceptually systematic design across participants, such as the standard SAFMEDS procedure (Lindsley, 1996). Another consideration is that while research has demonstrated most CALs show more improvements in fluency with SAFMEDS than with computer-based

PT programs (McDade et al., 1985), student preferences were much higher with computer-based PT programs. Instructors should consider this study when choosing which PT form to implement with CALs, and selections should take student preferences and general academic workload into account. Future research might benefit from component analyses of each PT format and further research on student preference.

Response Cards

Response cards are physical signs that are held up by all students in a classroom to display their response to questions and problems posed by the teachers and function to increase active student responding during group instruction. This approach derives from the theory that the active student response is the concept of “the learning trial” as the basic unit of instruction (Randolph, 2007). A learning trial is made up of three components: an instructional antecedent, a student response, and teacher feedback. While some argue that considering the learning trial to be the basic unit of instruction distorts and underestimates the nature of the teaching and learning process, research has demonstrated that the frequency and intervals of learning trials are tightly correlated with academic success (Randolph, 2007; Heward, 1994; Roseshine & Berlinger, 1978). Therefore, active student response interventions such as response cards may increase the number of learning trials that occur during classroom instruction.

A review of the research in response cards shows that this method has had significant results across various populations, settings, instructors, and behaviors. For instance, response cards have been found to increase active student responding (Gardner et al., 1994; Narayan et al., 1990), improve test scores (Cavanaugh et al., 1996), and decrease disruptive classroom behaviors (Armendariz & Umbreit, 1999; Chirstle & Schuster, 2003; Davis & O’Neil, 2004; Godfrey et al., 2003; Lambert et al., 2006). For example, Narayan et al. (1990) evaluated and compared the effects of write-on response cards and hand-raising with fourth grade students.

The authors indicated higher student participation and quiz scores occurred on the day's response cards were used. In a meta-analysis, Randolph (2007) found statistically significant effect sizes for test achievement, quiz achievement, participation, and reductions in problem behavior for all 18 studies that used response cards, regardless of the type of response cards used, place of publication, type of publication, and sample size. Unfortunately, there is very little research studying the effects of response cards to improve student performance with CALs. The results of several of these studies will be considered.

Response Cards with CALs

Response Cards for Improving Quiz Scores. Kellum et al. (2001) investigated the effects of response cards on the quiz scores of students in a community college education course. Using an alternating treatments design (Kazdin, 2011; Cooper et al., 2007), they compared the effects of presenting review questions with and without response cards. Answers to review questions were provided via hand raising in the review question only condition, and student responding was optional. In the response card condition, all students were prompted to raise cards with the individually selected responses. The authors found there were higher quiz scores in the response card condition when compared to the review only condition. Marmolejo et al. (2004) systematically replicated Kellum et al. (2001) and evaluated the effects of response cards on daily quiz scores with CALs. The authors also used an alternating treatments design to compare the effects of response cards and traditional lecture format (i.e., hand raising) on student quiz performance. This study resembled Kellum et al. (2001) in the design of response card and lecture conditions but differed because quizzes occurred daily rather than weekly. Marmolejo et al. (2004) found similar results to Kellum et al. (2001) and reported higher daily quiz scores in the response card condition. Results of these two studies demonstrate empirical support for using response cards to improve daily and weekly quiz scores with CALs. The finding of this research also suggests response cards can improve students' long-term maintenance (i.e., retention).

Malanga and Sweeney (2008) further extended the research conducted by Kellum et al. (2001) and Marmolejo et al. (2004). The authors used an alternating treatments design to evaluate and compare two different active student responding (ASR) methods on weekly quiz scores for CALs: daily assessments and response cards. In both Kellum et al. (2001) and Marmolejo et al. (2004), the baseline was identical to the standard lecture format condition, a non-ASR method, which they compared to response cards. On the other hand, Malanga and Sweeny (2008) compared two different ASR methods in the intervention phase relative to baseline, which was also identical to the standard lecture format condition in the previous studies. In the daily assessment condition, instructors provided five short questions related to lecture material at the end of class. After students finished answering questions, the teacher reviewed each item verbally and progressively provided correct answers via overhead transparency. In the response card condition, the instructor provided review questions at the end of the lecture, which was provided gradually via the projector. After time elapsed, which varied depending on question difficulty, the instructor prompted students to raise their cards with student-selected responses. Relative to baseline, results of students' average weekly quiz scores showed higher weekly quiz scores in the daily assessment when compared to those in the response card condition. However, evaluation of the individual data indicated results were idiosyncratic (Malanga & Sweeney, 2008). For some students, response cards were significantly more effective than daily assessments for improving weekly quiz scores. This study's findings extend the response card literature, and idiosyncrasies in individual data suggest that instructors select teaching methods carefully. It might be beneficial to expose students to multiple ASR methods and assess individual data on quiz performances and student preference on methods used.

Long Term Effects of Response Cards. Another study (Shabani & Carr, 2004) evaluated the long-term effects of response cards on academic achievements and participation in two experiments with university psychology majors in two Research Methods courses. The authors used a between-subjects quasi-experimental group design in the first experiment and a single-subject strategy in the second experiment. Both included the same three conditions: 1) review questions with response cards, 2) review questions without response cards and 3) a standard lecture format. Experiment one measured students' daily quiz scores, while experiment two measured students' percentage with final class "A" grades. They reported that while response cards improved performance on daily quizzes in Class A, the difference was only slightly higher compared to the other conditions. Response cards did not have an impact on the daily quiz scores of the students in Class B. Results of the second experiment indicated no difference in the percentages of "A" grades across conditions for either course section. The effects of response cards alone might not be sufficient to maintain performance over more extended periods (Shabani & Carr, 2004), which is contradictory to suggestions from previous findings (i.e., Kellum et al., 2001; Marmolejo et al., 2004). Although this study's results did not indicate that response cards were a reliable intervention for improving final grades, they did find the social validity of the instructional approach was significantly high. For instance, response cards improved participation significantly; students were more actively engaged, and reportedly enjoyed class more.

A limitation of Shabani and Carr (2004) is while they used a single-subject strategy to display graphs, which allows for clarity in analysis of the data, the authors did not provide figures or reports of the individual data like Malanga and Sweeney (2008). Without this, it is impossible to determine if some students' daily quiz scores benefited from response cards more than others, especially considering participation measures were significantly high in this condition. However, compared to the other studies, Shabani and Carr (2004) provided additional data on observable behaviors (e.g., response card raises, hand raises, and callouts) rather than just permanent products (e.g., daily and weekly quizzes, final exam grades). While permanent products are a good measure for student academic achievement, ongoing measurement of observable student behavior can help instructors monitor the subtle effects of teaching methods and advise when adjustments may be necessary.

While the effects of the previously discussed studies on response cards with CALs are relatively mixed, several studies outside of ABA offer this research area. Student response systems (SRS) are similar to response cards, but instead of using physical cards, they require clickers and other technological devices. With SRS, teachers still provide a question and allow students to respond. SRS research indicates its' effectiveness for increasing student participation (Dangel & Wang, 2008; Stowell & Nelson, 2007; Trees & Jackson, 2007), academic achievement (Yaoyuneyong & Thornton, 2011), and more preferred than traditional lecture format (Kaleta & Joosten, 2007). Some other online programs and applications that are functionally similar to response card technology, have developed in the past decade (i.e., Kahoot! ©, Poll Everywhere, etc.). While there is no research published evaluating this similar response card technology's effects, these programs are typically free and easy to access and download on various student-owned devices (i.e., smartphones and laptops). This other response card technology also takes data on correct and incorrect responses. It might be potentially more cost- and time-efficient for teachers to implement than response cards and some SRS methods (i.e., clickers). Another advantage of using this response card technology is it can easily be integrated into distance education (i.e., online learning) to increase academic performance and student participation, which is often lower in the latter by the nature of this educational environment.

Interteaching

Interteaching is a relatively new behavioral teaching method founded on principles of learning, utilizes features of other evidence-based practices, and is most used with CALs (Boyce & Hineline, 2002). The six essential components of interteaching (Querol et al., 2015) include:

a) The first is a preparation guide that outlines readings, consists of 10-12 practice questions, and functions as a study guide for students to prepare for and help facilitate high-quality group discussions in the subsequent class. These are typically distributed via hard copy at the end of each class, or several days prior in online student platforms (i.e., Blackboard).

b) The second is the "interteach" component which consists of small group peer guided discussions for reviewing study questions. The instructor actively supervises group discussions, responds to the problems within groups, and intervenes when necessary, to guide discussions.

c) The third component is an online or paper record sheet for students to provide instructors with frequent feedback on any problems with study questions, the quality of group discussions, and to report on the assistance supplied from group members. The instructor reviews the record sheet and addresses students' queries and other relative issues in a clarifying lecture.

d) The fourth component of interteaching is the clarifying lecture, which addresses the topics on the record sheet. Clarifying lectures will typically occur at the beginning of the following class, which allows for sufficient time for the instructor to review the record sheet and adequately prepare.

e) Another component is the small, frequent probes for mastery of the material (i.e., quizzes and exams). These are modeled after the questions and assigned reading in the preparation guides, and overall grades are not primarily affected by low grades on a particular probe. This component allows instructors to provide ongoing assessments of student maintenance and progress (Saville, Lambert, et al., 2011).

f) The final element of interteaching is an interdependent group contingency. All students in separate discussion groups can earn points if all members contribute to a high-quality discussion. Students can use points toward testing probe grades. The immediacy of quality point delivery as feedback was notable for promoting student participation and learning (Rosales, Soldner, & Crimando, 2014).

Overall, interteaching is an instructional approach that diverges from the lecture method and includes a package of procedures to create classroom contingencies that rely less on self-management strategies to promote student success (Querol et al., 2015). The interteaching teaching package emphasizes active student-driven learning, peer teaching, class preparedness, instructor facilitation, frequent tests for master, use of positive-reinforcement strategies, and immediacy of feedback on student learning (Querol et al., 2015).

Interteaching with CALs

Interteaching is a relatively newer behavioral teaching method, with research occurring in the last two decades. The majority of research has been conducted with CALs, as this method was primarily designed for higher education (Querol et al., 2015).

In a comprehensive review of the interteaching literature with CALs (Querol et al., 2015), 25 articles met inclusion criteria. The studies were evaluated based on participants' demographics, settings, experimental designs, dependent variables manipulated, academic discipline, and social validity. Compared to traditional lecture format (see Querol et al. for examples of these studies), interteaching demonstrated effects across various dependent variables including academic achievements in many forms (i.e., homework, participation, probe grades, final exam scores, and maintenance tests). Interteaching also demonstrated effects across various disciplines: computer programming, political science, nutrition, special education, nutrition, social welfare, and several psychology courses (i.e., introductory, abnormal, research methods, and behavior analysis). The few studies that conducted component analyses emphasized several components were critical for student success: 1) distribution of quality points in group contingencies (i.e., Saville & Zinn, 2009), 2) the clarifying lecture (Saville, Cox, et al., 2011), 3) preference assessments for discussion sizes (Truelove et al., 2013), and 4) frequent test probes (Felderman, 2014).

There is significant empirical support for the use of interteaching with CALs for improving academic achievement across various disciplines, with high social validity and student preference for interteaching over lecture (Querol et al., 2015). Interteaching incorporates several important learning principles for improving student performance, such as shaping (i.e., study guide questions progressively alter from easy to challenging) and reinforcement contingencies (i.e., quality points for group discussions). There are several recommendations for interteaching pedagogy. For example, future research could look at examining an adaptation of interteaching for distance education and blended online learning courses (Querol et al., 2015). Also, there is significantly less interteaching research devoted to component analyses of this instructional package. Since this method is relatively new, further research could confirm the critical components necessary for improving student learning and outcomes.

Equivalence-Based Instruction (EBI)

Equivalence-based instruction (EBI) is a teaching method derived from stimulus equivalence principles, which aims to teach academically relevant concepts. EBI developed from Sidman’s work with children with developmental disabilities (Rumbaugh, 1995) in which he discovered the emergence of equivalence relations (see Equivalence Relations section). For example, a learner establishes stimulus equivalence after the relations, A-B and

A-C, result in the student learning B-C and C-B without direct instruction. With the mathematical set theory application, Sidman later described the emergence of novel stimulus-stimulus relations in which each distinctive stimulus is jointly paired (stimulus equivalence; Sidman, 1994). Studies in basic behavioral research and ABA helped researchers further discover the principles governing equivalence relation formation and how to apply these concepts to socially significant behavior (Rehfeldt, 2011).

Subsequently, researchers used these principles to design curricula for CALs, which address instructional issues (e.g., aversive control in the lecture format) and increase efficiency for improving course study with low-performing students (Michael, 1991). EBI uses match-to-sample (MTS) procedures to teaches students to treat separate stimuli as functionally interchangeable by training overlapping conditional discriminations (Fienup et al., 2011). A conditional discrimination is one in which reinforcement provided for a response to a stimulus is dependent on the presence of other "conditional" stimuli (Cooper et al., 2007). In EBI, conditional discriminations are taught in order and to mastery. For example, instructors begin with teaching the baseline overlapping relations and observe the emergence of derived relations, including symmetry, transitivity, and equivalence (Sidman, 1994). A learner who has mastered all baseline and derived relations is said to have formed an equivalence class (Green & Saunders, 1998). Most research conducted in EBI has been with CALs and several studies that have produced many gains for this population are discussed below.

EBI with CALs

EBI Meta-Analyses Research. Brodsky and Fienup (2018) conducted the first meta-analysis of EBI in higher education, which is a review of several independent studies in the same subject to examine overall trends. About 28 studies met criteria for inclusion, and all used either group designs or single-subject designs. About 48% of the experiments assessed interobserver agreement and 13% took treatment integrity data. Publications occurred in various peer-reviewed behavior-analysis and educational journals (i.e., *Journal of Applied Behavior Analysis, Journal of Behavioral Education,* etc.). The authors reported the first EBI publication occurred in 1989, about 18 years after Sidman's seminal article on equivalence relations (Sidman, 1971). Roughly 65% of the studies were conducted on academically relevant learning in highly controlled laboratory settings. In contrast, the other studies reported EBI protocols implemented in higher education curricula. A significant portion of the studies conducted with CALs occurred in the last decade. The studies measured various response topographies and evaluated EBI effects on many dependent variables, and most studies used computer-based or written format MTS procedures (Brodsky & Fienup, 2018). EBI's impact in terms of percentage of correct responses on equivalence class formation or defined efficiency in the number of trials or time required to master equivalence classes (Brodsky & Fienup, 2018). The authors aimed to answer several questions about the EBI research conducted: 1) "Is EBI effective?", 2) "Are some EBI variations more effective than others?" and 3) "Is EBI more effective than other alternative instructional methods?".

To answer the first research question, Brodsky and Fienup (2018) examined the 23 studies that evaluated the effects of EBI compared to baseline or control. Group design studies with both within- and between-subject measures, found statistically significant results of EBI pre- and post-test scores compared to no-instruction controls. Large effect sizes were demonstrated in single-subject design studies, as well. Results of the second research question indicated few studies compared variations of EBI with between-subject controls and found low statistical significance in EBI variations and training sequences. However, one study (Fienup et al., 2015), which compared simple-to-complex and simultaneous training protocols, found they were more effective with four-member classes than three-member classes. The investigation results in their third research question indicated there aren't enough studies comparing EBI to alternative instructions to determine which is more effective (Brodsky & Fienup, 2018). EBI is a sufficient method for teaching equivalence relations for training small sets of stimuli across various disciplines (i.e., statistics, mathematics, neuroanatomy, music) (Brodsky & Fienup, 2018). However, more research is needed to examine EBI's effectiveness for teaching the full range of stimuli students are required to learn in an entire college semester course. Most EBI research conducted occurred in highly controlled laboratory settings; therefore, it falls more under basic behavioral and translational research (Brodsky & Fienup, 2018). However, several more recent EBI studies have been conducted in more naturalistic or 'applied' settings (Greville et al., 2016; O'Neill et al., 2015; Pytte & Fienup, 2012; Varelas & Fields, 2017).

EBI Research in Applied Settings. For instance, in a recent, more applied study, Ramos et al. (2018) evaluated EBI's effects on teaching basic music reading and piano skills to CALs. The authors reported that few studies examined the effects of EBI on teaching leisure skills. Fewer studies focused on using this method to teach music reading and piano skills (Arntzen et al., 2010; Hayes et al., 1989; Perez & de Rose, 2010). Since music instruction involves establishing classes of auditory (e.g., sounds of notes) and visual stimuli (e.g., written notes), and teaching students to play an instrument in these stimuli, EBI could be beneficial for teaching beginner-level music reading and piano skills (Ramos et al., 2018). This study tested a simple EBI procedure, MTS, for teaching basic piano skills to six female college students with no prior exposure to music education. The authors used a two-tier nonconcurrent multiple-baseline-across-participants design, to evaluate EBI procedure effects on the percentage of trials with correct responses during training and testing phases. Students were exposed to a simple visual-visual MTS task with common stimuli (i.e., elephant, balloon, apple) to familiarize them with the procedure and the computer software in the training phase. The music testing protocol involved a series of visual-visual and auditory-visual MTS tasks and probes for vocal responses and piano playing. The authors tested all possible relations before experimentation to rule out prior learning and to ensure students could not correctly select either of the following relations before training: the textual chord symbol for the musical symbol (e.g., BC) or the reverse (e.g., CB), select either the correct textual or musical notation when presented with the spoken name (e.g., AB, AC), vocally label the textual chord or musical symbol (e.g., BD, CD), nor play the correct chord piano keys when presented the spoken, textual, or musical notation of the chord (e.g., AE, BE, CE). To assess for equivalence formation, the authors re-evaluated the relations described above following MTS and piano training (Ramos et al., 2018, pp. 218). Mastery criteria for training and post-tests were 89% correct responding or above. The authors implemented piano playing post-tests and novel sequencing post-tests to assess for generalization of mastered musical and textual chord representation in song form (e.g., Amazing Grace).

Results of the pretests in Ramos et al. (2018) indicated all students scored below 56% on all relations except for the attending to verbally spoken chord, textual written chord, or vocally labeling (e.g., AB and BD relations). Although students scored higher on these relations (89% or above), the authors expected this response pattern because it is similar to MTS with letters of the alphabet. Results of the training and post-tests indicated that the procedure was sufficient for all students in training relations, including in auditory-visual (AB, AC, AB/AC mixed), visual-visual (BC, CB), and vocal (BD, CD) tests. Subsequently, all students met mastery criteria for the piano tests (AE) and transfer function test (BE, CE). Only one student required remedial training in the vocal, piano, and transfer function tests. However, the student quickly met the criteria following additional training. Therefore, all six students learned to select the textual and musical symbols of three different chords when they heard each chord's spoken name. Additionally, the students learned to play the chords when they heard or saw the textual or musical notation.

There are several strengths to this study. Ramos et al. (2018) used a multiple baseline design to demonstrate experimental control and confirm EBI procedures, MTS training, and remedial training were responsible for establishing equivalence relations among all students. The authors systematically replicated previous studies (Arntzen et al., 2010; Hayes et al., 1989; Perez & de Rose, 2010) and extended the findings by demonstrating that trained relations using EBI could also result in socially significant outcomes, such as students learning to play a song. However, they failed to systematically fade programmed consequences in training to maintain consistent responding despite the absence of reinforcement (Ramos et al., 2018). While this did not affect any of the students' performance of equivalence relations, future replications with CALs, and especially with children, should take considerations. Not all student responses will maintain in the absence of reinforcement without incorporating fading procedures.

After reviewing the available EBI research, there are several limitations to discuss. Most research takes place in highly controlled settings (i.e., experimental laboratories) and less in applied naturalistic settings (i.e., traditional classrooms, online learning, distance education). Additionally, EBI requires instructors to have significant training in teaching equivalence relations, and the time needed to develop courses incorporating EBI procedures is sufficient (Brodsky & Fienup, 2018). However, the present body of basic behavioral and translational research studies demonstrating the effectiveness of EBI improving student performance (i.e., Fienup et al., 2010), provides evidence that this method can be useful in traditional classrooms. More researchers have started to evaluate this teaching method in these settings (i.e., Ramos et al., 2018). However, to further provide empirical support for EBI in applied settings, researchers should collaborate with educators outside of ABA to examine adaptations and simplified versions of EBI to meet traditional classroom needs.

Personalized System of Instruction (PSI)

A personalized system of instruction (PSI), also known as the "Keller Plan," is a behavior analytic teaching model composed of small, self-paced units of instruction (i.e., "modules"). Study guides or teaching assistants help direct learners through the modules until they have mastered all the material (Keller, 1968; Twyman, 2014a). This mode of instruction relies on the teachers to design, individualize the teaching program for each student, and redesign the program in response to the student's performance on successive modules (Keller & Sherman, 1974). The learner cannot move forth through the PSI modules until they demonstrate mastery on each successive level, and the student's successful performance is then "reinforced" with access to the following material (Keller & Sherman, 1974). The learner contacts this contingency once they master the first module and access the next unit. The module-based learning also relies on the principles of equivalence relations and concept learning discussed in-depth earlier in this paper. By breaking up the content of a subject matter into smaller, discrete units or modules, and learning the necessary prerequisites, concept formation and the establishment of equivalence relations can occur (Keller & Sherman, 1974; Sidman, 1997; Layng, 2013). As described above, the initial PSI design contains five key components: unit mastery, self-pacing, lectures for motivational purposes, emphasis on the written word, and the use of student proctors (Keller & Sherman, 1974, p. 24).

Component analysis research in PSI over the past half-century contributed to adaptations, procedural adjustments, and identification of feature components critical for student success and academic achievement (Fox 2004, 2013; Eyre, 2007). For example, researchers have suggested PSI's key components be updated and redefined to compromise for administrative restrictions, traditional course semester constraints, and continuous advances in technology (Fox, 2004). These include unit mastery, flexible-pacing, on-demand course content, immediate feedback, and peer-tutoring (Fox, 2004).

One of the most widely recognized PSI adaptations is the computer-aided personalized system of instruction (CAPSI; www.webcapsi.com). CAPSI is an internet-based program that follows the same PSI tenets, which researchers developed at the University of Manitoba, where ongoing research continues today (Kinsner & Pear, 1988; Pear & Kinsner, 1988). In a CAPSI course, the units of study are set up by instructors, covering course reading and the curriculum with several study questions that correspond with each unit (Pear & Novak, 1996). Students enter the CAPSI system and take unit tests on the course readings and objectives at their own pace. During unit tests, the computer randomly selects several short-answer questions from a question bank for students to answer. Once completed, unit tests are marked for mastery by peer-reviewers who have already mastered the unit, student proctors who have completed the class, or the instructor. Besides CAPSI, there are additional online course platforms with built-in software that allow instructors to structure courses using a PSI component, such as WebCT® and Blackboard® (Chase, 2006). These programs can set up mastery-based multiple-choice unit quizzes or tests, in which instructors can set a minimum percentage for students to meet before they obtain access to the next unit tests. Instructors can also set maximum attempts on unit tests to achieve mastery (e.g., three attempts) or set deadlines for mastery achievements (e.g., bi-weekly). Other programs, such as ALEKS® (www.aleks.com/about_/aleks/overview), employ PSI principles and can be used to set mastery-based unit tests in mathematics and statistics (Eyre, 2007).

Research in PSI and functionally equivalent instructions (i.e., CAPSI) has investigated various empirical topics within the past half-century. These include PSI or CAPSI's overall effectiveness compared to the traditional lecture format (Kulik et al., 1979; Gifford & Vicks, 1982); component feature analyses (Kulik et al., 1978); differential effects of the immediacy, accuracy, and quality of proctor feedback (Haemmerlie, 1985; Pear, 2003; Martin et al., 2002a; Martin et al., 2002b; Chase & Houmanfar, 2009); and the effects of

PSI/blended courses (Svenningsen & Pear, 2011). This research has explored PSI's effectiveness in improving academic achievement (e.g., Springer & Pear, 2008), among other dependent variables, across various settings, disciplines, and populations (Fox, 2013). However, the majority of PSI research conducted has been with CALs (Eyre, 2007). To understand why PSI research focuses so heavily on CALs, it is important to understand the history of PSI.

History of PSI

A history of PSI begins with its founder, Fred S. 1968 who worked closely with J. Gilmour Sherman and other colleagues to formulate the instructional design as it is known today. The inception of PSI can be traced to a meeting in 1963 with several behavioral psychology researchers including Keller, Sherman, and Rodolpho Azzi and Carolina Martuscelli Bori from the University of São Paulo in Brazil (Keller & Sherman, 1974). The four investigators were invited to work together to develop the Department of Psychology at University of Brasilia and began developing a new plan of instruction (Keller & Sherman, 1974). They all agreed the instructional plan would include learning principles (e.g., reinforcement, extinction, punishment, discrimination, stimulus control, conditioned reinforcers, conditioned punishers, generalized reinforcers, and shaping), and use didactic procedures that operate under operant conditioning and reinforcement theory (Keller & Sherman, 1974, pp. 50-58). Martuscelli Bori, Azzi, and Sherman were all advisees to important behavioral researchers such as Ferster, Herrnstein, Lindsley, Schlosberg, Sidman, and Skinner, among others. Martuscelli Bori and Azzi had previously taken courses in Natural Science with Skinner at Harvard and Behavioral Technology with Ferster at the Institute for Behavioral Research (IBR); they planned to use the knowledge from these courses to inform their course design.

Initially, all four researchers were convinced that traditional teaching methods were out of date. Therefore, they designed what was promised to be “one of the most radical courses ever given in a university setting” (Keller & Sherman, 1974, p. 7). This course was to be a combination of multiple methods: laboratory teaching methods employed in psychology courses at Columbia, methods used at IBR, the use of the PI approach where possible, and the use of the five components discussed earlier (Keller & Sherman, 1974).

The theoretical account of the PSI course design outlined by Keller and Sherman (1974) was essentially a successive approximation to the course that took place at the University of Brasilia in 1964. In this course, the emphasis on lab work decreased and the proctor, or student teacher, was added. This design did not eliminate final exams and letter grades completely. However, the self-pacing feature, the performance requirement for advancement, the role of the lecture, and the general principles of the course remained in-tact. Social validity reports concluded that the students who participated successfully completed the first two parts of the course. All students reported that the self-pacing aspect of the course was their favorite part. The least desirable feature was the lack of opportunity to discuss materials with head professors rather than student proctors (Keller & Sherman, 1974). The pilot course discussed above was the initial implementation of PSI. It became a catalyst for thousands of successful PSI courses and hundreds of research studies evaluating the instructional method in just the 1970s alone (Fox, 2004, 2013).

The Fred S. Keller School. One of the schools that Keller developed is located in Yonkers, New York and operates under a Comprehensive Application of Behavior Analysis to Schooling (CABAS®) (Greer, 1980) program (Twyman, 1998). The CABAS® schools are self-correcting and self-sustaining and derive from similar principles underlying PSI and incorporate the science of teaching in every aspect of schooling. The first Keller School opened in 1986, and now there are eight different schools in the U.S. and one in the U.K. that use the CABAS® model (www.cabasschools.org). The Keller School functions as a technology-driven system of instruction that relies on the individualized instruction (i.e., PSI) of each student to influence the behavior of the entire education community. Student progress is continuously monitored, measured, graphed, and analyzed to inform necessary adjustments. Staff and parent training programs are also rooted in a behavior-analytic approach and occur frequently to adjust with the ever-changing academic performance of the students. Supervisors are well versed in behavior analysis and a scientific approach to teaching, which help ensure high treatment integrity.

Early PSI Research from 1970-1999 with CALs

Research studies between the early 1970s and late 1990s in PSI with CALs yielded positive results. Early PSI studies focused on several different areas of research. These include studies evaluating the differential effects of PSI compared to the traditional lecture method, PSI component analyses, studies addressing procrastination in PSI courses, studies examining the effectiveness of proctor feedback, and research on the development and efficacy of CAPSI. Several studies in each of these areas of PSI research will now be considered.

PSI Compared to Lecture

By 1979, there were over 3,000 papers, articles, and reports published on PSI with over 5,000 courses using PSI-based instructional designs (Sherman, 1982). A significant portion of the early studies in PSI researched and compared the differential effects of PSI and lecture course designs on student performance. Kulik et al. (1979) conducted a meta-analysis on outcome studies in PSI. The authors found 61 of the 75 studies reviewed, compared the effects of PSI and traditional lecture-formatted courses. Of the 61 studies that compared PSI and lecture courses, 57 reported the highest final examination scores using PSI. Only three studies' results favored conventional instruction and one study showed the average final exam scores were equivalent. Overall, the authors reported PSI is beneficial across various course settings with many different research designs. Kulik et al. (1979) is a seminal article in PSI research for its' most significant finding: when compared to conventional instruction, PSI raised final exam scores from the 50th to the 70th percentile. Researchers within and outside ABA continued to investigate the effects of PSI compared to lecture formats on the social validity and academic achievements of CALs.

For instance, Gray et al. (1986) conducted a two-experiment study in an introductory speech communication college course. The researchers used a group design to evaluate and compare the effects of adapted PSI-based course sections and lecture-recitation formatted sections on students' social validity and course material mastery. Both experiments indicated PSI-based formats were equivalent in some measures but generally more effective than lecture-recitation sections. The social validity measure showed students and instructors were more satisfied with the overall quality of the course's PSI-based sections. The PSI students exhibited higher overall grades on their final speeches, final examinations, and course grades. Other indirect measures indicated that PSI students reported less anxiety and an overall improvement in communication skills. This study had several limitations. The authors did not take any follow-up data. Therefore, they were unable to demonstrate whether PSI contributed to students' maintenance of the course material. They also relied heavily on statistical data, which makes it difficult to assess for variability in individual performances. However, higher final speech grades, exam scores, and course grades, and higher social validity scores of student and instructor satisfaction of the PSI-based format's overall quality, support the implications for using this approach in college speech and communication courses.

Jumpeter (1985) conducted a group design study that evaluated the pre- and post-test scores of 36 freshman and sophomore students enrolled in a university's music appreciation course. The author randomly assigned students to either the control group, a lecture format course, or to the experimental group, a PSI-based course. Results indicated no difference in their listening skill improvements between the PSI and control groups. PSI students reported the Learning Activity Packets to be more restrictive and time-consuming than attending to lecture notes or listening to a musical recording. While the study results did not demonstrate that PSI was more effective than the lecture format in improving the listening skills of college students, it did prove to be as effective. Therefore, PSI might be a valuable alternative teaching method for this material. The students in the lecture-formatted courses gained more from the course material potentially because students had a longer history learning in this manner (Jumpeter, 1985). Lecture notes supplemented study materials (e.g., textbooks) and may have contributed to students' mastering objectives more quickly than students in PSI sections (Jumpeter, 1985). Students in PSI sections potentially required more time to master objectives without lecture notes due to the self-paced feature (Jumpeter, 1985). Future research could examine the application of deadlines to increase motivation for timely course completion achievements and reduce procrastination (Jumpeter, 1985). A limitation of this study was the author did not indicate which mastery criteria were set for students to obtain on Learning Activity Packets before gaining access to the following material. Therefore, it is impossible to determine whether or not the author set sufficient mastery criteria for ensuring students did indeed master the material before moving through the course.

In another study, Gifford and Vicks (1982) evaluated the differential effects of PSI compared to a traditional lecture course for teaching biology to junior college freshman. This was the first study to examine PSI's impact at the junior college level, or post-secondary schooling for vocational skill training at a private junior college. The authors compared the pre- and post-test scores of the experimental group (PSI) and control group (traditional lecture). Results indicated the experimental group made significantly more gains in the post-test scores than the control group.

A significant amount of research has found PSI to be more effective than the traditional lecture method. However, another area of research examines PSI components to determine which is critical and how to maximize the effects of certain features.

PSI Component Analyses

Another groundbreaking article published in the *Journal of Personalized Instruction* was Kulik et al.'s (1978) comprehensive review of studies that conducted component feature analyses of PSI. There were several experimental problems that occurred in this area of research for future researchers to consider (Kulik et al., 1978). These include nonequivalence of comparison groups, inadequate experimental treatment, inappropriate measurement outcomes, and methodological conclusions. However, there were many valuable findings that helped guide future PSI studies and course designs (Kulik et al., 1978; Fox 2013). The unit mastery requirement in PSI consistently led to higher levels of student achievement. However, the number of proctor-student interactions did not have a significant effect on final exam scores; in fact, Kulik et al. (1978) suggested the fewer interactions, the better. Also, the authors reported immediate and more descriptive feedback improved student performance. While delaying feedback interfered with students' maintenance of course material, some studies showed delayed feedback was superior for some learning tasks. The authors found review units enhanced learning and student performance, and lectures were not necessary for student success. An essential finding was the self-pacing feature wasn't necessary for student achievement. Additionally, course policies that limit self-pacing and provided reinforcement for student progress successfully reduced procrastination while also lowering PSI withdrawal rates.

PSI Studies on Procrastination

PSI often receives criticism for contributing to low or zero student responding rates in initiating or progressing through course material (Kulik et al., 1974), which raises some ethical concerns. The self-paced feature can lead to the clustering of student completion requirements towards the end of the course, which can negatively impact academic performance due to potential stress increases, time-management issues, and problematic behaviors like procrastination. In pedagogical research, evenly distributed learning is preferred over heavily concentrated learning practices (Underwood & Ekstrand, 1966). An area of research in PSI explores ways to reduce procrastination.

Deadline Contingencies in PSI. Specifically, Ross and McBean (1995) investigated the use of deadlines for reducing procrastination in PSI-based college courses. The authors compared the differential effects of several deadline contingencies on students' unit-mastery test-taking behaviors. In four university classes using a PSI model, the authors implemented multiple and single deadline contingencies which involved issuing penalties for missing unit-mastery deadlines. They measured student test-taking behaviors in four 13-week courses; courses A and B were introductory behavior modification courses, and courses C and D were advanced behavior modification courses. All four courses used a PSI model, required students to work independently on material, and complete three review tests to mastery before proceeding to successive unit test. Test retakes contained question banks with new exemplars and students repeated unit or review tests until they met mastery criteria (80% or above). The single and multiple deadline contingencies used consisted of a variable-interval, fixed-interval, fixed-interval (VI FI FI) schedule sequence in course A, a VI FI VI sequence in B, and a VI schedule in courses C and D. Results indicated the least variability in test-taking rates were found during the VI schedules in courses C and D, which set multiple deadlines throughout the course.

In conclusion, this study demonstrated it is necessary to implement test-taking contingencies throughout time-sensitive courses to maintain student test-taking behaviors in PSI courses. This study also supports the implication of multiple contingencies in PSI courses to increase quiz and test-taking responses. Ross and McBean (1995) demonstrated that deadline contingencies could lead to more even distribution in students' course completion, effectively reducing student procrastination.

Pacing Procedures in PSI. Reisner (1984) evaluated three different pacing procedures on student withdrawal rates, rate of progress, final exam scores, and social validity measures in a PSI-based speech communication course for undergraduate students. The author used a group design with two experimental groups and one control in which all groups were in PSI courses but exposed to different pacing procedures. Students in one group, the reward group, earned two additional points per instructional unit mastered by the suggested deadlines. While students in another group, the penalty group, lost two points for each instructional unit they failed to master by the suggested deadlines. The final group, the control group, neither gained nor lost points for meeting or failing to meet suggested mastery deadlines. The results indicated no statistical difference in withdrawal rates between the groups. There was a statistically significant difference in the rate of progress between students in the penalty group compared to the control group, indicating slightly higher rates for the penalty group. However, there was very little difference between the reward and penalty groups' rate of progress. There was no difference in the final exam scores between groups, and social validity scores were high for all groups. This study demonstrated that procrastination in PSI courses could significantly reduce with deadlines without impairing social validity scores or final exam performance.

Feedback Research in PSI

Limiting PSI's self-paced feature and using more "flexible pacing" like Fox (2004) suggested is just one method researchers found for improving the effectiveness of PSI components. Another component researchers examined to maximize PSI's efficiency is the feedback provided on unit tests. More recent PSI research examined various aspects of feedback, such as immediacy, accuracy, and quality. Most early research on PSI feedback examined the differential effects of immediate and delayed feedback on student performance. One study found interesting results worth considering.

In an early feedback study, Haemmerlie (1985) conducted two experiments that examined the effects of various feedback intervals in a PSI course on the performance, maintenance, and preferences of CALs. Based on previous research, the author hypothesized that student performance would improve if feedback functioned as a reinforcer. Prior studies in applied teaching techniques indicated immediate feedback in PSI as empirically valid (Kulik et al., 1978). There was limited research evaluating the timing and frequency of feedback (Haemmerlie, 1985).

The first experiment in Haemmerlie (1985) compared the effects of feedback provided after-each-item and feedback given after-the-exam. The mastery criterion for unit tests was 90%, and students repeated tests until they met the requirements. The feedback type provided was dependent on the unit tests taken. Half of the participants were randomly assigned to receive one feedback condition, while the other half received the alternative feedback. In the after-each-item feedback condition, the proctor indicated if the answer provided was "correct" or "incorrect" immediately following the student's response. In the after-the-exam feedback condition, proctors provided answers in the same manner, except the student received all answers following completion of the unit test. Results in experiment one indicated students made more errors in the after-each-item condition compared to the after-the-exam condition. The number of preferences for the after-the-exam feedback was only slightly higher than for after-each-item feedback; however, 50% of students claimed the after-each-item condition disrupted performance when feedback indicated they responded incorrectly.

For experiment two, Haemmerlie (1985) used the same design, except the students were alternately exposed to four different feedback conditions: after-each-item feedback, after-the-exam feedback, overnight delay, and a minimal feedback condition. The order of conditions was the same for all students. The after-each-item and after-the-exam feedback conditions were the same as the first experiment. The overnight delay condition was the same as the after-the-exam condition except instead of directly following the session feedback was provided after a 24-hr delay. In the minimal feedback condition, the control condition, proctors scored exams in another room and only indicated whether students met the 90% mastery criterion. Results in experiment two indicated more errors occurred on first attempts, and more attempts overall occurred in the after-each-item feedback condition. Also, students' performances on maintenance tests were lowest in the after-each-item feedback condition. Students indicated higher preferences for the after-the-exam feedback and lower preferences for the minimal feedback and after-each-item feedback conditions in social validity measures. Additionally, students reported that missing items reduced their motivation and increased tension and anxiety in the after-each-item feedback condition.

A consistent finding throughout the Haemmerlie (1985) study was that the after-each-item feedback had a significantly more negative impact on student performance than with any other feedback types, including the control condition. The take-home point of this study is that not every form of feedback will function as reinforcement for all students, and some forms of feedback can potentially function as punishment. Therefore, researchers and instructors should take student feedback preferences into account when designing PSI courses.

PSI as a Component for Teacher Training

Another area of early research examined the effectiveness of PSI as a component for teacher training. As discussed earlier, The Keller School (Twyman, 1998) and others like it operate under a CABAS® approach, incorporating PSI as a significant component for training teachers to implement behavior-analytic teaching procedures.

For instance, in an extensive two-year study, Selinske et al. (1991) evaluated a CABAS® approach on several dependent variables in a small school including total and correct student trials taught by each of the four teacher groups and the number of objectives achieved. The authors also examined reliability measures for treatment package implementation, which included supervisor tasks and quizzes and PSI modules completed to mastery by teachers. The CABAS® treatment package involved applying behavior-analytic procedures and principles to the performance of students, teachers, and supervisors. The authors implemented the treatment package in a school for children with developmental disabilities. The treatment package included: a staff training program based on PSI, organizational behavior management (OBM) procedures for supervisors, regular assessment of teacher behaviors, and teacher assessment of all instructional trials of 38 children with individualized curricula designed to meet each child's needs.

This study's (Selinske et al., 1991) major strength is the authors' use of a multiple-baseline-across-participants design to examine the functional relationship between CABAS® and the performance of four groups of teachers and their students. There were three phases: baseline, training, and full treatment. Results for all groups of teachers in the PSI-based training course indicated they met the criteria of completing a set of 10 modules to earn $1,000 bonuses. Subsequently, all teachers earned bonuses by the following year of the study, a contingency set by the authors for meeting PSI module completion criteria. Data on supervisors provided an index on the reliability of treatment implementation. The authors found the number of weekly tasks performed increased systematically as the number of teacher groups in treatment increased. The experimental student performance data showed that increases in trials taught, correct responses, and objectives met were a function of implementing the CABAS® treatment package. The effects of treatment maintained in follow-up data.

One relevant limitation of Selinske et al. (1991) is that the authors did not evaluate whether teacher completion of the PSI modules or quizzes were an essential component for improving student performance. Also, each module required significant time and response effort. Therefore, the effectiveness of bonuses given for module completion might have been diminished due to the delay in delivery. Future research might examine the effectiveness of increasing the immediacy of bonus delivery following module completion, by dividing the bonuses into smaller rewards.

Early Research in CAPSI

Development of CAPSI. The computer-aided personalized system of instruction (CAPSI) was pioneered at the University of Manitoba, Canada, by Joseph J. Pear and W. Kinsner in the early 1980s. Their initial research focused primarily on the development of CAPSI, which was strongly influenced by the researchers' experience with PSI, and the existing body of research indicating PSI as practical, if not more effective than traditional teaching methods (Pear & Kinsner, 1988; Kinsner & Pear, 1988). The researchers found CAPSI to be a highly versatile method of instruction in that it could be used in numerous ways and across various courses. In their research developing the instructional practice, they found computers to be practical tools for elevating and improving education efficiency, which they referred to as "catalyst(s) for design-oriented learning" (Pear & Kisner, 1988, p. 214). Pear and Kinsner (1988) reviewed their final CAPSI course design and reported successfully using the method to teach five different undergraduate courses taught over 24 times to 650 students at the University of Manitoba.

Pear and Kinsner (1988) discussed how CAPSI could be used effectively and economically in distance-learning and traditional classes. The CAPSI design was discussed earlier in this section. In CAPSI, instead of proctors being students who have previously passed the course, unit tests are graded by student proctors in the present course who have already mastered them. The structure of CAPSI provides instructors with easy access to datasets for later analysis, including unit tests with instructor and proctor feedback. The authors suggested instructors periodically check proctor tests for errors, and provide feedback when necessary, which serves multiple functions: to ensure proctors offer high-quality feedback and check proctors' maintenance of the mastered material. Ultimately, CAPSI is an effective teaching method founded on PSI tenets that allow instructors to teach more students without compromising their quality of teaching or the students' learning.

Kinsner and Pear (1988) elaborated on the types of datasets the researcher found available for analysis in CAPSI courses. As mentioned before, an essential feature of CAPSI is that data saved describes interactions that occur during the course, which can be accessed and analyzed at any time. These data include marking transactions, test results, date, and time. In this article, the authors provided example graphs of several data sets from a previous CAPSI course they implemented. For example, they presented the number and percentages of students' canceled tests, unit test passes, conditional passes, and restudy tests. Kinsner and Pear (1988) explained the complete accessibility of testing and marking interaction records in CAPSI allows instructors and researchers to rigorously observe, analyze, and assess a significant portion of behavior and learning in the course. The early work of Pear and Kinsner (Pear & Kinsner, 1988; Kinsner & Pear, 1988) was groundbreaking and steered the way for a whole new area of research for investigators interested in PSI and CAPSI alike.

Efficacy of CAPSI. In another study at the University of Manitoba, Pear and Novak (1996) conducted a program evaluation of two undergraduate psychology courses taught using CAPSI. A CAPSI approach combines computer-assisted instruction (CAI) in which the computer provides direct teaching, computer-managed instruction (CMI) in which the computer structures learning tasks and lessons, and PSI. The program evaluation involved social validity measures, such as student evaluations and descriptive accounts of their experiences and analysis of performance determinants in the courses, including passes, conditional passes, restudy requirements, and final exam scores. Students received a pass result if they answered all three-unit test questions thoroughly and correctly. They required a restudy (retake) if they answered any questions incompletely or incorrectly. Finally, students received a conditional pass if clarification of some answers were necessary. Additionally, the authors only used data from students who completed pre-course and post-course questionnaires and set two criterion variables: "participation" and "achievement", as computers maintained proctoring records and marks on tests and exams. The authors operationally defined "participation" as the amount of proctoring a student engaged in measured by total proctor points obtained, and "achievement" as students' grades on the final exam.

The findings of Pear and Novak's (1996) CAPSI program evaluation on social validity measures indicated students reported the structure of the course design was highly convenient. While students were mildly dissatisfied with the requirement to use computers, they were comfortable using them, and overall, would retake the course. The regression analyses on proctoring and final exams showed performances in CAPSI did not differ much from traditional course types. The authors concluded that students that have done well in other course methods would do well in a CAPSI course. While it would have been interesting to see the display of individual data in a single subject design further to analyze the effects of CAPSI and variability across individuals, this study aimed to determine whether CAPSI was an effective alternative to the traditional lecture method.

Crosbie and Kelley (1993) evaluated the effects of CAPSI on the academic performances of 51 students in an introduction to applied behavior analysis course with minimal background in behavioral psychology. Previous research indicated that PSI courses are challenging to maintain without proctors (Green, 1974). The CAPSI design in this study incorporated the traditional PSI components except they eliminated the use of proctors by employing an automated computerized software for scoring unit tests, providing feedback, and keeping all records in the course. The authors set unit test mastery criteria to 90%. To prevent rote memorization on unit test retakes, the computer software was programmed to present multiple exemplars of the unit material. Presentation of numerous question models in unit test retakes served several functions; to improve students' ability to attend to subtle discriminations and promote generalization of material. Students were also required to take four review tests once they completed the preliminary prerequisite unit tests. Additionally, the study included an external-pacing structure to apply assigned deadlines for the four review tests. This pacing contingency involved students receiving 3% towards their grade when they mastered review tests by target dates, and 1% deductions occurring for each day between deadlines and review test mastery dates.

Results of Crosbie and Kelley (1993) indicated all but two students completed the course, met the review test deadlines and met mastery on the first attempts, and therefore received the 3% bonus. Student overall grades were based on the cumulative scores of their unit tests, review tests, and assignment grades. The authors indicated the grade distribution was negatively skewed which is typical of PSI courses (e.g., Croft et al., 1976). Social validity measures were high, and students reported enjoying the CAPSI course structure, despite using a computerized version of proctors.

One limitation of this study was that it was difficult to evaluate the deadline contingencies' effectiveness because it lacked comparison to a control group or comparison of conditions in which individual students act as their own control. Regardless, this study adds to the existing PSI and CAPSI research and provided evidence for an effective computerized alternative to proctors. The course utilized computers as a reliable and cost-efficient way to test, score, and keep up with PSI records.

General Advantages, Disadvantages, & Considerations of Early PSI Research

To summarize, early PSI studies examined various areas of research. Researchers found PSI to be at least, if not more, effective than conventional teaching methods for improving the student academic performances of CALs (Kulik et al., 1979; Gray et al.,1986; Jumpeter, 1985; Gifford & Vicks, 1982). Early component analysis research made significant findings that directed later research, including the strict self-pacing feature and proctor-student interactions, as not necessary for student achievement (Kulik et al., 1978). Some studies examined the efficacy of deadline contingencies and different pacing procedures in PSI courses and found them successful for reducing procrastination, improving course completion, and decreasing withdrawal rates (Ross & McBean, 1995; Reisner, 1984). Other studies explored the role of feedback in PSI, providing evidence that immediate feedback provided directly after each item on unit tests may not function as reinforcement for students and may negatively impact performance (Haemmerlie, 1985). A smaller research area found PSI to be a practical component in teacher training treatment package in a CABAS® approach to a school for children with developmental disabilities (Selinske et al., 1991). Additionally, early research focused on the development of CAPSI and found it a compelling adaptation of PSI (Pear & Kinsner, 1988; Kinser & Pear, 1988). CAPSI has much to offer data-driven pedagogical research due to the nature of the instructional method’s ability to record various student behaviors for instructors to monitor, analyze, and evaluate any time (Pear & Novak, 1996; Crosbie & Kelley, 1993).

Overall, early researchers found PSI and CAPSI useful for improving academic achievements across various disciplines such as human growth and development (Lu, 1976), economics (Specter, 1976), writing skills (Allen, 1985), biology (Gifford & Vicks, 1982), Medical School biochemistry (Weisman & Shapiro, 1973), introductory psychology (Pear & Novak, 1996), applied behavior analysis (Crosbie & Kelley, 1993), among many others. Early researchers also found PSI an effective instructional method for improving socially valid behaviors and vocational skills, such as reducing Navy Training dropout rates (McMichael et al., 1976) and teaching checking account skills to adults (Zencius et al., 1990). Another essential finding in early PSI research provided empirical support for assessing student preferences of different course design procedures for improving social validity in PSI courses (Bono & McAvoy, 1977).

Recent PSI Research from 2000-Present with CALs

PSI research within the past two decades has slowed down significantly. Several researchers have discussed the possible reasons for the decline (Fox, 2013; Eyre, 2007; Grant & Spencer, 2003). While PSI research has decreased since the early 1970s, today, researchers continue producing beneficial research. Several areas of PSI that more recent researchers have explored include the following: advancements in CAPSI development (Springer & Pear, 2007); further investigation of the effectiveness of proctor feedback, such as immediacy, accuracy, and quality (Pear, 2003; Martin et al., 2002a; Martin et al., 2002b; Chase & Houmanfar, 2009); improving academic achievements of CALs and the use of PSI and CAPSI to teach various course disciplines such as weight training (Pritchard et al., 2012) and even a behavioral application to higher-order thinking (Crone-Todd & Pear, 2001; Svenningsen & Pear, 2011); implementing PSI for teacher job specialist professional training (Sulzer-Azaroff et al., 2008; Mayer et al., 2014); and finally the integration of PSI into other course methods (Paiva et al., 2017; Svenningsen & Pear, 2011).

Further CAPSI Research

More recent research in PSI has focused on further developments in CAPSI. Advancements in technology and software have opened more opportunities for researchers to take advantage of the CAPSI course designs’ ability to save student performance data for later analysis, a suggestion made by earlier CAPSI researchers (Pear & Kinsner, 1988; Kinsner & Pear, 1988; Crosbie & Kelley, 1993).

Retroactive Analysis in CAPSI. Springer and Pear (2007) retroactively examined the data sets of four online undergraduate psychology CAPSI courses. The authors evaluated the relationship between students' final grades, peer-reviewing participation, and progress rates on unit tests using a less traditional self-pacing procedure. The authors were interested in empirically investigating if a correlation existed between lower unit test progress rates, thus higher procrastination, and lower final exam performance. They were also concerned with determining whether higher peer-reviewing participation correlated with increased final exam performances. The authors conducted data analyses of students' progress rates by dividing them into three comparison groups based on progress scores: an early-massed group (scores above 507), a late-massed group (scores below 393), and a distributed-progress group (scores between 393-507). The authors also rescored final exam grades with clearly defined operational definitions. Interobserver agreement was high across student progress rate comparisons and final exam grade rescoring.

The results in Springer and Pear (2007) indicated students who completed all unit tests performed better on final exams. Still, little difference in exam scores existed between those students with high progress rates and low progress rates on unit tests. Peer-reviewing behaviors had little correlation with final exam performance, however, students with more peer-reviewing points tended to perform better on their final exams (Springer & Pear, 2007). Scores on final exams might have underestimated students' learning with more peer-reviewing experience, but it was difficult to determine if peer-reviewing contributed to improved final exam scores (Springer & Pear, 2007). Also, limited student procrastination might have been due to the deadline contingencies imposed on specific assignments, such as midterms (Springer & Pear, 2007).

This study (Springer & Pear, 2007) supports applying the more traditional self-pacing feature with minimal deadline contingencies in CAPSI courses when administratively feasible. This study's findings also extend the work of earlier researchers in semester-less PSI studies suggesting that students can perform similarly on final exams despite varying levels of course progress rates. This study also suggests there might be two distinct categories of procrastination: students who finish all unit tests and those who do not (Springer & Pear, 2007). The most significant finding of this study was the existing positive correlations between high peer-reviewing participation, progress rate, and final exam scores. This supports the implication for students to begin work earlier and engage in more peer-reviewing, leading to increased overall performances.

Enhancement of Feedback Research

Another area of recent PSI research focuses on examining the effectiveness of different feedback aspects, such as accuracy, quality, and even measures of student compliance with feedback.

General Feedback Research. Research has provided support for the implication of enhancing feedback in CAPSI courses (Pear, 2003). For instance, large classes limit the amount of verbal and written engagement instructors can provide to students on a particular subject matter, resulting in limited quality feedback, which can significantly reduce comprehension, learning, maintenance, and generalization of course material. Therefore, teachers are suggested to give students more short and frequent tests as well as quizzes rather than longer, infrequent exams as previous research demonstrated (Conner-Green, 2000). Contingencies for contributing to online discussions, commonly found in online college courses, make it difficult for instructors to evaluate and provide effective feedback to large amounts of verbal material (Pear, 2003). Minimal feedback is defined as consisting of short statements like "good answer," while substantive feedback as consisting of specific reference to the actual content of the answer (Pear, 2003). Another suggestion was instructors employ methods for standardizing course material to improve the accessibility of monitoring student activities and providing high-quality feedback (Pear, 2003). However, there is evidence to support the amount of substantial input given in CAPSI courses is more significant than in traditional lecture courses.

Feedback Accuracy, Quality, & Compliance. Studies conducted at the University of Manitoba developed procedures for and analyzed proctor marking accuracy (Martin et al., 2002a). They also examined student compliance with proctor feedback (Martin et al., 2002b) in a CAPSI course.

Proctor Marking Accuracy Procedure Development. In their initial study, Martin et al. (2002a) developed software and identified procedures for measuring student proctor-marking feedback accuracy for 33 students in an undergraduate course taught using CAPSI. Students became eligible to be a proctor for a particular unit when they met mastery for that unit. The procedure involved several steps:

1. Upon request, students working through CAPSI units received randomly generated short-essay questions for appropriate units electronically
2. Instructors assigned markers to each completed unit test
3. Instructors electronically delivered completed unit tests to the assigned student proctors for marking
4. And instructors electronically returned the tests to the student with written feedback from the markers

The authors used a sample of 101-unit tests selected for inclusion. Two observers independently assessed the correctness of the 302 answers using the same criteria as the proctors. Interobserver agreement was high for all questions except for 3, which the authors excluded. Therefore there was a total sample of 299 answers. The results of this study, along with student compliance to proctor feedback, were reported in Martin et al. (2002b) and is described below.

Analysis of Proctor Feedback Compliance. Martin et al. (2002b) was the first study to assess the effectiveness of feedback provided in a CAPSI-taught course. Subsequently, the authors analyzed the feedback-related performances of the same 33 students enrolled in the CAPSI course as mentioned above. The authors were interested in investigating students' accuracy in providing feedback on unit tests while acting as proctors and how well the students complied with proctor feedback on unit test retakes. The course contained 10 study units and study questions from the textbook, 10-unit tests, a midterm, and a final. There were two proctors or markers for a student's unit test, and proctors were students that had previously passed the unit. The authors determined the criterion for meeting accuracy for an instance of feedback (IOF) and several responsibilities of the marker involved correctly identifying errors and classifying them with either models, suggestions, examples, questions, or page references. The authors defined Type A errors as an IOF not based on an accurate reading of the answer and Type B errors as an IOF not consistent with the information provided in the textbook or course materials. Additionally, trained observers assessed students' compliance with IOFs. The observers scored feedback compliance if students improved answers consistent with the IOF suggestions.

The results of Martin et al. (2002b) indicated that IOF accuracy was 87% overall. Feedback accuracy was slightly lower when proctors classified test question errors with models and questions compared to the other types. The percentage of students in full compliance with IOFs was 55%, and this percentage increased to 61% when the authors included IOFs met with partial agreement. One limitation of this study is the use of descriptive data analyses, thus difficult to identify the variables functionally related to feedback accuracy and compliance. For example, due to the descriptive nature of the data, there is no way to determine whether or not training materials (i.e., task analysis on how to correctly provide IOFs or instructor models) had an impact on students' feedback accuracy. Visual analysis of single subject data might be a good addition for assessing the variables functionally related to feedback accuracy and compliance (Kazdin, 1982). However, these studies contributed a reliable method for categorizing feedback types (Martin et al., 2002a) and producing measurable feedback accuracy and feedback compliance data (Martin et al., 2002b) in a CAPSI-taught course. Additional research is needed to evaluate effective methods for improving student feedback accuracy and feedback compliance in a CAPSI course.

Analysis of Feedback Quality. In a more recent study, Chase and Houmanfar (2009) used a between-group design to evaluate the differential effects of elaborate and basic feedback on several dependent variables related to students' academic performances in an introductory to psychology course using a modified PSI design. These dependent variables included response accuracy, learning gains from the comparisons of first and second unit test attempts, performance on difficult questions, grade distribution, and social validity measures. The authors operationally defined elaborate feedback as explanatory feedback indicating which answers are correct and incorrect while also providing conceptual information, definitions, and references for clarification. They described basic feedback as corrective and numerical (0-100%), indicating whether answers were correct. Feedback was delivered electronically and immediately following the submission of quizzes using WebCT® software. Students from two of five sections of the course participated, and the authors randomly assigned each section to experimental or control conditions. Course section designs were the same for both conditions except students in the experimental condition received elaborate feedback on unit tests, while students in the control condition received basic feedback.

Chase and Houmanfar (2009) indicated the response accuracy results were undifferentiated between groups on the first unit test attempts. Still, students in the elaborate feedback group had more significant improvements on their second attempts. The elaborate feedback group had an overall 3% higher average on learning gains than the basic feedback group, including on difficult questions. The grade distribution was more concentrated with As and Bs for the elaborate feedback group. Also, social validity scores were high for both PSI sections, especially the experimental group, due to student reports of improved academic performance.

This study (Chase & Houmanfar, 2009) demonstrated a functional relationship between feedback type, question difficulty, and response accuracy, with significantly more improvements occurring for students who received elaborate feedback relative to students who received basic feedback. A significant strength of Chase and Houmanfar (2009) was they substituted descriptive statistical analyses with visual analyses of the students' data, which allowed for further identification of the variables responsible for the differences between groups (Sidman, 1960). As mentioned earlier in the paper, the founder of PSI, Fred S. Keller, said, "the student is always right" (Keller, 1968), meaning if the student isn't learning, evaluation of teaching methods or rearrangement of instructional contingencies may be necessary. The present study demonstrated that one way to improve student learning in PSI courses is by providing effective, elaborate feedback.

PSI for Improving Academic Achievements

PSI for Weight Training. Several recent studies have examined the effectiveness of PSI for improving academic achievements in diverse disciplines. For instance, Pritchard et al. (2012) evaluated the effects of a PSI model in a university physical activity weight training course. Based on the statistics reported by the Centers for Disease Control and Prevention (CDC), leisure-time activity decreased to 36.6% (CDC, 2006), while obesity rates in young adults aged 18 to 29 tripled from 8% in 1971-1974 to 24% in 2005-2006 (CDC, 2009). The authors' proposed solution was to offer empirically based instructional methods for teaching physical activity programs in colleges and universities. The authors developed a physical education adapted PSI model based on Siedentop (1973), which included: content selection, managerial control, task presentation, engagement patterns, instructional interaction, pacing, and task progression. While previous research evaluated PSI in the physical education setting (Siedentop, 1973), this was the first study to examine PSI's effectiveness in a fitness-oriented physical education course. The authors assessed the pre-test scores of students on the FITNESSGRAM and knowledge tests to establish baseline performance and post-test scores on the measures following the PSI course completion. The FITNESSGRAM included a series of tests implemented to measure aerobic capacity, muscular strength and endurance, flexibility, and body composition. The knowledge test consisted of 50 questions designed to assess overall weight training knowledge. It included items related to various unit topics such as types of exercises used, lifting techniques, and the muscles' anatomy.

The results of Pritchard et al. (2012) indicated no statistical difference between pre- and post-scores on some FITNESSGRAM tests, including the progressive aerobic cardiovascular endurance run (PACER), the back saver-left leg lift, the back saver-right leg lift, and the trunk lift tests. However, there were significant statistical differences between the pre- and post-test scores in the number of curl-ups and push-ups performed and a decrease in the percentage of body fat and body mass index following the PSI course. The authors also reported there was a significant increase in knowledge test scores from pre- to post-tests. The adapted PSI model allowed the instructors to provide sufficient direct one-on-one contact with the students with minimal experience, thus enabling them to provide adequate individual feedback (Pritchard et al., 2012). This study demonstrates a PSI model can effectively be used to teach fitness and weight training courses in colleges and universities.

PSI for Critical Thinking. Researchers have even used PSI's empirical application to teaching to courses in critical or higher-order thinking. Crone-Todd and Pear (2001) developed a modified form of Benjamin Bloom's Taxonomy (Bloom, 1956) to behaviorally define and teach higher-order thinking skills to university students using CAPSI. Bloom's Taxonomy consists of six successive levels of critical thinking components, which the authors broke down in two sections: the "lower" levels and the "higher" levels. The "lower" levels of critical thinking are composed of knowledge or memorization and comprehension. A behavioral description of knowledge involves responding to questions with answers in the form of intraverbal chains that are word-for-word from the textual material (Crone-Todd & Pear, 2001). Comprehension refers to the responses a learner emits via intraverbals from the text but clarifies them appropriately to a particular situation of their own condition (Crone-Todd & Pear, 2001). The "higher" levels of critical thinking are composed of application, analysis, synthesis, and evaluation. Application refers to when the learner applies lower-level knowledge to a new situation or problem, essentially generalizing learned material to new exemplars (Crone-Todd & Pear, 2001). The analysis level involves a learner responding to parts of a definition or concept; in other words, smaller response units come under the control of smaller stimulus units (Crone-Todd & Pear, 2001). Synthesis consists of several variables coming together to produce new behavior. For example, the learner can break down a situation into component parts (i.e., task analyses) before providing a synthesized answer to a question. The final and highest level of critical thinking, evaluation combines all the previous categories for providing logical or rational arguments to defend a given position, which involves students employing complex sets of interconnected concepts to generate verbal behavior about new novel concepts (Crone-Todd & Pear, 2001).

Crone-Todd and Pear's (2001) behavioral application was later used in a study by Svenningsen and Pear (2011) to examine the pedagogical effects of CAPSI as a component of a blended learning method for developing and teaching critical thinking skills with CALs. There are many definitions of critical thinking. However, Skinner's behavioral definition of 'critical thinking' for including in this study (Svenningsen & Pear, 2011); critical thinking is verbal behavior that occurs in a speaker who is also their own listener (Skinner, 1957, pp. 438-452). The study was composed of two experiments conducted two successive years' fall introductory to university courses designed to ease students' transition from high school to college. For the first experiment, two experimental course sections received two APA-formatted essays (1,000 words each) plus one CAPSI-based assignment. The other two non-CAPSI course sections received three essay assignments. The CAPSI assignment required students to pass 20-unit tests and to peer review 10-unit tests, in which each unit corresponded to a single chapter of the course textbook. The unit tests consisted of three randomly generated short-answer questions composed according to the six levels of Bloom's Taxonomy and designed to assess students' critical thinking skills and learning of the textbook material. The authors used the Applied Critical Thinking Measure (ATCM), a measure consistent with a behavioral view of critical thinking, to assess students' overall critical thinking development at the end of the course (Svenningsen & Pear, 2011). Other dependent variables included students' final exam scores, performance on the critical thinking exam question, and performance on course content questions. The first experiment's results indicated CAPSI sections scored higher than the non-CAPSI sections on all measures, including the final exam, the critical thinking exam question, course content questions, and the ACTM.

Experiment two was conducted in the fall semester following experiment one (Svenningsen & Pear, 2011). The authors attempted to control for the variability between course conditions in experiment one, since a different instructor taught each section. In this experiment, the course only contained two sections, a CAPSI section and a non-CAPSI section, which were both taught by the same instructor. One course section received a CAPSI assignment similar to the one used in the first experiment, while the other section received a research paper assignment. The only difference in the CAPSI design of this experiment was that the authors developed unit test short-answer questions according to only the sixth and highest level of Bloom's Taxonomy, thus requiring students to evaluate a response supported by information from the course textbook. The authors evaluated the differential effects of CAPSI and traditional lecture sections on the pre- and post-ACTM scores. The results showed that pre-ACTM scores were slightly higher in the CAPSI section than the non-CAPSI sections; however, they were approximately equal prior to the study. The ACTM scores increased significantly in the CAPSI sections, and there was no difference between pre- and post-ACTM scores in the lecture sections.

Svenningsen and Pear (2011) attempted to control for variability in experiment two. However, it may have been better to evaluate CAPSI effects on the same measures studied in experiment one (i.e., final exam scores, critical thinking question responses, course content questions). Assessing the impact of CAPSI on the same dependent variables would have provided a better comparison of the results of the two experiments. Regardless, this study supports earlier researchers' claims (Grant & Spencer, 2003) that PSI can be a sufficient method for teaching critical and higher-order thinking skills in post-secondary courses. Additionally, this study employed an effective CAPSI application to Bloom's Taxonomy (Crone-Todd & Pear, 2001) to develop and teach critical thinking skills to CALs. The authors also included CAPSI as a component in a blended learning course, which adds to the PSI literature by providing additional empirical support for students performing better in a PSI-based hybrid teaching method compared to those in a traditional lecture course.

PSI for Mathematical Concepts. Additionally, Paiva et al. (2017) developed a computer tutorial system based on the tenets of PSI for teaching and maintaining mathematical concepts and evaluated its effects on students' progress and academic performance in an undergraduate quantitative methods course. Undergraduate students with varied socioeconomic backgrounds have higher failure and dropout rates (Paiva et al., 2017). The proposed solution was to design a PSI self-study program with formative assessment and automated feedback (Paiva et al., 2017). Therefore, the authors developed and designed the PSI-based tutorial system for implementation as a mathematical course component. They aimed to create a PSI-based math tutorial system to integrate into higher education programs for students majoring in engineering, economics, political science, management, or other similar degrees. The tutorial system incorporated several different software programs. The program's PSI mastery format involves the student writing answers in algebraic form and receiving immediate feedback with possible suggestions for assistance consisting of video or noted recommendations. When students failed to pass an e-assessment, they earned access to another e-assessment containing a question bank with new exemplars. This would continue until they reached mastery criteria (60% or above). In this study, 72 students were randomly assigned to either the experimental or control groups. The experimental group received the mastery criterion requirement for advancing to the next level in the tutorial system. The control group received the same recommendations; however, there were no restrictions for advancement to successive levels.

The results of Paiva et al. (2017) indicated that the mean number of accesses, or e-assessment attempts, was statistically higher in the experimental group than in the control group. Additionally, students in the experimental group had higher average grades than the control group. Because the only difference between groups was the restriction on access to the following level in the tutorial system, this was concluded as the contributing variable responsible for improved performance in the experimental group. This study demonstrated that the PSI mastery component's implementation could be successfully integrated into an advanced level college mathematics course.

PSI for Implementing Job Training

Another area more recent research has examined is in using PSI-based methods for implementing job specialist training programs with adults with college degrees. For example, Sulzer-Azaroff et al. (2008) developed objectives for a series of CAPSI-taught courses in applied behavior analysis designed for behavior intervention specialists. The authors chose the course objectives based on a survey of experts in ABA intended to identify the most critical topics behavioral specialists are required to learn. Some of the authors later described the results of the course designed in Sulzer-Azaroff et al. (2008) in Mayer et al. (2014). The authors evaluated the effects of a series of CAPSI-taught courses in applied behavior analysis on behavior intervention specialists' pre- and post-test scores. The course consisted of required participation in discussions and field and laboratory activity assignments. Also, study questions were given to students to help them prepare for weekly quizzes. Students could take the unit tests at their own pace and when they felt adequately prepared. Results on unit quizzes and relevant feedback were provided to students immediately. Students were required to meet a mastery criterion of at least 80 percent before continuing to the following unit. Students could retake quizzes as often as necessary; however, question exemplars varied with each subsequent retake.

The results of Sulzer-Azaroff et al. (2008) indicated that the PSI course effectively improved average scores on pre-test from below 60 percent to above 80 percent class average on post-test scores. This study demonstrates the effectiveness of CAPSI for job specialist training in courses in applied behavior analysis. Sulzer-Azaroff et al. (2008) provided an effectively designed course based on critical components identified from expert behavior analysts' surveys.

Integration of PSI with Other Teaching Methods

PSI research within the past decade shows a trend in interest in evaluating the effectiveness of using PSI complimented with other empirical methods. For instance, Svenningsen and Pear (2011), as discussed earlier, used a CAPSI component in a blended learning course, which combines classroom instruction with Web-based learning. Researchers have also found that blended learning to be an empirically supported instructional method, and studies have demonstrated its' superiority for improving student academic achievement compared to the traditional lecture course (e.g., Dzubian et al., 2007; Pereira et al., 2007). Also, Paiva et al. (2017) demonstrated the PSI mastery component was a successful integration into a computerized tutorial system for teaching upper-level college mathematical concepts in a quantitative methods course for undergraduates in a marketing program.

Advantages, Disadvantages, & Considerations with Comparisons to Early PSI Research

Although PSI research has significantly slowed down since its conception and the early 1970s when studies were booming, recent studies within the past two decades have demonstrated PSI to be an empirically sound instructional method for teaching CALs across diverse academic disciplines (Pritchard et al., 2012; Crone-Todd & Pear, 2001; Svenningsen & Pear, 2011; Paiva et al., 2017). Advancements in technology have improved the CAPSI design (Pear, 2003; Springer & Pear, 2007), which is efficient for data collection of various student behaviors, including proctor feedback accuracy and student feedback compliance (Martin et al., 2002a; Martin et al., 2002b).

A majority of recent PSI research followed in earlier researchers' footsteps and focused on establishing procedures for measuring and examining the effects of different feedback aspects. For example, researchers demonstrated students in PSI courses have more proficient learning gains from elaborate feedback containing corrective input, references, examples, and additional explanations than from simple corrective feedback (Chase & Houmanfar, 2009). Researchers have also continued exploring PSI's impact on vocational training procedures and have found it an effective instructional method for educating behavioral intervention specialists (Sulzer-Azaroff et al., 2008; Martin et al., 2014).

Upon review, these studies appeared in various peer-reviewed journals, including several behavior-analytic journals such as the *Journal of Applied Behavior Analysis*, *The Behavior Analyst*, *Journal of Behavioral Education*, and in several educational journals such as *Computers & Education*, *The Physical Educator*, and the *Journal of Computer Assisted Learning*. The appearance of studies in various journals ensures that at least some minimal dissemination of the empirical method has occurred since earlier PSI research.

However, a review of this literature from the last two decades yielded less than a dozen studies. Therefore, researchers should continue to work towards disseminating PSI and CAPSI further. It seems recent PSI research has focused less on evaluating the differential effects of various deadline contingencies on students' academic progress rates and performance. Therefore, researchers should continue to explore this area, as well as PSI and CAPSI's effectiveness as a component of other empirical instructional methods for teaching CALs.

Discussion

In this paper, I reviewed the available research on behavior-analytic teaching methods empirically supported for improving various academic behaviors for CALs. This literature review indicated several behavioral teaching methods have demonstrated their effectiveness for improving student performances with CALs. These include direct instruction (DI), precision teaching (PT), response cards, interteaching, equivalence-based instruction (EBI), and personalized system of instruction (PSI).

Advantages & Disadvantages of Each Behavioral Teaching Method

DI

The review of DI research with CALs yielded limited results, as most DI studies were with early learners (Kinder & Carnine, 1991). While the available DI studies with CALs found this method beneficial for improving student academic measures with high social validity (Cummings et al., 2010; Cazazza, 1993) there was only one study without crucial design flaws or not lacking in demonstration of experimental control (Kits & Thorpe, 1995).

PT

On the other hand, the review of PT research with CALs yielded significantly more research than DI. A PT-based method called SAFMEDS has demonstrated a valuable instructional component for improving student mastery and fluency in course material across various disciplines (Quigley et al., 2017). Several universities and college programs have used or actively use PT and PT-based methods (e.g., SAFMEDS) to increase mastery and fluency learning in their courses (Johnson & Layng, 1992; McDade & Brown, 2001).

Response Cards

Research in response cards with CALs was skim. Most studies examining the effectiveness of response cards for improving students' academic performances on daily and weekly quizzes yields either undifferentiated results (Shabani & Carr, 2004) or were only slightly higher compared to baseline or control groups (Kellum et al., 2001; Marmolejo et al., 2004; Malanga & Sweeney, 2008). However, recent technological advancements and the development of functionally similar electronic response cards via online applications and programs (e.g., Poll Everywhere, Kahoot! ©, etc.) create opportunities for researchers to explore the efficacy of this behavioral teaching method with CALs.

Interteaching

Several newer behavioral teaching methods developed within the last couple of decades, interteaching and EBI, have demonstrated their effectiveness with improving academic performance with CALs. Interteaching is a package of multiple components that relies heavily on student discussions based on instructor-built study guide questions and course objectives (Querol et al., 2015). While a recent comprehensive review of interteaching (Querol et al., 2015) found at least two dozen studies demonstrating this instructional method effectively improves various academic behaviors with high social validity, there are very few studies that conducted component analyses. More research is needed to investigate which components of this instructional package are critical for student achievement.

EBI

Similarly, researchers have found EBI to be an effective instructional method for teaching equivalence relations to CALs across various stimulus classes (i.e., neuroanatomy, piano skills, mathematics). However, a significant portion of the EBI research conducted was in highly controlled laboratories (Brodsky & Fienup, 2018). Although some studies exist (Greville et al. 2016; O'Neill et al., 2015; Pytte & Fienup, 2012; Varelas & Fields, 2017; Ramos et al., 2018), much less of the EBI research conducted was in applied or more naturalistic classroom settings (Brodsky & Fienup, 2018). Therefore, more research is needed to adapt and develop curricula that effectively integrate EBI methods and examine the effectiveness of this EBI-based curriculum for improving the academic behaviors of CALs.

PSI

The final behavior-analytic instructional method reviewed, PSI, was developed by a renowned researcher in behavior analysis, Fred S. Keller (1968), and yielded the most significant results with CALs. PSI and its computerized adaptation, CAPSI, have an extensive research history demonstrating their effectiveness for improving various student academic behaviors and adult vocational skills in diverse disciplines across numerous settings for over half a century.

Early PSI. The early 1970s to the late 1990s was a prime era for a considerable amount of PSI research. Investigators found PSI to be as effective, but generally more effective than traditional lecture for improving academic achievements (Kulik et al., 1979; Gray et al., 1986; Jumpeter, 1985; Gifford & Vicks, 1982). Early researchers also conducted component analyses that helped to identify the PSI components critical for student success (Kulik et al., 1978). They also found that limiting PSI's self-paced feature by using deadline continencies reduced student procrastination and improved course progress and academic performances (Ross & McBean, 1995; Reisner, 1984). Early PSI research indicated feedback was most effective if delivered following the completion of unit tests (Haemmerlie, 1985). Researchers developed and found CAPSI a successful adaptation of PSI that is an efficient data-yielding instructional method (Pear & Kinsner, 1988; Kinsner & Pear, 1988; Pear & Novak, 1996; Crosbie & Kelley, 1993). Lastly, researchers found PSI a practical component in teacher training procedures for improving treatment integrity in a CABAS®-based school for children with developmental disabilities (Selinske et al., 1991).

Recent PSI. While recent research in PSI has significantly declined since the late 1990s, researchers have examined PSI and CAPSI's effectiveness across numerous student academic behaviors in the past couple of decades. More recent PSI research has produced advancements in CAPSI (Springer & Pear, 2007) and developed effective procedures for measuring student proctor feedback accuracy and feedback compliance (Martin et al., 2002a, 2002b). Current PSI researchers have demonstrated the efficacy of delivering elaborate feedback rather than basic corrective feedback (Chase & Houmanfar, 2009). More recent research has shown improved student learning in widely different academic disciplines (Pritchard et al., 2012; Crone-Todd & Pear, 2001; Svenningsen & Pear, 2011; Paiva et al.

2017). Also, studies have found PSI to be an effective procedure for job specialist skill training (Sulzer-Azaroff et al., 2008; Martin et al., 2014). Other studies in PSI have even shown some promise when combined with other empirical teaching methods (Svenningsen & Pear, 2011). Less recent PSI research focused on component analyses or investigating the effectiveness of deadline contingencies for improving course progress or student academic performances.

Evaluation & Comparison of Behavioral Teaching Methods

There are many evidence-based practices derived from principles of behavior analysis available for university and college instructors to choose from when selecting a teaching method to use with their students. However, based on the present review of the literature with CALs, I suggest further research examine the effectiveness of several teaching methods with this population, including DI, response cards, interteaching, and EBI. Also, PT-based methods, such as SAFMEDS, have demonstrated this method as an effective additional component to other teaching methods (Kubina et al., 2009); however, little research showed this to be sufficient as a stand-alone procedure for improving learning in traditional college settings (Quigley et al., 2017). Additionally, little research has examined the combined effects of the evidence-based practices above. Future research in combining methods could be especially valuable for developing effective college curricula that include practices typically used as components in other instructional methods, such as response cards, EBI, and PT.

Finally, based on this review's findings, of all the behavioral teaching methods discussed above, PSI and the sibling method CAPSI had the most robust literature supporting their effectiveness for improving various academic behaviors with CALs. These evidence-based instructional methods show the most promise in terms of adaptability and flexibility for instructors to integrate into disparate course disciplines in a traditional classroom, blended learning, or distance learning setting. In the following sections, I will discuss the common misconceptions about PSI, the advantages and barriers to using PSI-based methods, and suggestions for future research.

PSI with CALs: Misconceptions & Reasons for Research Decline

Despite the significant body of research supporting PSI's efficacy to teach CALs, there has been a considerable decline in PSI research conducted within the past couple of decades. Several investigators have examined the possible reasons for this decline.

PSI has often been overlooked due to educational establishments resistance to change and the significant time demanded setting up and maintaining PSI courses (Grant & Spencer, 2003). A contributing factor is the implementations of ineffective teaching methods called "PSI" that do not contain the critical PSI components (Grant & Spencer, 2003). Another reason is there are several misconceptions about the nature of PSI in academic literature, including that PSI is not appropriate for teaching higher-order thinking (Grant & Spencer, 2003). However, several research studies supported a behavioral approach to using PSI to teach critical thinking skills with CALs (Crone-Todd & Pear, 2001; Svenningsen & Pear, 2011).

The debate of the decline of PSI continued and Eyre (2007) agreed with several previous arguments (Grant & Spencer, 2003). A possible reason for the fall of PSI is because PSI courses diverge from “true” PSI to “something-like-it” in that they are missing key components such as the unit mastery or self-pacing features, which makes it difficult to assess the effectiveness (Eyre, 2007). Educational establishment resistance has been seen through university and college administrations blocking PSI based on the belief teachers are not teaching unless they are standing in front of a class (Eyre, 2007). Additionally, educational researchers might have lost interest in PSI due to complaints of the course's time-sensitive nature (Eyre, 2007).

Fox (2013) reiterated the likely reasons for PSI's decline mentioned by earlier researchers. Moreover, instructors often have difficulty adapting the self-paced model to traditional academic calendars (Fox, 2013). However, although the exact reasons for why PSI fell from favor are unknown, it is not due to a lack of student achievement with this method (Fox, 2013).

Advantages of PSI with CALs

Limiting Self-Pacing with Deadline Contingencies. There are many advantages to a PSI approach for college students and adult learners. The self-paced feature of PSI allows for flexibility in course structure (Fox, 2004). This component permits instructors to individualize the course in terms of student's needs, instructors' time, and adaptability to various course disciplines (Querol et al., 2015; Mayer et al., 2014). Instructors can also limit the self-paced feature by using deadline contingencies to reduce student procrastination and improve course progress rates without compromising PSI's efficacy for improving student academic achievements (e.g., Ross & McBean, 1995; Reisner, 1985).

Technological Improvements. Also, in an ever technologically advancing society, PSI's computerized adaptation CAPSI can be implemented across many classroom settings, including distance education (Grant & Spencer, 2003). CAPSI also offers instructors the capability to keep records easily and access data sets of various student academic behaviors (Pear & Kinsner, 1988; Kinsner & Pear, 1988; Pear & Novak, 1996; Springer & Pear, 2007). Easy access to data permits instructors to actively analyze student performance patterns and make individual adjustments to instructional procedures as necessary.

PSI Combined with other Evidence-Based Practices. Additionally, PSI can be combined with other evidence-based practices (e.g., Svenningsen & Pear, 2011) or even implemented as job specialist training procedures (e.g., Sulzer-Azaroff et al., 2008). PSI-based methods can be used successfully across a diverse range of course disciplines including upper-level college mathematics (Paiva et al., 2017), physical fitness and weight training (e.g., Pritchard et al., 2012), and even higher-order thinking courses (e.g., Svenningsen & Pear, 2011) despite doubts from early researchers (e.g., Meek, 1977). Furthermore, PSI can be implemented at a variety of scales, from within one classroom, across entire undergraduate or graduate programs (e.g., Martinez-Diaz & Wilder, 2016), or on a school-wide basis (e.g., Selinske et al., 1991).

PSI Components with Most Empirical Support. Finally, the PSI components with the most empirical support and found to be most critical for students' academic success are the unit mastery and immediate feedback features (Fox, 2013). The unit mastery requirement ensures students have learned previous material before moving on to more advanced course material (Keller, 1968). The most beneficial aspect of this feature is students can take unit tests as often as necessary to master the material. However, instructors can easily set limits on the number of unit test attempts allotted and vary the mastery criterion accordingly based on the difficulty of the material. The feedback component is essential for students learning from one unit test attempt to the next (Fox, 2004, 2013). Elaborate feedback is the most effective for producing the most learning gains (i.e., Chase & Houmanfar, 2009). PSI courses can be structured so that student proctors provide feedback (e.g., Martin et al. 2002a, 2002b); however, computerized program alternatives to proctors have also been successful (e.g., Crosbie & Kelley, 1993).

Barriers

There are several potential barriers to implementing PSI with college students and adult learners that may be worth considering. First, some college course structures might limit the feasibility of the implementation of a PSI model. For example, the self-paced feature might be more challenging to integrate into traditional semesterly courses. Second, resistance from administrative staff in colleges and universities might present a barrier to PSI's application with CALs for several reasons. One is the time required to develop and maintain PSI programs (Buskist et al., 1991; Cracolice & Roth, 1996; Lloyd & Lloyd, 1986). Another reason for administrative resistance to incorporating PSI methods is the educational establishment's reluctance to change (Buskist et al., 1991; Cracolice & Roth, 1996; Lloyd & Lloyd, 1986; Sherman, 1992) or altogether prohibiting PSI courses (Sherman, 1992).

Proposed Solution to Barriers. I propose that a solution to the limitations and barriers posed by administrative resistance to implementing PSI for CALs may be a hybrid PSI/flipped-learning course model. Flipped learning is a pedagogical approach that involves direct instruction moving from the group learning space to the individual learning space. The resulting group space becomes a dynamic, interactive learning environment where the educator guides students to apply concepts and engage creatively in the subject matter (Flipped Learning Network, 2014).

Hybrid PSI/Flipped Learning Course. For example, semesterly courses could be designed based on either traditional classroom or blended learning settings. Instructors could develop PSI-based learning modules for students to work through course objectives between class meetings. Instructors could dedicate class time to having student-led group discussions on the course material students worked through each week. Instructors could guide and facilitate discussions as needed. Course grades could be calculated based on the total percentages of students' participation grades from class discussions and unit completion grades from the PSI-based modules.

Potential Benefits. There are many student benefits of the flipped learning component of the PSI/flipped learning hybrid model described above (Sota, 2016a). Students who require the most help on material covered in PSI modules could get assistance from peers and instructors during in-class activities and review. Flipped learning requires that students learn time-management skills and incorporates student choice in in-class activities (2016a). By combining a flipped learning approach to a PSI course, behavioral repertoires are more likely to be complete. Any concepts not learned in at-home material can improve during in-class meetings (2016a). In addition to PSI coursework, flipped learning further promotes individualized, personalized, and discrimination learning due to in-class opportunities to observe and participate in demonstrations of learned concepts (2016a).

Pilot Research. One initial application of a flipped-PSI hybrid model was implemented by a faculty member in the Applied Behavior Analysis and Clinical Science (ABACS) program at Rollins College, Dr. Stephanie Kincaid (personal communication, June 22, 2018). Dr. Kincaid is a committee member and advisor on the present review. In a pilot study, Kincaid designed a computerized PSI course used in conjunction with in-class meetings structured based on a flipped-learning design. Kincaid measured the quiz attempts to mastery in each successive PSI module and manipulated the question bank sizes. Students could retake module quizzes an infinite number of times, but a requirement of 100% correct responding allowed students to move on to the following module. Kincaid also measured the in-class quiz scores, which included new exemplars of the concepts covered in the modules assigned for each week's class. It is hypothesized that high in-class quiz scores will be a function of higher module quiz attempts when the bank size was increased, containing more exemplars of the concepts of those modules.

Future Research

Future research examining the effectiveness of this hypothetical PSI/flipped-learning hybrid model proposed above might focus on conducting a descriptive analysis of retrospective data sets from Kincaid's (personal communication, June 22, 2018) pilot research. Descriptive analyses of Kincaid's pilot research might help inform future research examining the utility of this proposed PSI-based model with participating undergraduate or first-year graduate students in various potential courses. One empirical question worth studying might be: Will systematically increasing the bank size of module quizzes and exposing students to multiple exemplars in a hybrid PSI/flipped-learning promote generalization of material and improve in-class quiz scores? I hypothesize that by increasing the bank size of PSI-module based quizzes, we will be exposing students to multiple exemplars, thus, increasing the probability of high quiz scores.

General Conclusion

The paper aimed to provide a comprehensive review of evidence-based practices in education shown to demonstrate empirical support for improving various academic behaviors and learner outcomes. Research in behavior analysis in education has declined in the last half-century. Even more, the learner population neglected most in this declining era of behavior-analytic educational research is college students.

Behavior-analytic instructional methods found to be generally effective with this learner population include direct instruction (DI), precision teaching (PT), response cards, interteaching, and equivalence-based instruction (EBI). However, personalized system of instruction (PSI) yielded the most significant results in terms of studies demonstrating this methods' effectiveness for improving academic achievement with college-aged and adult learners. In this review, I discussed possible reasons PSI has been overlooked in education and the advantages and barriers to using this method. Finally, I proposed using a hybrid PSI/flipped learning course method as a potential solution to some of the barriers instructors might face when implementing PSI-based practices. I suggest future research investigate the efficacy using this proposed instructional model. In conclusion, PSI-based instructional methods have a strong research foundation of over half a century that educators can and should build on.

References

Allen, G. J. (1984). Using a personalized system of instruction to improve the writing skills of undergraduates. *Teaching of Psychology*, *11*(2), 95–98. https://doi.org/10.1207/s15328023top1102_10

Aninao, T., Acevedo, D., Newsome, K., & Newsome, D. (2015). Putting the precision in precision teaching: Using the standard celeration chart for ongoing assessment. *Behavioral Development Bulletin*, *20*(2), 168–172. https://doi.org/10.1037/h0101312

Armendariz, F., & Umbreit, J. (1999). Using active responding to reduce disruptive behavior in a general education classroom. *Journal of Positive Behavior Interventions*, *1*(3), 152–158.

Arntzen, E., Halstadtro, L.-B., Bjerke, E., & Halstadtro, M. (2010). Training and testing music skills in a boy with autism using a matching-to-sample format. *Behavioral Interventions*, *25*(2), 129–143. https://doi.org/10.1002/bin.301

Austin, J. L. (2000). Behavior approaches to college teaching. In J. L. Austin & J. E. Carr (Eds.), *Handbook of applied behavior analysis* (pp. 449–472). Context Press.

Axelrod, S. (1991). The problem: American education the solution: Use behavior analytic technology. *Journal of Behavioral Education*, *1*(3), 275–282. https://doi.org/10.1007/BF00947182

Axelrod, S. (1992). Disseminating an effective educational technology. *Journal of Applied Behavior Analysis*, *25*(1), 31–35. https://doi.org/10.1901/jaba.1992.25-31

Ayllon, T., & Michael, J. (1959). The psychiatric nurse as a behavioral engineer. *Journal of the Experimental Analysis of Behavior*, *2*, 323–334.

Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *1*(1), 91–97. PubMed. https://doi.org/10.1901/jaba.1968.1-91

Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy*, *20*(2), 261–282. https://doi.org/10.1016/S0005-7894(89)80073-5

Barlow, D. H., Nock, M., & Hersen, M. (2009). *Single case experimental designs: Strategies for studying behavior for change* (3rd ed). Pearson/Allyn and Bacon.

Barrett, B. H., Beck, R., Binder, C., Cook, D. A., Engelmann, S., Greer, R. D., Kyrklund, S. J., Johnson, K. R., Maloney, M., McCorkle, N., Vargas, J. S., & Watkins, C. L. (1991). The right to effective education. *The Behavior Analyst*, *14*(1), 79–82. https://doi.org/10.1007/BF03392556

Beck, R. J. (1976). *Report for the Office of Education Dissemination Review Panel, 1976, Special Education* (Precision Teaching Project). Great Falls Public Schools.

Beck, R. J. (1981). *High School Basic Skills Improvement: Validation Report for ESEA Title IV-C* (Precision Teaching Project). Great Falls Public Schools.

Becker, W. (1977). Teaching reading and language to the disadvantaged—What we have learned from field research. *Harvard Educational Review*, *47*(4), 518–543. https://doi.org/10.17763/haer.47.4.51431w6022u51015

Becker, W. C., & Carnine, D. W. (1980). Direct instruction. In B. B. Lahey & A. E. Kazdin (Eds.), *Advances in Clinical Child Psychology* (pp. 429–473). Springer US. https://doi.org/10.1007/978-1-4613-9805-9_11

Becker, W. C., & Gersten, R. (1982). A follow-up of follow through: The later effects of the direct instruction model on children in fifth and sixth grades. *American Educational Research Journal*, *19*(1), 75–92. https://doi.org/10.2307/1162369

Bijou, S. W., & Baer, D. M. (1961). *Child Development: A systematic and empirical theory*. Appleton-Century-Crofts. https://books.google.com/books?id=mf8MAQAAIAAJ

Binder, C. (1994). Measurably superior instructional methods: Do we need sales and marketing? In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. W. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 21–31). Brooks-Cole.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives: Affective domain, by D. R. Krathwohl, B. S. Bloom [and] B. B. Masia*. D. McKay. https://books.google.com/books?id=mRWcAAAAMAAJ

Bono, S. F., & McAvoy, R. M. (1977). Student preference and performance under a self-grading and a proctor-grading procedure in a personalized system of instruction approach. *Journal of Personalized Instruction*, *2*(1), 28–34.

Boyce, T. E., & Hineline, P. N. (2002). Interteaching: A strategy for enhancing the user-friendliness of behavioral arrangements in the college classroom. *The Behavior Analyst*, *25*(2), 215–226. https://doi.org/10.1007/BF03392059

Branch, M. N. (1977). On the role of "memory" in the analysis of behavior. *Journal of the Experimental Analysis of Behavior*, *28*(2), 171–179. https://doi.org/10.1901/jeab.1977.28-171

Brodsky, J., & Fienup, D. M. (2018). Sidman goes to college: A meta-analysis of equivalence-based instruction in higher education. *Perspectives on Behavior Science*, *41*(1), 95–119. https://doi.org/10.1007/s40614-018-0150-0

Bushell, D., & Baer, D. (1994). Measurably superior instruction means close, continual contact with the relevant outcome data. Revolutionary. In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 3–10). Brooks/Cole.

Buskist, W., Cush, D., & DeGrandpre, R. J. (1991). The life and times of PSI. *Journal of Behavioral Education*, *1*(2), 215–234. https://doi.org/10.1007/BF00957005

Calkin, A. B. (2003). The course of precision teaching. *European Journal of Behavior Analysis*, *4*(1–2), 87–95. https://doi.org/10.1080/15021149.2003.11434222

Cameron, M. J., & Cappello, M. J. (1993). “We’ll cross that hurdle when we get to it”: Teaching athletic performance within adaptive physical education. *Behavior Modification*, *17*(2), 136–147. https://doi.org/10.1177/01454455930172003

Carmignani, K. (2010). Training tigers: The cat’s meow. *San Diego Zoonooz*, *February*, 29–31.

Carnine, D. (1997). Bridging the research-to-practice gap. *Exceptional Children*, *63*(4), 513–521. https://doi.org/10.1177/001440299706300406

Carnine, D. W. (2000). *Why education experts resist effective practices; Report of the Thomas B. Fordham Foundation*. Thomas B. Fordham Foundation.

Carr, E. G., Newsom, C. D., & Binkoff, J. A. (1980). Escape as a factor in the aggressive behavior of two retarded children. *Journal of Applied Behavior Analysis*, *13*(1), 101–117. https://doi.org/10.1901/jaba.1980.13-101

Carr, J. E. (2016). *On the diversity of practice areas in applied behavior analysis*. Florida Association of Behavior Analysis Conference, Marriott Resort Ft. Lauderdale Harbor Beach.

Casazza, M. E. (1993). Using a model of direct instruction to teach summary writing in a college reading class. *Journal of Reading*, *37*(3), 202–208.

Cavanaugh, R. A., Heward, W. L., & Donelson, F. (1996). Effects of response cards during lesson closure on the academic performance of secondary students in an earth science course. *Journal of Applied Behavior Analysis*, *29*(3), 403–406.

Center for Disease Control and Prevention. (2009, February 18). *Health habits of adults aged 18-29 highlighted in report on nation's health* [Press Release]. https://www.cdc.gov/media/pressrel/2009/r090218.htm

Center for Disease Control Prevention. (2006). *Health, United States, 2006 with chartbook on trends in the health of Americans*. U.S. Department of Health and Human Services. https://www.cdc.gov/nchs/data/hus/hus06.pdf#highlights

Center of Innovations in Learning & Twyman, J. S. (2014). Envisioning Education 3.0: The fusion of behavioral analysis, learning science and technology. *Revista Mexicana de Análisis de La Conducta*, *40*(2), 20–38. https://doi.org/10.5514/rmac.v40.i2.63663

Chase, J. A. (2006). *Differential effects of elaborate feedback and basic feedback on student performance in a modified PSI course.* [Unpublished Master's Thesis]. University of Nevada.

Chase, Jared A., & Houmanfar, R. (2009). The differential effects of elaborate feedback and basic feedback on student performance in a modified, personalized system of instruction course. *Journal of Behavioral Education*, *18*(3), 245–265. https://doi.org/10.1007/s10864-009-9089-2

Christle, C. A., & Schuster, J. W. (2003). The effects of using response cards on student participation, academic achievement, and on-task behavior during whole-class, math instruction. *Journal of Behavioral Education*, *12*(3), 147–165.

Connor-Greene, P. A. (2000). Assessing and promoting student learning: Blurring the line between teaching and testing. *Teaching of Psychology*, *27*(2), 84–88. https://doi.org/10.1207/S15328023TOP2702_01

Cooper, Heron, Heward, 2007. (n.d.). *Applied behavior analysis*.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed). Pearson/Merrill-Prentice Hall.

Cracolice, M. S., & Roth, S. M. (1996). Keller's "old" personalized system of instruction: A "new" solution for today's college chemistry students. *The Chemical Educator*, *1*(1), 1–18. https://doi.org/10.1007/s00897960004a

Croft, R. G., Johnson, W. G., Berger, J., & Zlotlow, S. F. (1976). The influence of monitoring on PSI performance. *Journal of Personalized Instruction*, *1*(1), 28–31.

Crone-Todd, D. E., & Pear, J. J. (2001). Application of Bloom's Taxonomy to PSI. *The Behavior Analyst Today*, *2*(3), 204.

Crosbie, J., & Kelly, G. (1993). A computer-based personalized system of instruction course in applied behavior analysis. *Behavior Research Methods, Instruments, & Computers*, *25*(3), 366–370. https://doi.org/10.3758/BF03204527

Cummings, R., Maddux, C. D., Cladianos, A., & Richmond, A. (2010). Moral reasoning of education students: The effects of direct instruction in moral development theory and participation in moral dilemma discussion. *Teachers College Record*, *112*(3), 621–644.

Dangel, H. L., & Wang, C. X. (2008). Student response systems in higher education: Moving beyond linear teaching and surface learning. *Journal of Educational Technology Development and Exchange*, *1*(1), 93–104.

Davis, L. L., & O'Neill, R. E. (2004). Use of response cards with a group of students with learning disabilities including those for whom English is a second language. *Journal of Applied Behavior Analysis*, *37*(2), 219–222.

Dziuban, C. D., Hartman, J., & Moskal, P. (2007). Everything I need to know about blended learning I learned from books. In A. G. Picciano & C. D. Dziuban (Eds.), *Blended Learning: Research Perspectives* (pp. 265–286). Sloan Center for Online Education.

Edwards, J. S. (1969). *Precisely teaching children labelled learning disabled* [Unpublished Master's Thesis]. University of Kansas.

Embry, D. D., & Biglan, A. (2008). Evidence-based kernels: Fundamental units of behavioral influence. *Clinical Child and Family Psychology Review, 11*(3), 75–113. https://doi.org/10.1007/s10567-008-0036-x

Engelmann, S., & Carnine, D. (1991). *Theory of instruction: Principles and applications* (Rev. ed). ADI Press.

Eyre, H. L. (2007). Keller's personalized system of instruction: Was it a fleeting fancy or is there a revival on the horizon? *The Behavior Analyst Today*, *8*(3), 317–324. https://doi.org/10.1037/h0100623

Felderman, T. A. (2014). Preliminary analysis of interteachings Frequent examinations component in the community college classroom. *Journal of College Teaching & Learning (TLC)*, *11*(4), 149–156. https://doi.org/10.19030/tlc.v11i4.8851

Fienup, D. M., Covey, D. P., & Critchfield, T. S. (2010). Teaching brain—Behavior relations economically with stimulus equivalence technology. *Journal of Applied Behavior Analysis*, *43*(1), 19–33. https://doi.org/10.1901/jaba.2010.43-19

Fienup, D. M., Hamelin, J., Reyes-Giordano, K., & Falcomata, T. S. (2011). College-level instruction: Derived relations and programmed instruction. *Journal of Applied Behavior Analysis*, *44*(2), 413–416. https://doi.org/10.1901/jaba.2011.44-413

Fienup, D. M., Wright, N. A., & Fields, L. (2015). Optimizing equivalence-based instruction: Effects of training protocols on equivalence class formation: Optimizing equivalence-based instruction. *Journal of Applied Behavior Analysis*, *48*(3), 613–631. https://doi.org/10.1002/jaba.234

Fink, E. R. (1968). *Performance and selection rates of emotionally disturbed and mentally Retarded Preschoolers on Montessori Materials* [Unpublished Master's Thesis]. University of Kansas.

Flipped Learning Network. (2014). *The four pillars of F-L-I-P*. Flipped Learning Network. www.flippedlearning.org/definition

Fox, E. J. (2013). Keller's personalized system of instruction. In J. Hattie & E. M. Anderman (Eds.), *International guide to student achievement* (pp. 359–361). Routledge. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=1114642

Foxx, R. M. (1996). Translating the covenant: The behavior analyst as ambassador and translator. *The Behavior Analyst*, *19*(2), 147–161. https://doi.org/10.1007/BF03393162

Ganschow, L., & Sparks, R. (2001). Learning difficulties and foreign language learning: A review of research and instruction. *Language Teaching*, *34*(2), 79–98. https://doi.org/10.1017/S0261444800015895

Gardner III, R., Heward, W. L., & Grossi, T. A. (1994). Effects of response cards on student participation and academic achievement: A systematic replication with inner-city students during whole-class science instruction. *Journal of Applied Behavior Analysis*, *27*(1), 63–71.

Gersten, R. (1985). Structured immersion for language minority students: Results of a longitudinal evaluation. *Educational Evaluation and Policy Analysis*, *7*(3), 187–196. https://doi.org/10.3102/01623737007003187

Gifford, V. D., & Vicks, J. (1982). A comparison of the personalized system of instruction and a conventional biology course on the achievement of junior college freshmen. *Journal of Research in Science Teaching*, *19*(8), 659–664. https://doi.org/10.1002/tea.3660190805

Godfrey, S. A., Grisham-Brown, J., Schuster, J. W., & Hemmeter, M. L. (2003). The effects of three techniques on student participation with preschool children with attending problems. *Education and Treatment of Children*, 255–272.

Graf, S., & Auman, J. (2005). SAFMEDS: A tool to build fluency. *Graf Implements*.

Grant, L. K., & Spencer, R. E. (2003). The personalized system of instruction: Review and applications to distance education. *The International Review of Research in Open and Distributed Learning*, *4*(2). https://doi.org/10.19173/irrodl.v4i2.152

Gray, P. L., Buerkel-Rothfuss, N. L., & Yerby, J. (1986). A comparison between PSI-based and lecture-recitation formats of instruction in the introductory speech communication course. *Communication Education*, *35*(2), 111–125. https://doi.org/10.1080/03634528609388329

Green, B. F. JR. (1974). Fifteen reasons not to use the Keller plan. In J. G. Sherman (Ed.), *PSI: Personalized System of Instruction: 41 germinal papers* (pp. 117–119). W. A. Benjamin.

Green, G., & Saunders, R. R. (1998). Stimulus equivalence. In K. A. Lattal & M. Perone (Eds.), *Handbook of Research Methods in Human Operant Behavior* (pp. 229–262). Springer US. https://doi.org/10.1007/978-1-4899-1947-2_8

Greer, R. D. (1997). The comprehensive application of behavior analysis to schooling (CABAS®). *Behavior and Social Issues*, *7*(1), 59–63. https://doi.org/10.5210/bsi.v7i1.300

Greville, W. J., Dymond, S., & Newton, P. M. (2016). The student experience of applied equivalence-based instruction for neuroanatomy teaching. *Journal of Educational Evaluation for Health Professions*, *13*, 32. https://doi.org/10.3352/jeehp.2016.13.32

Griffith, K. R., Ramos, A. L., Hill, K. E., & Miguel, C. F. (2018). Using equivalence-based instruction to teach piano skills to college students: Equivalence and music. *Journal of Applied Behavior Analysis*, *51*(2), 207–219. https://doi.org/10.1002/jaba.438

Haemmerlie, F. M. (1985). Role of immediate feedback in a personalized system of instruction: Evidence of a negative impact. *Psychological Reports*, *56*(3), 947–954. https://doi.org/10.2466/pr0.1985.56.3.947

Harris, F. R., & Wolf, M. M. (1967). Effects of adult reinforcement on child behavior. In R. J. Karracker, N. B. King, & D. M. Baer (Eds.), *Readings in Classroom Management* (pp. 88–97). MSS Information Corporation. https://books.google.com/books?id=hOeNm2Yc8NwC

Hattie, J., & Anderman, E. M. (2013). *International guide to student achievement*. Routledge. http://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=1114642

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112. https://doi.org/10.3102/003465430298487

Haughton, E. (1971). Great gains from small starts. *TEACHING Exceptional Children*, *3*(3), 141–146. https://doi.org/10.1177/004005997100300310

Hayes, L. J., Thompson, S., & Hayes, S. C. (1989). Stimulus equivalence and rule following. *Journal of the Experimental Analysis of Behavior*, *52*(3), 275–291. https://doi.org/10.1901/jeab.1989.52-275

Heron, T. E., Heward, W. L., Neef, N. A., Peterson, S. M., Sainato, D. M., Cartledge, G., Gardner III, R., Peterson, L. D., Hersh, S. B., & Dardig, J. C. (Eds.). (1994). *Focus on behavior analysis in education: Achievements, challenges, and opportunities*. Pearson/Merrill/Prentice Hall.

Herrnstein, R. J., & Loveland, D. H. (1964). Complex visual concept in the pigeon. *Science*, *146*(Whole No. 3643), 549–550. https://doi.org/10.1126/science.146.3643.549

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes*, *2*(4), 285–302. https://doi.org/10.1037/0097-7403.2.4.285

Heward, W. L. (1994). Three “low-tech” strategies for increasing the frequency of active student response during group instruction. In R. Gardner III, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior* *analysis in education: Focus on measurably superior instruction.* (pp. 283–320). Thomson Brooks/Cole Publishing Co.

Hodges, A., Davis, T., Crandall, M., Phipps, L., & Weston, R. (2017). Using shaping to increase foods consumed by children with autism. *Journal of Autism and Developmental Disorders*, *47*(8), 2471–2479. https://doi.org/10.1007/s10803-017-3160-y

Holland, J. G., & Skinner, B. F. (1961). *The analysis of behavior: A program for self-instruction.* (pp. x, 337). McGraw-Hill.

Honig, W. K. (1966). *Operant behavior: Areas of research and application.* Appleton-Century-Crofts.

Horner, R. H., & Albin, R. W. (1988). Research on general-case procedures for learners with severe disabilities. *Education and Treatment of Children*, *11*(4), 375–388.

J. Fox, E. (2004). The personalized system of instruction. In *Evidence-Based Educational Methods* (pp. 201–221). Elsevier. https://doi.org/10.1016/B978-012506041-7/50013-9

The Individuals with Disabilities Education Act (IDEA), Pub. L. No. 101–476, 20 U.S.C. (2004).

Jacob Truelove, Bryan Saville, & Ryan Van Patten. (2013). Interteaching: Discussion group size and course performance. *Journal of the Scholarship of Teaching and Learning*, *13*(2). https://scholarworks.iu.edu/journals/index.php/josotl/article/view/3200

Johnson, K. R., & Layng, T. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, *47*(11), 1475–1490. https://doi.org/10.1037/0003-066X.47.11.1475

Johnson, K. R., & Layng, T. V. (1994). The Morningside model of generative instruction. In R. Gardner, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education: Focus on measurably superior instruction* (pp. 173–197). Brooks-Cole.

Johnson, N. J. (1971). *Acceleration of Inner-city Elementary Schools' Pupils Reading Performance* [Unpublished Doctoral Dissertation]. University of Kansas.

Jumpeter, J. (1985). Personalized system of instruction versus the lecture-demonstration method in a specific area of a college music appreciation course. *Journal of Research in Music Education*, *33*(2), 113. https://doi.org/10.2307/3344731

Kaleta, R., & Joosten, T. (2007). Student response systems. *Research Bulletin*, *10*(1), 1–12.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings* (2nd ed). Oxford University Press.

Keats, D., & Schmidt, J. P. (2007). The genesis and emergence of Education 3.0 in higher education and its potential for Africa. *First Monday*. https://doi.org/10.5210/fm.v12i3.1625

Keller, F. S. (1968). Good-bye, teacher... *Journal of Applied Behavior Analysis*, *1*(1), 79–89. PubMed. https://doi.org/10.1901/jaba.1968.1-79

Keller, F. S., & Sherman, J. G. (1974). *PSI, the Keller Plan Handbook: Essays on a personalized system of instruction*. WA Benjamin Advanced Book Program.

Kellum, K. K., Carr, J. E., & Dozier, C. L. (2001). Response-card instruction and student learning in a college classroom. *Teaching of Psychology*, *28*(2), 101–104.

Kincaid, S. (n.d.). [Unpublished data sets].

Kinder, D., & Carnine, D. (1991). Direct instruction: What it is and what it is becoming. *Journal of Behavioral Education*, *1*(2), 193–213. https://doi.org/10.1007/BF00957004

Kinsner, W., & Pear, J. (1988). Computer-aided instruction for the virtual classroom. *Canadian Journal of Education Communication*, *17*(1), 21–36.

Kitz, W. R., & Thorpe, H. W. (1995). A comparison of the effectiveness of videodisc and traditional algebra instruction for college-age students with learning disabilities.

Remedial and Special Education, *16*(5), 295–306. https://doi.org/10.1177/074193259501600506

Koenig, C. H. (1967). *Precision Teaching with Emotionally Disturbed Children* [Unpublished Master's Thesis]. University of Kansas.

Kratochwill, T. R., Albers, C. A., & Steele Shernoff, E. (2004). School-based interventions. *Child and Adolescent Psychiatric Clinics of North America*, *13*(4), 885–903. https://doi.org/10.1016/j.chc.2004.05.003

Kubina Jr, R. M., Commons, M. L., & Heckard, B. (2009). Using precision teaching with direct instruction in a summer school program. *Journal of Direct Instruction*, *9*(1), 1–12.

Kulik, J. A., Jaksa, P., & Kulik, C. C. (1978). Research on component features of Keller's personalized system of instruction. *Journal of Personalized Instruction*, *3*(1), 2–14.

Kulik, J. A., Kulik, C. C., & Cohen, P. A. (1979). A meta-analysis of outcome studies of Keller's personalized system of instruction. *American Psychologist*, *34*(4), 307–318. https://doi.org/10.1037/0003-066X.34.4.307

Kulik, J. A., Kulik, C.-L., & Carmichael, K. (1974). The Keller Plan in science teaching. *Science*, *183*(4123), 379–383. JSTOR.

Lambert, M. C., Cartledge, G., Heward, W. L., & Lo, Y. (2006). Effects of response cards on disruptive behavior and academic responding during math lessons by fourth-grade urban students. *Journal of Positive Behavior Interventions*, *8*(2), 88–99.

Larsson, E. (2014). Focus on science: Some cautions on the exclusive use of standardized assessments in recovery-oriented treatment. *Science in Autism Treatment*, *11*(3), 22–25.

Layng, T. V. J. (2013). *Understanding Concepts: Implications for Science Teaching*. https://doi.org/10.13140/RG.2.2.19003.41765

Layng, T. V. J., & Twyman, J. S. (2014). Education + technology + innovation = learning? In *Handbook on innovations in learning.* (pp. 133–148). IAP Information Age Publishing.

Layng, T. V. J., Twyman, J. S., & Stikeleather, G. (2004). Engineering discovery learning: The contingency adduction of some precursors of textual responding in a beginning reading program. *The Analysis of Verbal Behavior*, *20*(1), 99–109. https://doi.org/10.1007/BF03392997

Layng, T. V., Twyman, J., & Stikeleather, G. (2003). Headsprout Early Reading™: Reliably teaching children to read 1. *Behavioral Technology Today*, *3*.

Lengel, J. G. (2015). *Education 3.0 Seven Steps to Better Schools.* Teachers College Press.

Lindsley, O. R. (n.d.). Theoretical basis for behavioral modification. In C. E. Pitts (Ed.), *Operant Conditioning in the Classroom: Introductory Readings in Educational Psychology* (pp. 54–60). Thomas Y. Crowell Co.

Lindsley, Ogden R. (1991). Precision teaching's unique legacy from B. F. Skinner. *Journal of Behavioral Education*, *1*(2), 253–266. https://doi.org/10.1007/BF00957007

Lindsley, Ogden R. (1992a). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis*, *25*(1), 51–57. https://doi.org/10.1901/jaba.1992.25-51

Lindsley, Ogden R. (1992b). Why aren't effective teaching tools wildly adopted? *Journal of Applied Behavior Analysis*, *25*(1), 21–26. https://doi.org/10.1901/jaba.1992.25-21

Lindsley, Ogden R. (1996). The four free-operant freedoms. *The Behavior Analyst*, *19*(2), 199–210.

Lloyd, M. E., & Lloyd, K. E. (1986). Has lightning struck twice? Use of PSI in college classrooms. *Teaching of Psychology*, *13*(3), 149–151. https://doi.org/10.1207/s15328023top1303_14

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, *55*(1), 3.

Lu, P. H. (1976). Teaching human growth and development by the personalized system of instruction. *Teaching of Psychology*, *3*(3), 127–128. https://doi.org/10.1207/s15328023top0303_6

Malanga, P. R., & Sweeney, W. J. (2008). Increasing active student responding in a university applied behavior analysis course: The effect of daily assessment and response cards on end of week quiz scores. *Journal of Behavioral Education*, *17*(2), 187–199.

Markle, S. M. (1983). *Designs for instructional designers*. Stipes Pub. Co.

Marmolejo, E. K., Wilder, D. A., & Bradley, L. (2004). A preliminary analysis of the effects of response cards on student performance and participation in an upper division university course. *Journal of Applied Behavior Analysis*, *37*(3), 405–410.

Martin, T. L., Pear, J. J., & Martin, G. L. (2002a). Analysis of proctor marking accuracy in a computer-aided personalized system of instruction course. *Journal of Applied Behavior Analysis*, *35*(3), 309–312. https://doi.org/10.1901/jaba.2002.35-309

Martin, T. L., Pear, J. J., & Martin, G. L. (2002b). Feedback and its effectiveness in a computer-aided personalized system of instruction course. *Journal of Applied Behavior Analysis*, *35*(4), 427–430. https://doi.org/10.1901/jaba.2002.35-427

Martinez-Diaz, J. A., & Wilder, D. A. (2016). Conceptualization and development of the behavior analysis programs at the Florida Institute of Technology. *The Behavior Analyst*, *39*(1), 149–156. https://doi.org/10.1007/s40614-016-0060-y

Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. (2014). *Behavior analysis for lasting change* (Third Edition). Sloan Publishing.

Mayer, G. R., Sulzer-Azaroff, B., & Wallace, M. (2014). Teaching new behavior: Shaping. In *Behavior Analysis for Lasting Change* (Third Edition, pp. 248–269). Sloan Publishing.

McDade, C., & Brown, J. (2001). Celerating student skills: Basic and advanced. *Jacksonville, AL: Author*.

McDade, C. E., Austin, D., & Olander, C. (1985). Technological advances in precision teaching: A comparison between computer testing and SAFMEDS. *Journal of Precision Teaching*, *6*(3), 49–53.

McMichael, J. S., Brock, J. F., & Delong, J. (1976). Job-relevant navy training and Keller's personalized system of instruction: Reduced attrition. *Journal of Personalized Instruction*, *1*(1), 41–44.

Meek, R. L. (1977). The traditional in non-traditional learning methods. *Journal of Personalized Instruction*, *2*(2), 114–118.

Merbitz, C., Vietez, D., Merbitz, N. K., & Binder, C. (2004). Precision teaching: Application in education and beyond. In D. J. Moran & R. W. Malott (Eds.), *Evidence-based educational methods* (pp. 47–62). Elsevier Academic Press.

Merbitz, N. K., Merbitz, C., Binder, C., & Vietez, D. (2016). *Applications of Precision Teaching in education and beyond*. https://doi.org/10.13140/RG.2.1.4584.2167

Michael, J. (1991). A behavioral perspective on college teaching. *The Behavior Analyst*, *14*(2), 229–239. https://doi.org/10.1007/BF03392578

Morvac, J. W. (2008, February 15). *Moving beyond education 2.0*. https://www2.educationfutures.com/blog/2008/02/moving-beyond-education-20/

Narayan, J. S., Heward, W. L., Gardner III, R., Courson, F. H., & Omness, C. K. (1990). Using response cards to increase student participation in an elementary classroom. *Journal of Applied Behavior Analysis*, *23*(4), 483–490.

No Child Left Behind Act (NCLB) of 2001, Pub. L. No. 107–110, 1425 Stat. (2002).

Northup, J., Vollmer, T. R., & Serrett, K. (1993). Publication trends in 25 years of the Journal of Applied Behavior Analysis. *Journal of Applied Behavior Analysis*, *26*(4), 527–537. https://doi.org/10.1901/jaba.1993.26-527

O'Neill, J., Rehfeldt, R. A., Ninness, C., Muñoz, B. E., & Mellor, J. (2015). Learning Skinner's verbal operants: Comparing an online stimulus equivalence procedure to an assigned reading. *The Analysis of Verbal Behavior*, *31*(2), 255–266. https://doi.org/10.1007/s40616-015-0035-1

Owens, C. J. (2016). Toward an analysis of concepts and solutions to training concepts. *Behavioral Development Bulletin*, *21*(1), 63–73. https://doi.org/10.1037/bdb0000021

Paiva, R. C., Ferreira, M. S., & Frade, M. M. (2017). Intelligent tutorial system based on personalized system of instruction to teach or remind mathematical concepts: Intelligent tutorial system. *Journal of Computer Assisted Learning*, *33*(4), 370–381. https://doi.org/10.1111/jcal.12186

Patching, W., Kameenui, E., Carnine, D., Gersten, R., & Colvin, G. (1983). Direct instruction in critical reading skills. *Reading Research Quarterly*, *18*(4), 406. https://doi.org/10.2307/747376

Pear, J. (2003). Enhanced feedback using computer-aided personalized system of instruction. In W. Buskist, B. K. Hevern, B. K. Saville, & T. E. Zinn (Eds.), *Essays from excellence in teaching* (pp. 68–72).

Pear, J. J., & Novak, M. (1996). Computer-aided personalized system of instruction: A program evaluation. *Teaching of Psychology*, *23*(2), 119–123. https://doi.org/10.1207/s15328023top2302_14

Pear, J., & Kinsner, W. (1988). Computer-aided personalized system of instruction: An effective and economical method for short- and long-distance education. *Machine-Mediated Learning*, *2*.

Pennypacker, H. S., & Hench, L. L. (1997). Making behavioral technology transferable. *The Behavior Analyst*, *20*(2), 97–108. https://doi.org/10.1007/BF03392767

Pereira, J. A., Pleguezuelos, E., Merí, A., Molina-Ros, A., Molina-Tomás, M. C., & Masdeu, C. (2007). Effectiveness of using blended learning strategies for teaching and learning human anatomy. *Medical Education*, *41*(2), 189–195. https://doi.org/10.1111/j.1365-2929.2006.02672.x

Perez, W. F., & de Rose, J. C. (2010). Recombinative generalization: An exploratory study in musical reading. *The Analysis of Verbal Behavior*, *26*(1), 51–55. https://doi.org/10.1007/BF03393082

Potts, L., Eshleman, J. W., & Cooper, J. O. (1993). Ogden R. Lindsley and the historical development of precision teaching. *The Behavior Analyst*, *16*(2), 177–189.

Pritchard, T., Penix, K., Colquitt, G., & McCollum, S. (2012). Effects of a weight training personalized system of instruction course on fitness levels and knowledge. *The Physical Educator; Vol 69, No 4 (2012)*. https://js.sagamorepub.com/pe/article/view/2897

Pytte, C. L., & Fienup, D. M. (2012). Using equivalence-based instruction to increase efficiency in teaching neuroanatomy. *Journal of Undergraduate Neuroscience Education: JUNE: A Publication of FUN, Faculty for Undergraduate Neuroscience*, *10*(2), A125–A131. PubMed.

Querol, B. I. D., Rosales, R., & Soldner, J. L. (2015). A comprehensive review of interteaching and its impact on student learning and satisfaction. *Scholarship of Teaching and Learning in Psychology*, *1*(4), 390–411. https://doi.org/10.1037/stl0000048

Quigley, S. P., Peterson, S. M., Frieder, J. E., & Peck, K. M. (2018). A review of SAFMEDS: Evidence for procedures, outcomes and directions for future research. *Perspectives on Behavior Science*, *41*(1), 283–301. https://doi.org/10.1007/s40614-017-0087-8

Randolph, J. J. (2007). Meta-analysis of the research on response cards: Effects on test achievement, quiz achievement, participation, and off-task behavior. *Journal of Positive Behavior Interventions*, *9*(2), 113–128.

Rehfeldt, R. A. (2011). Toward a technology of derived stimulus relations: An analysis of articles published in the Journal of Applied Behavior Analysis, 1992–2009. *Journal of Applied Behavior Analysis*, *44*(1), 109–119. https://doi.org/10.1901/jaba.2011.44-109

Reiser, R. A. (1984). Reducing student procrastination in a personalized system of instruction course. *ECTJ*, *32*(1), 41–49. https://doi.org/10.1007/BF02768768

Ricciardi, J. N., Luiselli, J. K., & Camare, M. (2006). Shaping approach respones as intervention for specific phobia in a child with autism. *Journal of Applied Behavior Analysis*, *39*(4), 445–448. https://doi.org/10.1901/jaba.2006.158-05

Rogers, E. M. (1995a). *Diffusion of Innovations, 4th Edition.* Free Press.

Rogers, E. M. (1995b). *Diffusion of Innovations, 4th Edition.* Free Press.

Rosales, R., Soldner, J. L., & Crimando, W. (2014). Enhancing the impact of quality points in interteaching. *Journal of the Scholarship of Teaching and Learning*, 1–11. https://doi.org/10.14434/josotlv14i5.12746

Rosenshine, B. V., & Berliner, D. C. (1978). Academic engaged time. *British Journal of Teacher Education*, *4*(1), 3–16.

Ross, L. L., & McBean, D. (1995). A comparison of pacing contingencies in classes using a personalized system of instruction. *Journal of Applied Behavior Analysis*, *28*(1), 87–88. https://doi.org/10.1901/jaba.1995.28-87

Rumbaugh, D. M. (1995). Emergence of relations and the essence of learning: A review of Sidman’s equivalence relations and behavior: A research story. *The Behavior Analyst*, *18*(2), 367–375. https://doi.org/10.1007/BF03392726

Saville, B. K., Cox, T., O'Brien, S., & Vanderveldt, A. (2011). Interteaching: The impact of lectures on student performance. *Journal of Applied Behavior Analysis, 44*(4), 937–941. https://doi.org/10.1901/jaba.2011.44-937

Saville, B. K., Lambert, T., & Robertson, S. (2011). Interteaching: Bringing behavioral education into the 21st century. *The Psychological Record*, *61*(1), 153–165. https://doi.org/10.1007/BF03395752

Saville, B. K., & Zinn, T. E. (2009). Interteaching: The effects of quality points on exam scores. *Journal of Applied Behavior Analysis*, *42*(2), 369–374. https://doi.org/10.1901/jaba.2009.42-369

Selinske, J. E., Greer, R. D., & Lodhi, S. (1991). A functional analysis of the comprehensive appliation of behavior analysis to schooling. *Journal of Applied Behavior Analysis*, *24*(1), 107–117. https://doi.org/10.1901/jaba.1991.24-107

Shabani, D. B., & Carr, J. E. (2004). An evaluation of response cards as an adjunct to standard instruction in university classrooms: A systematic replication and extension. *North American Journal of Psychology*, *6*(1), 85–100.

Sherman, J. G. (1982). PSI today. In F. S. Keller & J. G. Sherman (Eds.), *The PSI Handbook: Essays on personalized instruction* (pp. 72–78). TRI.

Sherman, J. Gilmour. (1992). Reflections on PSI: Good news and bad. *Journal of Applied Behavior Analysis*, *25*(1), 59–64. https://doi.org/10.1901/jaba.1992.25-59

Sidman, M. (1960). *Tactics of scientific research: Evaluating experimental data in psychology*. (pp. x, 428). Basic Books.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research*, *14*(1), 5–13. https://doi.org/10.1044/jshr.1401.05

Sidman, M. (1994). *Equivalence relations and behavior: A research story*. (pp. xvii, 606). Authors Cooperative.

Sidman, M. (1997). Equivalence relations. *Journal of the Experimental Analysis of Behavior*, *68*(2), 258–266. https://doi.org/10.1901/jeab.1997.68-258

Siedentop, D. (1973). How to use personalized system of instruction in college teaching. In L. L. Gedvilas (Ed.), *National College Physical Education Association For Men. Proceedings of Annual Meeting (77th, Kansas City, Missouri, December 26-29, 1973)* (pp. 116–125). Office of Publication Services. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.929.9067&rep=rep1&type=pdf#page=120

Skinner, B. F. (1957). *Verbal behavior*. Appleton-Century-Crofts.

Skinner, B. F. (1968). *The technology of teaching.* Appleton-Century-Crofts.

Slezak, S. A. (1969). *Two Years of Precisely Teaching with Orthopedically Handicapped Pupils* [Unpublished Master's Thesis]. University of Kansas.

Sokolove, H. (1978). *Blueprint for Productive Classrooms* (ESSA Title IV-C). Kansas State Department of Education.

Sota, B. (2016a). Co-designing instruction with students. In M. Murphy, S. Redding, & J. Twyman (Eds.), *Handboook on personalized learning for states, districts, schools* (pp. 57–71). Temple University, Center on Innovations in Learning. www.centeril.org

Sota, B. (2016b). Flipped learning as a path to personalization. In M. Murphy, S. Redding, & J. Twyman (Eds.), *Handbook on personalized learning for states, districts, and schools* (pp. 73–87). Temple University, Center on Innovations in Learning. www.centeril.org

Spector, L. C. (1976). The effectiveness of personalized system of instruction in economics. *Journal of Personalized Instruction*, *1*(2), 118–122.

Springer, C. R., & Pear, J. J. (2008). Performance measures in courses using computer-aided personalized system of instruction. *Computers & Education*, *51*(2), 829–835. https://doi.org/10.1016/j.compedu.2007.09.002

Starlin, A. (1971). TIX 43: Charting group and individual instruction. *TEACHING Exceptional Children*, *3*(3), 135–140. https://doi.org/10.1177/004005997100300309

Starlin, C. M. (1970). *The use of daily direct recording as an aid in teaching oral reading* [Doctoral dissertation]. University of Oregon.

Stowell, J. R., & Nelson, J. M. (2007). Benefits of electronic audience response systems on student participation, learning, and emotion. *Teaching of Psychology*, *34*(4), 253–258.

Sulzer-Azaroff, B. (2004). The shaping of behaviorists: B.F. Skinner's influential paper on teaching machines. *European Journal of Behavior Analysis*, *5*(2), 129–135. https://doi.org/10.1080/15021149.2004.11434239

Sulzer-Azaroff, B., Fleming, R., Tupa, M., Bass, R., & Hamad, C. (2008). Choosing objectives for a distance learning behavioral intervention in autism curriculum. *Focus on Autism and Other Developmental Disabilities*, *23*(1), 29–36. https://doi.org/10.1177/1088357607311442

Sulzer-Azaroff, B., & Gillat, A. (1990). Trends in behavior analysis in education. *Journal of Applied Behavior Analysis*, *23*(4), 491–495. https://doi.org/10.1901/jaba.1990.23-491

Sulzer-Azaroff, B., & Harshbarger, D. (1995). Putting fear to flight: While enhancing quality of performance. *Quality Progress*, *28*(12), 61–65.

Svenningsen, L., & Pear, J. J. (2011). Effects of computer-aided personalized system of instruction in developing knowledge and critical thinking in blended learning courses. *The Behavior Analyst Today*, *12*(1), 34–40. https://doi.org/10.1037/h0100709

Trees, A. R., & Jackson, M. H. (2007). The learning environment in clicker classrooms: Student processes of learning and involvement in large university-level courses using student response systems. *Learning, Media and Technology*, *32*(1), 21–40.

Twyman, J. S., Layng, T. V. J., Stikeleather, G., & Hobbins, K. A. (2005). A non-linear approach to curriculum design: The role of behavior analysis in building and effective reading program. In W. L. Heward, T. E. Heron, N. A. Neef, S. M. Peterson, D. M. Sainato, G. Cartledge, R. Gardner III, L. D. Peterson, S. B. Hersh, & J. C. Dardig (Eds.), *Focus on behavior analysis in education: Achievements, challenges, and opportunities* (pp. 55–68). Pearson/Merrill/Prentice Hall.

Twyman, Janet S. (1998). The Fred S. Keller School. *Journal of Applied Behavior Analysis*, *31*(4), 695–701. https://doi.org/10.1901/jaba.1998.31-695

Twyman, Janet S. (2014). Behavior analysis in education. In *The Wiley Blackwell handbook of operant and classical conditioning.* (pp. 533–558). Wiley-Blackwell. https://doi.org/10.1002/9781118468135.ch21

Twyman, Janet S. (2014). Competency-based education: Supporting personalized learning. Connect: Making learning personal. *Center on Innovations in Learning, Temple University*. https://files.eric.ed.gov/fulltext/ED558055.pdf

Twyman, Janet S., & Heward, W. L. (2018). How to improve student learning in every classroom now. *International Journal of Educational Research*, *87*, 78–90. https://doi.org/10.1016/j.ijer.2016.05.007

Underwood, B. J., & Ekstrand, B. R. (1966). An analysis of some shortcomings in the interference theory of forgetting. *Psychological Review*, *73*(6), 540–549. https://doi.org/10.1037/h0023883

U.S. Department of Education. (n.d.). *Competency-based learning or personalized learning* [Ed.gov]. https://www.ed.gov/oii-news/competency-based-learning-or-personalized-learning

Vance Hall, R. (1991). Behavior analysis and education: An unfulfilled dream. *Journal of Behavioral Education*, *1*(3), 305–315. https://doi.org/10.1007/BF00947185

Varelas, A., & Fields, L. (2017). Equivalence based instruction by group based clicker training and sorting tests. *The Psychological Record*, *67*(1), 71–80. https://doi.org/10.1007/s40732-016-0208-x

Vargas, E. A., & Vargas, J. S. (1991). Programmed instruction: What it is and how to do it. *Journal of Behavioral Education*, *1*(2), 235–251. JSTOR.

Weisman, R. A., & Shapiro, D. M. (1973). Personalized system of instruction (Keller method) for medical school biochemistry. *Academic Medicine*, *48*(10). https://journals.lww.com/academicmedicine/Fulltext/1973/10000/Personalized_system_of_instruction__Keller_method_.7.aspx

White, W. A. (1988). A meta-analysis of the effects of direct instruction in special education. *Education and Treatment of Children*, 364–374.

Willis, B. (1974). *Progress Report IV SST for Precise Education Managers of Specific Learning Disabilities* (Part G; pp. 91–230). Child Service Demonstration Programs.

Yaoyuneyong, G., & Thornton, A. (2011). Combining peer instruction and audience response systems to enhance academic performance, facilitate active learning and promote peer-assisted learning communities. *International Journal of Fashion Design, Technology and Education*, *4*(2), 127–139.

Zencius, A. H., Davis, P. K., & Cuvo, A. J. (1990). A personalized system of instruction for teaching checking account skills to adults with mild disabilities. *Journal of Applied Behavior Analysis*, *23*(2), 245–252. https://doi.org/10.1901/jaba.1990.23-245