



Precision Teaching and Tap Dance Instruction

Marisela Pallares¹  · Kendra Brooks Newsome² · Patrick M. Ghezzi¹

Published online: 14 August 2020

© Association for Behavior Analysis International 2020

Abstract

The purpose of this study was (a) to use a precision-teaching (PT) framework to design, train, and evaluate a tap-dancing training sequence and (b) to evaluate fluency outcomes as a function of training tap dance components to optimal frequencies. The study trained a series of 8 tap-dancing steps to 4 novice dancers and evaluated the effects on untrained components and probes of retention, stability, endurance, and application. The study also included a control participant who only completed application probes. Weekly probes examining the facilitative effects of training on the untrained components revealed improvements for some untrained steps, but not all. Retention probes revealed little difference in frequencies from the last data point in training. Stability and endurance probes revealed marked increases in the frequency of corrects and decreases in the frequency of errors. The results of application probes showed improvements to some degree for experimental participants; however, the control participant also made gains in performance. This makes it difficult to draw conclusions regarding application. The study demonstrates how a PT framework may be useful to those interested in enhancing sports performance training. We discuss limitations and future directions.

Keywords Component-composite analysis · Dance · Fine motor · Fluency · Fluency outcomes · Frequency building · Precision teaching · Sports performance

Applied behavior analysis has an expansive history of using behavioral techniques to enhance the training and performance of athletes in a long list of sports. The list includes football (Allison & Ayllon, 1980; Komaki & Barnett, 1977; Smith & Ward, 2006; Stokes, Luiselli, Reed, & Fleming, 2010; Tai & Miltenberger, 2017; Ward & Carnes, 2002), tennis (Allison & Ayllon, 1980; Ziegler, 1987), swimming (Hume & Crossman, 1992; Koop & Martin, 1983; Rogers, Hemmeter, & Wolery, 2010; Schonwetter, Miltenberger, & Oliver, 2014), soccer (Brobst & Ward, 2002; Ziegler, 1994), gymnastics (Allison & Ayllon, 1980; Boyer, Miltenberger, Batsche, & Fogel, 2009), track and field (Scott & Scott, 1997), basketball (Kladopoulos & McComas, 2001), in-line speed skating (Anderson & Kirkpatrick, 2002), martial arts (Harding, Wacker, Berg, Rick, & Lee, 2004), rugby (Mellalieu, Hanton, & O'Brien, 2006), and many more (Schenk & Miltenberger, 2019).

Behavioral techniques have also enhanced skills using many procedures and intervention packages (Luiselli, Woods, & Reed, 2011; Schenk & Miltenberger, 2019). In a review by Schenk and Miltenberger (2019), they found that the enhancement of sports performance has been targeted using procedures such as consequence and antecedent manipulations, feedback procedures, skills training, and unique procedures such as habit reversal and acceptance and commitment therapy. Some examples of intervention packages used include positive and negative reinforcement, video modeling, teaching with acoustical guidance, public posting, goal setting, verbal and video feedback, self-monitoring, simulated practice, behavioral rehearsal, relaxation, discrimination training, and many more (Luiselli et al., 2011; Schenk & Miltenberger, 2019).

Among this volume of literature, there is a small number of published studies focused on improving dance behavior. Auditory feedback (Quinn, Miltenberger, Abreu, & Narozanick, 2017; Quinn, Miltenberger, & Fogel, 2015; Quinn, Sherman, Sheldon, Quinn, & Harchik, 1992), error correction procedures (Fitterling & Ayllon, 1983; Vintere & Poulson, 2010), and self-instruction procedures (Vintere &

✉ Marisela Pallares
sela@nevada.unr.edu

¹ University of Nevada, Reno, NV, USA

² Fit Learning, LLC, Reno, NV, USA

Poulson, 2010) have been shown to improve the accuracy of dance performance. However, only one study on dance instruction has adopted a rate-based measure. Lokke, Lokke, and Arntzen (2008) recognized the importance of shaping both the accuracy and speed of responding to improve dance performance. Lokke et al. (2008) worked with a ballet dancer to correct a faulty component (i.e., three jumps on the left foot that ended in a 90-degree turn) that interfered with the dancer's accurate performance of a dance routine. Specifically, they used timed-practice intervals (i.e., timings) of 15 s along with performance feedback and modeling to help the dancer master the faulty component. They found that once the dancer was performing the jumps accurately and at the same rate as that of an experienced dancer, she was able to complete the routine flawlessly. Moreover, the amount of practice time totaled 21 timings across 9 days, demonstrating the efficiency of the intervention.

The importance of the Lokke et al. (2008) study is that it introduced the framework of precision teaching (PT) to dance instruction. Though many of the interventions for enhancing sports performance have been effective in building the accuracy of specified responses during interventions, there have been limitations in terms of maintenance of trained responses and generalization to untargeted responses (Luiselli et al., 2011). PT has a long history of producing important learning outcomes in classrooms and other educational settings with a variety of learners using a framework that emphasizes rate of response (frequency) displayed on the standard celeration chart (SCC), continuous real-time monitoring for effective data-based decision making, and an approach to instructional design based on a component-composite analysis (Binder, 1996; Johnson & Street, 2013; Kubina & Yurich, 2012). Therefore, a PT approach to dance instruction, and sports performance more generally, may yield the learning outcomes that, according to Luiselli et al. (2011), have been missing from this literature base.

Precision Teaching and Behavioral Fluency

Simply put, PT involves making instructional decisions on continuous performance frequencies displayed on the SCC (Lindsley, 1992a, 1992b). This emphasis on sensitive measurement gave rise to a general process of data-based decision making that can be summed up in four steps: pinpoint, record, change, and try again (Johnson & Street, 2013; Kubina & Yurich, 2012). This process is a continuous cycle of defining objectives (pinpoint), providing practice opportunities and recording frequencies (record), changing strategies or interventions based on performance (change), and in cases in which the first change does not work, persisting in systematically changing variables and observing their effects on learner performance (try again). This process is a perpetual exploration

and refinement of instructional strategies tailored to maximize efficient learning. Through this process, precision teachers have discovered that building responses to optimal rates of responding leads to behavioral fluency (Haughton, 1972, 1980; Johnson & Layng, 1992).

Behavioral fluency is defined by the important learning outcomes that emerge at optimal performance frequencies (Binder, 1996; Johnson & Layng, 1992, 1996; Johnson & Street, 2013; Kubina & Yurich, 2012). In other words, the assessment of mastery is post hoc, and certain observable features participate in the evaluation. Specifically, a mastered skill will *retain* over time without practice, *endure* when performed for sustained periods, *persist* (i.e., *stability*) in conditions with competing stimuli, and can be *applied* to (i.e., facilitates) learning related skills (RESA; Berens, Boyce, Berens, Doney, & Kenzer, 2003; Berens & Hayes, 2007; Brady & Kubina, 2010; Dembek & Kubina, 2018; Haughton, 1972; Carl Hughes, Beverley, & Whitehead, 2007; Kubina, 2005; Newsome, Berens, Ghezzi, Aninao, & Newsome, 2014; Stocker, Schwartz, Kubina, Kostewicz, & Kozloff, 2019; Twarek, Cihon, & Eshleman, 2010; Weiss, Foley, Pearson, & Pahl, 2010).

Anyone who has played a sport, learned a musical instrument, or danced recognizes the importance of practice to master the skills of those domains (Binder, 1996). One way of practicing that precision teachers commonly use to shape fluent behavior is frequency building. Frequency building involves timed repetition of behavior and performance feedback after the timed interval to build the frequency of correct responses to specified frequency aims, while simultaneously establishing low frequencies of incorrect responses (Kubina & Yurich, 2012). Practicing in this manner typically leads to behavior that seems effortless, automatic, or masterful—words typically used to describe fluent behavior (Binder, 1996; Haughton, 1972).

In PT, the explicit purpose of practice is to build the frequencies of component skills rather than composite skills (Johnson & Street, 2013). Research has shown that building the frequency of component skills in a training sequence generally facilitates the acquisition of more complex skills and can have facilitative effects across nontargeted repertoires (Berens & Hayes, 2007; Brosnan, Moeyaert, Newsome, & Healy, 2018; Haughton, 1972, 1980; Carl Hughes et al., 2007; Johnson & Layng, 1992; Johnson & Street, 2013; Lokke et al., 2008; McDowell, McIntyre, Bones, & Keenan, 2002; McTiernan, Holloway, Healy, & Hogan, 2016; Newsome et al., 2014; Twarek et al., 2010). For example, the performance of basic arithmetic computations (i.e., addition, subtraction, single-digit multiplication, and division) that is both accurate and sufficiently fast aids in the acquisition of related, but more complex skills, such as long division. Similarly, in sports, a basketball player who learns how to dribble a ball at an optimal frequency will more easily learn to pass and shoot

the ball from various positions on the court in order to play effectively in a game. Thus, component-composite approaches to instructional design are advantageous in that the arrangement of skills in an instructional sequence is an important variable that impacts the efficiency of learning. Addressing dysfluency at the root cause prevents cumulative dysfluency (the compound effects of layering fluency-deficient component skills on top of one another), which can limit or even halt the acquisition of composite skills (Binder, 1996; McDowell & Keenan, 2001).

Precision Teaching Research

The majority of PT and fluency literature focuses on education and academic skills (Berens et al., 2003; Brady & Kubina, 2010; Brosnan et al., 2018; Cavallini & Perini, 2009; Chiesa & Robertson, 2000; Griffin & Murtagh, 2015; Carl Hughes et al., 2007; Lambe, Murphy, & Kelly, 2015; McDowell & Keenan, 2001; McTiernan et al., 2016; Mercer, Campbell, Miller, Mercer, & Lane, 2000; Milyko, Berens, & Ghezzi, 2012; Newsome et al., 2014; Sleeman, Friesen, Tyler-Merrick, & Walker, 2019). Research on nonacademic skills is small by comparison and limited to a handful of studies on fine-motor behavior and sports (Fabrizio, Schirmer, King, Diakite, & Stovel, 2007; Lokke et al., 2008; McDowell et al., 2002; Twarek et al., 2010; Weiss et al., 2010). The most widely taught nonacademic skills are fine-motor responses necessary to improve handwriting and daily living skills. Eric Haughton (1972) and his colleagues (Binder, 1996) pioneered the work on the “Big 6 + 6.” They discovered that strengthening a series of 12 fine-motor responses (i.e., reach, touch, point, grasp, place, release, push, pull, shake, squeeze, tap, twist) leads to better acquisition of complex skill sets of daily living. Innovative applications of the Big 6 + 6 approach includes one study by Twarek et al. (2010), on teaching young children with autism to dress themselves, and a second study by Fabrizio et al. (2007), on teaching a young boy with autism how to squeeze an object, a deficit that prevented the child from playing with toys that required this response. The Big 6 + 6 approach has also proven to be successful in strengthening hand dexterity for sign language purposes with people suffering from traumatic brain injury (Chapman, Ewing, & Mozzoni, 2005).

PT has also contributed to the area of sports, though in a limited capacity. Aside from the Lokke et al. (2008) article on dance, there is one other demonstration of PT and fluency-based instruction applied to sports. McDowell et al. (2002) improved the swing of golfers by strengthening components of a golf swing in isolation (e.g., backswing, follow-through). The participant’s golf swing improved even though it was never directly targeted. Both studies demonstrated a PT-oriented solution to a sports performance problem.

These studies are a promising start toward answering the call from the sports performance literature to optimize learning outcomes, a hallmark of PT. Moreover, given the outcomes that the PT framework has yielded in other areas, it is time to expand this approach to sports and other motor-learning repertoires. Tap dancing is an ideal domain for investigating fluency through the PT framework, as it requires expert performers to tap at high rates. Applying a PT framework to tap dance instruction may reveal relevant component-composite relations that can help improve the efficiency of tap dance training, as well as provide an additional demonstration of how the major concepts of PT apply to sports performance.

The current study, therefore, sought to evaluate fluency outcomes with motor learning and to explore whether a PT approach to the teaching of tap dancing would yield efficient learning across a training sequence. Specifically, the purpose of this study was (a) to use a PT framework in the design, teaching, and evaluation of an instructional sequence to teach tap dancing and (b) to examine the facilitative effects of frequency building on tap dance components on untrained steps and probes of retention, endurance, stability, and application (RESA).

Method

Participants

Five adults participated in the study. The experimental participants included two females and two males. Daisy and Tina, the female participants, were both 23 years old at the time of the study. Andy was 25 years old, and Berry was 20 years old; both identified as male. Jack, the constant-series probe participant, was an 18-year-old male. No participants reported that they had dance or music experience, except for Andy, who reported he had played the trumpet in grade school. All participants reported that they had played organized sports as children and teenagers. Only one participant (Andy) continued to play an organized sport (basketball) as an adult.

Materials and Setting

All sessions took place at the University of Nevada, Reno, in a dance room equipped with one wall lined with mirrors and a wooden dance floor. Materials included data collection sheets, writing utensils, a video camera, scripts for the instructor, tap dance shoes for the participants and the instructor, a ThinkPad laptop computer with tap dance analysis software, an external microphone, wireless headphones, and an iPhone.

Response Targets

The tap dance steps used in the study are from the beginning curriculum at the Fascinating Rhythm School of Performing Arts, located in Reno, Nevada. Table 1 contains detailed descriptions of basic tap steps, the selected target responses, and the application combinations in this study. Note that not all training steps are part of the application combinations, but all training steps are in one or the other combination. Training steps were organized in pairs and placed in a presumed hierarchy of increasing complexity and taught sequentially. Toe taps and heel taps constituted Pair 1. Tip steps and dig steps formed Pair 2. Pair 3 included shuffles right and shuffles left. Shuffle steps and windshield wipers made up Pair 4. Two of the participants, Daisy and Andy, received some training on the fourth pair, but researchers revised the training sequence as a function of their data, and from that point forward, all participants experienced the training sequence in the order mentioned previously.

Dependent Variable

The frequencies of correct and incorrect responding on weekly probes of the untrained components after the current training component in the training sequence, as well as the frequencies of correct responding and incorrect responding on RESA probes, served as the dependent variables. Weekly probes served to evaluate generalization to untrained components in the training sequence, and RESA probes served to evaluate the emergence of fluency as the participant reached the performance criterion for each component step in the training sequence.

Measurement

Frequency of correct sounds per minute and frequency of incorrect sounds per minute were the measures of performance for both training steps and probes. A correct sound was any tap sound with the part of the foot, the rhythm, and the order appropriate for the response target. There were two types of errors: missing sounds and extra sounds. A missing sound occurred when a participant attempted a step but did not contact the floor and therefore failed to make the sound associated with a given response target. Observers also counted a missing step when a step that was part of the sequence did not occur at all. If a participant made an error and restarted a combination, the steps that would have followed were not considered errors. Extra sounds were sounds produced by the dancer that were not appropriate for a given response target. This included sounds that accompanied topographical errors made by the participants.

Researchers used a software program designed to capture and analyze the sound of the taps. The software was also used

to provide performance feedback to the participants. The apparatus audio-recorded each tap and counted the total number of taps that occurred during a timing. Because the software could not capture the absence of a tap or count a tap that occurred when it should not have occurred, the experimenter counted the total number of incorrect taps.

The software program provided a prompt after each timing so that the experimenter could enter the number of errors made. The software then calculated the total number of correct responses and the total number of incorrect responses based on the total number of taps recorded and the total number of incorrect taps inputted by the experimenter.

The software gave a visual display of the number of total taps, the number of correct taps, and the number of extra or missing taps. Figure 1 shows an example of the screen the experimenter saw at the end of a timing. The participants would see the screen in the frequency-building phase only. The sound wave produced by the taps would appear in blue. Vertical green lines were overlaid on the sound wave to indicate the position of the taps. The closer the green lines appeared, the faster the participant tapped. The farther apart the lines appeared, the slower the participant tapped. Evenly spaced lines indicated a steady pace, whereas unevenly spaced lines indicated a variable pace.

The distance between the green lines revealed whether the participant tapped at the appropriate pace. For example, a shuffle required the participant to perform a syncopation. If the participant executed the timing with accuracy, the pattern of syncopation would appear like the one in Fig. 1, where two green lines (taps) are close together (a syncopated shuffle), followed by a comparatively larger gap, which represents the time between the end of one shuffle and the beginning of the next shuffle.

This display was useful for delivering intertiming performance feedback to the participant. It allowed the instructor to speak specifically to instances in the timing when a participant paused, missed a sound, or made an extra sound. The display could also show where the participant may have slowed down, sped up, maintained, or fell out of rhythm, along with the participant's performance feedback in terms of the number of correctly executed sounds and the number of missing and/or extra sounds.

Research Design and Analysis

The study used a multiple-probe across-responses single-subject design. Researchers selected this design specifically because the training responses were presumably part of the same operant class, and thus expected the baselines of the untrained responses (as well as the frequencies of the fluency outcome probes) to change as a function of training on other responses in the training sequence (Horner & Baer, 1978). The design included one constant-series control participant

Table 1 Tap Steps Terminology

Response	Description
Basic Terminology	
Heel tap	Lifting the heel from the ground and allowing it to drop to the floor without lifting any other part of the foot (1 sound)
Toe tap	Lifting the toe from the ground and allowing it to drop to the floor without lifting any other part of the foot (1 sound)
Dig	Jabbing the heel into the floor (1 sound)
Tip	Jabbing the toe into the floor (1 sound)
Brush	Striking the ball of the foot in a forward motion (1 sound)
Spank	Striking the ball of the foot in a backward motion (1 sound)
Step	Stepping on the ball of the foot with full weight (1 sound)
Basic Combinations	
Flap	[Brush + Step] traveling forward or backward (2 sounds)
Shuffle	[Brush + Spank] remaining stationary (2 sounds)
Split shuffle	[Shuffle] with the accent on the front beat and a slight pause between the brush and the spank (2 sounds)
Training Pairs	
Toe taps	[Toe taps] alternating from one foot to the other (1 sound)
Heel taps	[Heel taps] alternating from one foot to the other (1 sound)
Dig steps	[Dig + Step] alternating from one foot to the other (2 sounds)
Tip steps	[Tip + Step] alternating from one foot to the other (2 sounds)
Shuffles right	[Shuffles] continuous on the right foot while balancing on the left foot (2 sounds)
Shuffles left	[Shuffles] continuous on the left foot while balancing on the right foot (2 sounds)
Shuffle step	[Shuffle + Step] alternating from one foot to the other (3 sounds)
Windshield wipers	[Toe tap (outward) + Toe tap (inward) + Heel tap (outward) + Heel tap (inward)] alternating from one foot to the other (4 sounds)
Application Combinations	
Exercise 6 (traveling)	[Shuffle] + [Heel] + [Flap + Heel] traveling forward and backward (6 sounds)
Exercise 8 (stationary)	[Split shuffle + Shuffle + Step + Heel] + Dig (8 sounds)

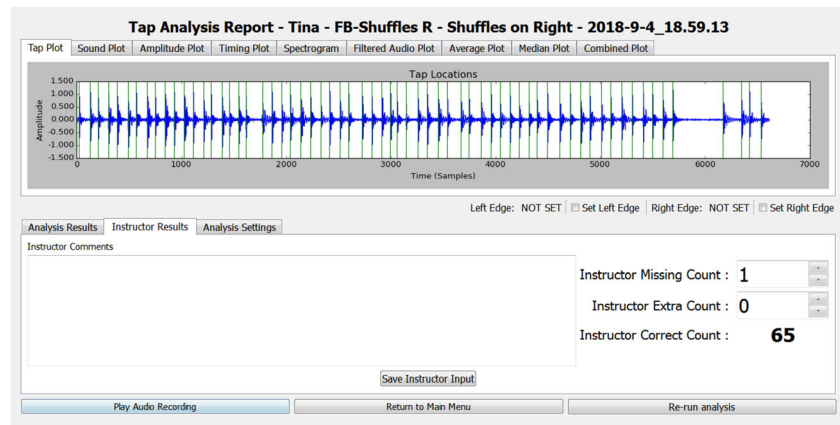
Note. The table provides definitions of basic terminology first. Then these basic terms explain basic combinations, training pairs, and application combinations. Brackets indicate movement on the same foot. A change in bracket indicates a change to the other foot. Addition signs indicate the next movement. For example, to perform Exercise 6, the participant would start with a shuffle on the right foot, then switch to a heel on the left foot, then do a flap followed by a heel on the right foot. The participant would then repeat this pattern starting with the left foot. The number of sounds that each step or combination makes is included in parentheses

who served as an uninstructed comparison to the effects of instruction for application combinations only (Hayes, Barlow, & Nelson-Gray, 1999). All steps entered baseline at the same time. Once the steps met the inclusion criteria, the first pair of training steps entered frequency building, and the next step pair simultaneously underwent weekly probes. Once the training step pair met the performance criterion, it entered RESA probes for retention, endurance, and stability. Application probes on the application combinations were also conducted during the RESA probes phase. Every training pair followed this same progression. The constant-series participant underwent probes for application combinations only.

First-timing data yielded celerations, frequency multipliers, accuracy ratios, and improvement index metrics for training and probe data. Researchers selected first-timing data because

these data are the most conservative measure of change across sessions. The authors used CR PrecisionX (CentralReach, LLC, 2019) online software, formerly known as Chartlytics, to generate celeration, first-last frequency multipliers, and improvement indices (also known as the accuracy improvement measure; see Pennypacker, Gutierrez, & Lindsley, 2003, for definitions of these measures). Accuracy ratios represent the directed distance from the incorrect frequency to the correct frequency. Accuracy ratios were calculated by dividing the frequency of corrects by the frequency of errors (Pennypacker et al., 2003). All accuracy ratios in this study were assigned a multiplicative sign because the frequency of errors was never higher than the frequency of corrects. Any additional frequency multipliers that appear in the Results section were calculated by dividing

Fig. 1 Screenshot of the software output screen. This is an example of the software's output screen after a timing. The screen shows what the participant and the experimenter saw at the end of the timing. The vertical green lines overlaid on the blue sound wave represent tap sounds. The output screen also indicates the number of correct responses, the number of missing sounds, and the number of extra sounds



the larger frequency (F1) by the smaller frequency (F2) in the comparison and assigning a sign depending on the direction of change on the chart to get from F1 to F2 (Pennypacker et al., 2003).

The study evaluated retention by calculating the percentage of change from the first data point of the last training session to the first data point in the retention probes for each training step (Bucklin, Dickinson, & Brethower, 2000). The percentage-of-change formula involved dividing the frequency of the training data point by the frequency of the retention probe data point and multiplying by 100. On the graphs, a percentage of change that represents growth in the frequency of corrects from the training phase to the retention phase appears on the positive scale, and a percentage of change that represents decay appears on the negative scale. This measure provides a rapid visual analysis of retention across training responses for each participant.

Procedures

Normative sampling Ten experienced tap dancers enrolled in advanced tap classes and, either competing on a tap dance team or working as a tap dance instructor, established performance criteria for all the dance step pairs used in the current study. The dancers completed two 15-s timings for all the tap steps and sequences. Then, the researchers calculated the average frequencies for each step based on the 10 dancers' frequencies and set performance criteria at ranges from 10% above the average to 10% below the average for each step. The inclusion criterion for steps entering training was 60% within the average frequency or less. Table 2 shows the aim ranges and the baseline criterion for each step.

Sessions The first author served as the instructor and conducted individual sessions that lasted approximately 10–15 min, four times per week, with each participant at the University of Nevada, Reno. All sessions started with scripted instructions that described the step and its important features. Then, the

instructor modeled the step for the participant and asked the participant to engage in the step. The instructor provided praise or corrective feedback after the participant performed the step. Once the participant demonstrated the step correctly for one eight-count, timings would begin. When a participant was ready to engage in a timing, the software would begin a countdown with four tones plus a visual display of numerals counting down. The participant began tapping at the fourth tone for the length appropriate for the condition. During timings, the software would simply display a countdown of the timing duration. The software would indicate the end of a timing with another tone. Then the participant either would or would not receive feedback, depending on whether it was a baseline session, a frequency-building session, or a probe session. The number of timings per response varied depending on the session condition. The standard baseline and frequency-building timing length was 15 s. Researchers chose this timing length for several reasons. First, in a natural setting, tap dancers are not expected to perform the same tap step for more than a few seconds. During normative sampling, we discovered that experienced dancers had difficulty sustaining their rates for more than approximately 15 s. Furthermore, the PT literature has specified that practicing in shorter timing lengths (sprints) has the same benefits as practicing with longer timing lengths, and it is more efficient because sprints take less time (Binder, 1996; Haughton, 1980; Kostewicz & Kubina, 2010). Finally, Lokke et al. (2008) also used 15-s timings.

Baseline Baseline sessions included three 15-s timings for each step. Participants did not receive feedback of any kind during baseline. This phase continued until one or all of the following four conditions were met: (a) the celeration of correct steps was equal to or less than $x1.2$, (b) variability was equal to or greater than $x2.0$, (c) there were relatively high rates of inaccurate steps, or (d) frequencies were no greater than 60% of the average frequencies established by the normative sampling group. The authors selected these criteria because baselines with one or more of these features would

Table 2 Normative Sampling Data

Tap Step	Average Frequency	10% Below Average	10% Above Average	60% of Aim
Toe taps	360	324	396	216
Heel taps	400	360	440	240
Tip steps	300	270	330	180
Dig steps	300	270	330	180
Shuffles right	320	288	352	192
Shuffles left	320	288	352	192
Shuffle steps	360	324	396	216
Windshield wipers	260	234	286	156
Exercise 6	320	288	352	192
Exercise 8	280	252	308	168

Note. The table shows the averages generated across each step based on the data from the 10 normative sampling dancers. The lower end of the aim range was determined by reducing the average by 10%, and the upper end by increasing the average by 10%. The inclusion criterion for steps entering training was 60% within the average frequency or less. The table shows where those frequencies would have to fall to meet this criterion

likely indicate dysfluent performance. Celerations at or below $\times 1.2$ are considered practice effects (Kubina & Yurich, 2012). Fluency outcomes typically emerge once frequencies of correct responding stabilize within the fluency aim, so responses with a high degree of variability are unlikely to yield fluency outcomes. Moreover, high rates of errors can also impact the emergence of fluency outcomes, and low error rates are a critical feature of dance performance. Finally, responses occurring at frequencies above 60% of the presumed fluency aim would have little room for growth; celerations naturally flatten as response frequencies approach the optimal frequency.

Frequency building Following the baseline phase, the steps chosen for training entered the frequency-building phase in pairs. Table 1 shows the pairs and the order in which they were trained for all participants under the heading “Training Pairs.” During frequency-building sessions, participants completed three 15-s timings for each step in the pair they were currently training on. The instructor provided goals before each timing. If the participant met the goal, the instructor provided performance feedback and praise after the timing. If the participant did not meet the goal, the instructor provided performance feedback and corrective feedback. Performance feedback included the total number of correct responses, the total number of incorrect responses, and a brief description of intertiming performance while looking at the visual display provided by the software. Praise involved short and concise statements (not exceeding five words) delivered in an enthusiastic voice that indicated the learner did well. Some examples of praise statements include “You did it,” “That was awesome,” or “Great dancing.” Corrective feedback involved describing the nature of the learner’s errors, modeling a strategy to correct the error, and providing the learner with an opportunity to rehearse the step again before moving to the next timing. The duration of feedback did not exceed 2 min.

Goal setting. The instructor established participants’ goals according to their overall performance in the study. The study used “personal best” goals. “Personal best” goals require participants to beat their best performance up to that point in the study (Binder & Sweeney, 2002; Carl Hughes et al., 2007; Lokke et al., 2008; Milyko et al., 2012). Personal best goals could be set for increasing the frequency of correct steps, decreasing the frequency of incorrect steps, or the combination of the two. For example, before a timing, if a participant was training on windshield wipers and the participant’s best performance up to that point had been 100 sounds per minute with no errors with a frequency aim of 234–286 sounds per minute, then the instructor would likely set the goal at 101 sounds per minute for that timing. If the participant beat that goal in the proceeding timing, then the instructor would use this new frequency to set a new goal for the next timing. The instructor reviewed previous performance on the SCC with each participant at the beginning of every session so that participants were aware of their overall performance and the basis for their goal. Visual inspection of learning pictures served to make determinations regarding goals. The instructor would set accuracy-only goals if the participants’ data trends showed accelerations in errors or high variability in errors. If data trends showed a deceleration in errors or low frequencies and low variability in errors, then the instructor would set frequency and accuracy goals.

Goal failures. If a participant failed to meet his or her goal for three consecutive timings (i.e., one session), the instructor would implement an intervention. Interventions included giving a prompt, isolating a component of the failed step(s), shortening the timing length, or revising a goal. The instructor would select the intervention based on the nature of the performance that led to

a failed goal. For example, a commonly used prompt for participants who had steady but low rates during timings included presenting an auditory model of the speed and pace required to meet a goal via a metronome. Another typical prompt used in the study involved putting a marking such as a colored sticker on the floor that the participant could use as a target. This prompt helped participants make their steps smaller and hence increase the number of steps they could perform in a timing. If high errors occurred due to a specific component of the step, the component was isolated and practiced in the timings. Once the isolated component reached the performance criterion (typically the same as the aim of the step it was a component of), timings of the target step resumed. If a step occurred at a high rate at the beginning of a timing but slowed down at the end of the timing, or vice versa, or if high and low rates occurred sporadically throughout a timing, the timing length was shortened.

Removing the intervention was contingent on reaching the performance criterion with the modification in place. Once the participant met the criteria for removing the modification, session contingencies reverted to the original procedure. Therefore, participants had to reach the performance criterion without the modification to move on in the experiment.

Performance criterion. The frequency-building phase ended when steps reached and sustained frequencies within the performance criterion (i.e., the frequency range demonstrated by the normative sampling participants) across two or more timings for three consecutive sessions. Once they demonstrated this criterion, trained steps entered the RESA probes to ensure the steps had achieved functional mastery.

On some occasions, however, steps were moved to the next phase before meeting the performance criterion based on improvements in the weekly probes coupled with flattened accelerations within 80% of the performance criterion. Performance aims are estimates of mastery that are based on performers who demonstrate proficiency in the skill and are generally thought of as guidelines (Binder, Haughton, & Bateman, 2002; White, 1984). Learners can demonstrate fluency below or above the specified fluency aim. Ultimately, what defines fluency is the demonstration of fluency outcomes (Binder, 1996; Binder et al., 2002). For this reason, and given our small normative sample size, we made data-based decisions guided by weekly probe data for participants whose performance stabilized below the performance criterion to move forward in the training sequence.

Weekly probes. While one pair entered frequency building, the next pair in the training sequence entered weekly

probes. For example, when Pair 1 (i.e., toe taps and heel taps) entered frequency building, Pair 2 (i.e., tip steps and dig steps) entered weekly probes. Weekly probes occurred in the first session of the week and consisted of two 15-s timings with no programmed feedback. Weekly probe conditions were exactly like baseline conditions because the purpose was to evaluate growth on untrained responses. All pairs in the training sequence underwent weekly probes except for the first pair.

RESA probes. Once a training pair reached the performance criterion, RESA probes followed. The recently mastered training pair would enter retention probes. All steps in the training sequence including the recently mastered steps would undergo endurance and stability probes, and the application combinations underwent application probes. Retention, stability, and application probes involved two 15-s timings per probe session. The timing length for retention, stability, and application probes remained the same as the training timing length to avoid confounding the duration of timing with the variables specific to each probe condition. Retention probes occurred after 2 weeks and 4 weeks with no practice following mastery. Endurance probes involved two 30-s timings per probe session. The timing length for endurance probes was double the training length so that there was a clear difference between the duration in training conditions and the duration in endurance probes so that growth in endurance could be more clearly evaluated. During stability probes, participants would listen to a tap dance choreography that was unlike the beat of the step they were performing transmitted through wireless headphones, and would face away from the mirror to simulate a performance environment (distracting conditions). During application probes, participants performed the application combinations. There was no programmed feedback of any kind for any of the RESA probes.

Interobserver Agreement and Treatment Fidelity

The experimenter conducted sessions and collected the error data that the software did not collect. The instructor video-recorded all sessions for the purpose of assessing interobserver agreement (IOA). Trained secondary observers viewed videos of the sessions and collected the data independently of the instructor. The study collected IOA for 25% of all sessions per step, per participant. Researchers calculated IOA by dividing the number of exact agreements for each timing by the total number of agreements plus disagreements and multiplying by 100%. The average IOA for each step is as follows: toe taps, 88.8% (range 83.3%–91.7%); heel taps, 84.7% (range 72%–93.3%); tip steps, 97.5% (range 90%–100%); dig

steps, 84.2% (range 63.3%–93.3%); shuffles right, 67.1% (range 36.7%–61.3%); shuffles left, 59.1% (range 36.7%–80%); shuffle steps, 71.7% (range 61.7%–76.7%); windshield wipers, 80.4% (range 68.3%–86.7%); Exercise 6, 70% (range 33.3%–100%); and Exercise 8, 52.7% (range 30%–63.3%).

The high degree of variability in IOA for errors is a limitation of this study. It is virtually impossible, however, to manually count responses that occur at extremely high rates and do not leave a permanent product (e.g., the 6th log of the SCC, which ranges from 100 to 1,000 per minute). Future research could target methods for improving the measurement of high-rate behavior. Advances in technology may remove this limitation in future studies.

Checklists that included the steps required of the instructor for each type of session (baseline, probe, frequency building) were developed to assess treatment fidelity. Trained observers completed the checklists for 25% of all sessions. On the checklists, “Y” indicated a correct procedural step, “N” indicated an incorrect step, and “N/A” indicated the step was not applicable during the session. Dividing the total number of Y selections by the total number of Y plus N selections and multiplying by 100% yielded a fidelity score for each session. On average, the instructor demonstrated 99.5% fidelity (range 82.2%–100%) regarding the implementation of study procedures.

Results

All experimental participants, except for Berry, met inclusion criteria and underwent training for all tap pairs. Berry did not meet the inclusion criteria for the first pair. His frequencies at baseline for these steps were nearly at aim and highly accurate. Therefore, Berry’s training began with the second pair in the training sequence. Each participant generated over 20 charts including training and probe data, an amount too expansive to include in its entirety. We have provided a sample of Tina’s data to demonstrate the design. However, we have included all data generated by weekly probes on untrained steps and RESA probes.

Sample Training Data

Figure 2 shows a portion of Tina’s training data. The panel shows data for four of the eight training steps. Note that each step shown in this panel went through the training sequence with its pair. For example, training for toe taps occurred simultaneously with heel taps, tip steps with dig steps, shuffles right with shuffles left, and shuffle steps with windshield wipers.

Tina met the inclusion criteria for training for all steps in the training sequence. She had decelerations in the frequency of corrects for heel taps (/1.51), dig steps (/1.31), and shuffles

left (/1.27). For windshield wipers, her acceleration of corrects was less than $x1.2$ ($x1.13$). Her frequencies were also below 60% of the performance criterion. Heel taps entered frequency building and, after an initial acceleration, met criteria for intervention. The instructor reduced the timing length to 10 s. Once Tina met the performance criterion, the intervention was removed, and heel taps returned to 15-s timings. Retention probes showed frequency of corrects above the performance criterion with no errors.

Dig steps entered weekly probes at the same time as heel taps entered frequency building. Weekly probes for dig steps showed an acceleration of corrects ($x1.13$) and a deceleration of errors (/1.39). Once dig steps entered frequency building, Tina met the performance criterion within eight sessions without interventions. Retention probes showed maintenance of frequencies above the performance criterion for this step as well.

Shuffles left also showed improvements in acceleration of corrects during weekly probes ($x1.18$) but simultaneously showed increases in errors ($x1.13$). To avoid shaping errors, the instructor set accuracy-only goals at the beginning of the frequency-building phase, which produced a deceleration of /1.49 in errors. However, Tina experienced several more interventions before she met the performance criterion for shuffles left. Shuffles left took notably longer to reach the performance criterion than previous steps. This pattern was common across all experimental participants for both shuffles left and right. All participants also met the criteria for interventions for both steps.

The data for weekly probes for windshield wipers show an initial level change from baseline to weekly probes. The frequency multiplier from the last data point of baseline to the first data point of weekly probes equals $x1.66$. After this initial jump up, the acceleration of corrects hovered close to the performance criterion, and errors remained low. To demonstrate the design, we have displayed the weekly probes on the daily per-minute chart here. However, during the experiment, the instructor monitored these on weekly per-minute charts. The next section includes all data for weekly probes across participants.

Weekly Probes

Figure 3 shows participants’ acceleration collections for weekly probes, and Table 3 shows acceleration values. Daisy, Andy, and Tina had the greatest accelerations of frequency of corrects for the second pair in the training sequence (tip steps and dig steps). Additionally, Tina had a substantial deceleration in errors for dig steps (/3). Because Berry started training with tip steps and dig steps, there were no weekly probes for these steps. Daisy and Berry showed very flat accelerations for the third pair (shuffles right and shuffles left), and Daisy had an acceleration of $x1.52$ for errors on shuffles left. Andy only had

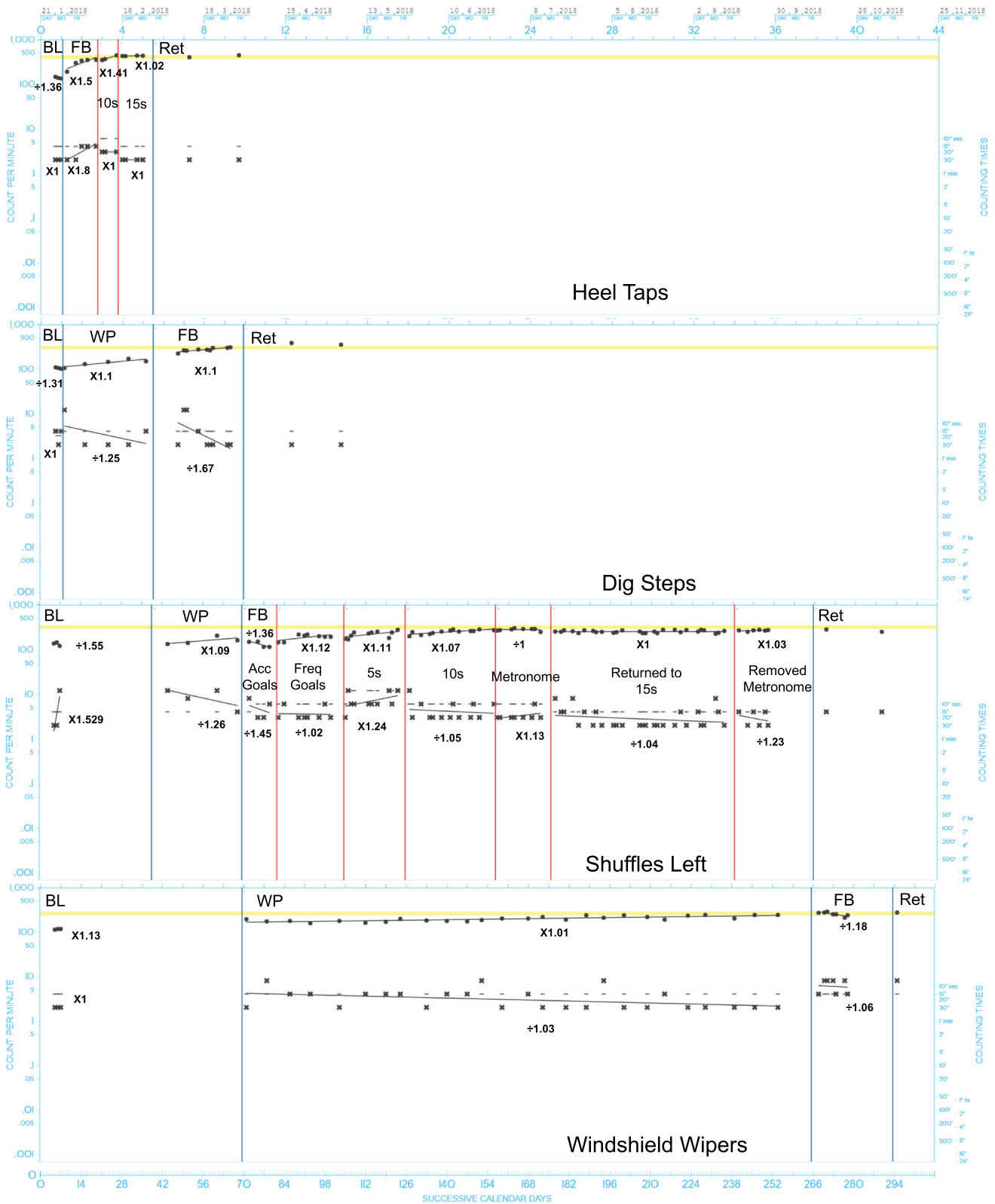
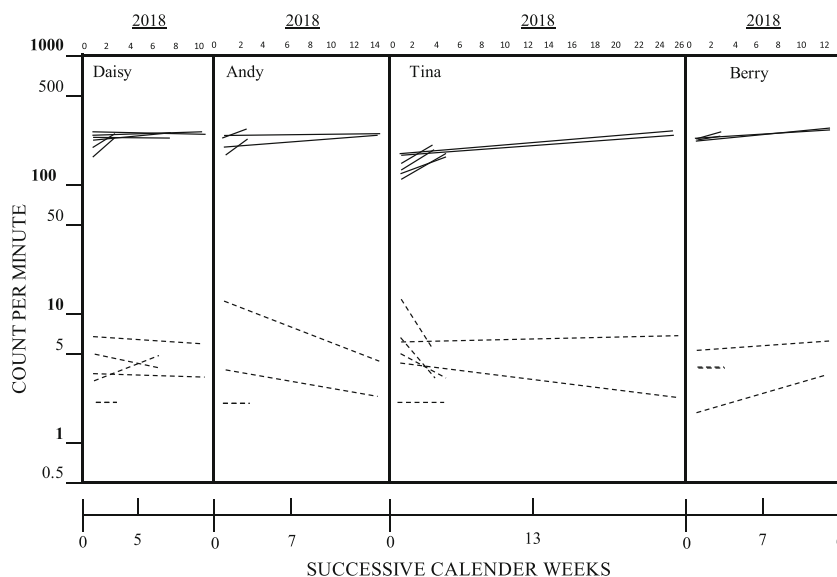


Fig. 2 Sample of Tina’s training data. The daily per-minute charts represent Tina’s training data for one half of the training pairs. Dots represent frequencies of correct responses, and x’s represent frequencies of errors. The yellow bands across the charts represent the performance

criterion. Horizontal black lines represent celerations. Vertical blue lines represent phase changes, and vertical red lines represent intervention changes

Fig. 3 Weekly probes celeration collections. Solid lines indicate the celeration of correct steps, and dashed lines indicate the celeration of errors



two probe data points for the third pair, which was not enough to generate celeration values.

Tina, on the other hand, demonstrated significant improvement on the third pair. Her shuffles right and shuffles left both had accelerations of $\times 1.61$ for frequency of corrects and $/5.21$ and $/4.26$, respectively, for frequency of errors. Andy showed some gains in accuracy for the fourth pair, with decelerations in errors for shuffle steps ($/1.49$) and windshield wipers ($/1.22$). Otherwise, there were no other major improvements across celerations for the rest of the participants. However, it is important to note that participants took much longer to master shuffles right and shuffles left in comparison to the previous steps, which extended the amount of time shuffle steps and windshield wipers were in the weekly probe phase, therefore naturally flattening celeration. Some participants did have jump-ups in frequency from baseline sessions to weekly probes for these steps. Tina’s windshield wipers (as seen previously) were one example of this. She had a $\times 1.66$ frequency multiplier from the last baseline data point to the first probe

data point. Andy and Daisy both had frequency multipliers of $/4$ for errors on shuffle steps.

Retention

Figure 4 presents the percentage of change in frequency from the last training data point to the first retention data point for the eight training steps for each participant. Data points falling below the 0 line indicate a decrease in frequency, whereas data points above the 0 line indicate an increase in frequency during the first retention probe. The smaller the difference, the more stability at retention for that step.

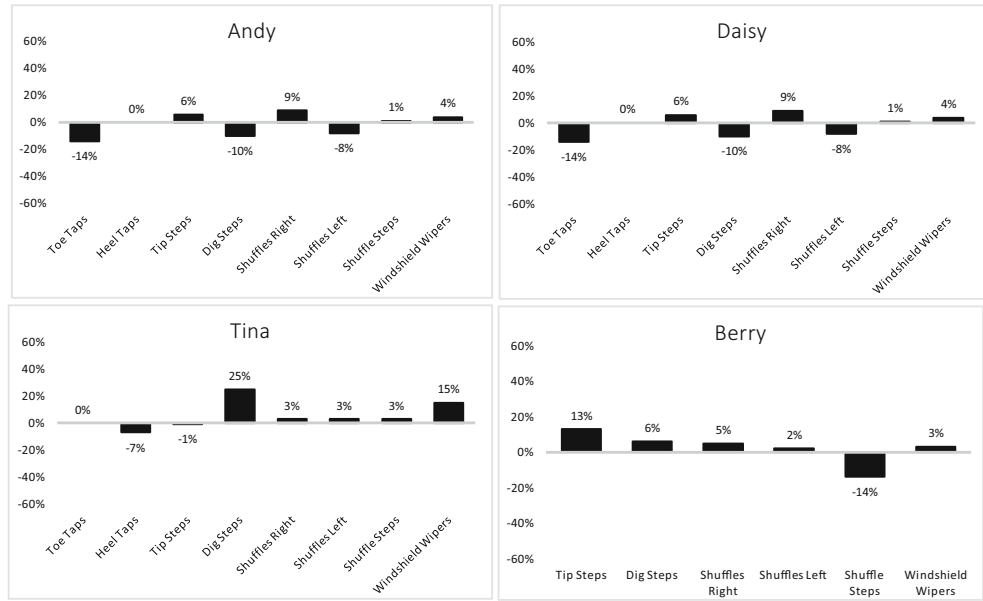
All four participants demonstrated high levels of retention across trained steps. Daisy, Tina, and Berry exhibited the most stable patterns of retention with the most steps falling above the 0 line, indicating that the change in frequency was an increase from training rather than a decrease. Though Andy and Daisy had the most steps with a loss in frequency at retention, these were all below a 15% loss. In fact, no step across all four participants suffered more than a 14% loss in

Table 3 Celeration Values for Weekly Probes

Response	Daisy		Andy		Tina		Berry	
	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.
Tip steps	$\times 1.81$	$\times 1$	$\times 1.72$	$\times 1$	$\times 1.3$	$/1$	—	—
Dig steps	$\times 1.63$	$\times 1$	$\times 2.05$	$\times 1$	$\times 1.63$	$/3$	—	—
Shuffles right	$\times 1.05$	$/1.19$	—	—	$\times 1.61$	$/5.21$	$\times 1.05$	$\times 1.05$
Shuffles left	$/1$	$\times 1.52$	—	—	$\times 1.61$	$/4.26$	$\times 1.09$	$\times 1$
Shuffle steps	$\times 1.02$	$/1.04$	$\times 1.06$	$/1.49$	$\times 1.07$	$\times 1.06$	$\times 1.12$	$\times 1.05$
Windshield wipers	$/1.03$	$/1.02$	$\times 1.01$	$/1.22$	$\times 1.07$	$/1.14$	$\times 1.12$	$\times 1.44$

Note. Andy only had two probes for shuffles right and left, which were not enough to generate celeration values. Berry does not have celeration values for tip steps and dig steps because he started training with this pair

Fig. 4 The percentage of change from each participant’s last frequency in training to the first retention probe 2 weeks after termination of training. Data above 0 indicate a percentage increase in frequency, whereas data below 0 indicate a percentage decrease



frequency from the last training frequency to first retention probe frequency.

Stability

Figure 5 shows celeration collections for stability probes for all eight training steps across participants, and celeration values can be found in Table 4. All participants showed an acceleration

in the frequency of corrects across all stability probes. With few exceptions, most also experienced decelerations in errors or maintained stability in low rates of errors across steps.

Daisy had decelerations in errors across all steps except for dig steps (x1.49) and windshield wipers (x1.98), and she maintained low errors on toe taps (x1). Andy also maintained low errors across toe taps, heel taps, and tip steps and showed decelerations for dig steps, shuffles left, and windshield wipers.

Fig. 5 Stability probes celeration collections. Solid lines indicate the celeration of correct steps, and dashed lines indicate the celeration of errors

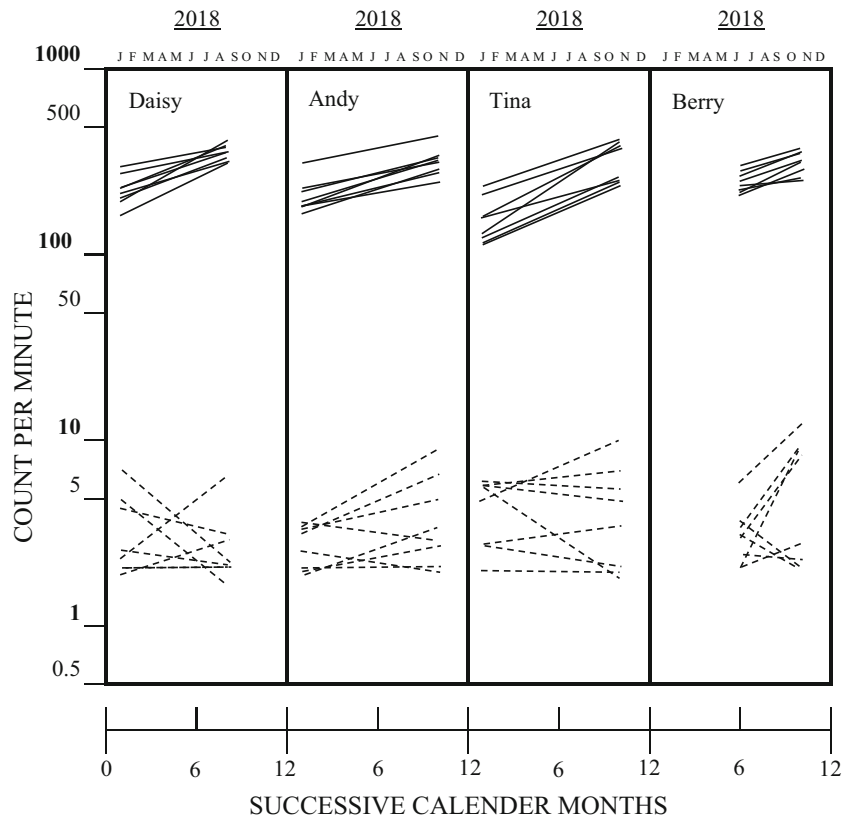


Table 4 Celeration Values for Stability Probes

Response	Daisy		Andy		Tina		Berry	
	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.
Toe taps	x1.25	x1	x1.17	x1	x1.23	/2.1	x1.32	x1
Heel taps	x1.24	/1.22	x1.18	x1	x1.47	x1	x1.44	/2.08
Tip steps	x1.8	/1	x1.52	x1	x1.72	/1.75	x1.48	/3.13
Dig steps	x1.66	x1.49	x1.36	/1.17	x1.86	/2.56	x1.53	/2.08
Shuffles right	x1.51	/2.43	x1.34	x2.3	x1.69	/1.01	x1.2	x9.79
Shuffles left	x1.69	/2.43	x1.39	/1.48	x1.54	x1.22	x1.35	x13.7
Shuffle steps	x1.57	/1.31	x1.35	x1.2	x1.67	/1.2	x1.8	x3.17
Windshield wipers	x1.37	x1.98	x1.4	/1.61	x1.55	x1.06	x1.79	x1.17

Tina had the steepest accelerations for frequency of corrects overall. Six out of eight steps had accelerations higher than x1.5. However, it is important to point out that her baseline frequencies were substantially lower than those of the other participants, giving her more room to grow toward the aims. Berry most notably had accelerations in errors on various steps (shuffles right, shuffles left, shuffle steps, windshield wipers). These steps had consistent errors throughout Berry’s training as well. Nonetheless, Berry demonstrated decelerations in errors in three steps (heel taps, tip steps, dig steps) and maintained a celeration of x1 for toe taps during stability probes.

Endurance

Figure 6 displays the celeration collections for endurance probes, and Table 5 displays the associated values. All four participants showed improvements in endurance probes. Daisy and Andy showed a very consistent acceleration in the frequency of corrects. Daisy also showed decelerations in the frequency of errors across all steps except two (tip steps and shuffles left). Andy showed acceleration in the frequency of errors across three steps: heel taps (x3.97), shuffles right (x1.26), and shuffles left (x1.27). Tina and Berry started with higher rates of errors during endurance probes. For various

Fig. 6 Endurance probes celeration collections. Solid lines indicate the celeration of correct steps, and dashed lines indicate the celeration of errors

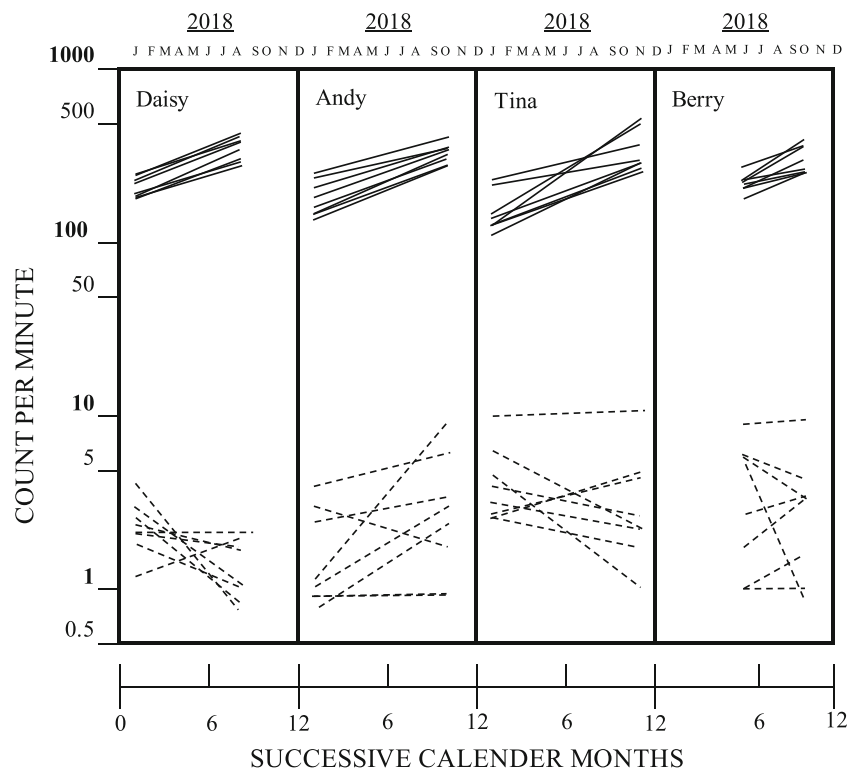


Table 5 Celeration Values for Endurance Probes

Response	Daisy		Andy		Tina		Berry	
	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.
Toe taps	x1.61	/2.43	x1.63	/1.61	x1.93	/1.68	x1.43	x1
Heel taps	x1.52	/1.22	x1.37	x3.97	x1.39	x1.62	x1.63	x1
Tip steps	x1.61	x1.36	x1.53	x1	x1.93	/1.68	x1.82	x1.51
Dig steps	x1.49	/1.58	x1.41	x1	x2.19	/2.51	x2.29	/39.6
Shuffles right	x1.96	/2.2	x1.63	/1.61	x1.53	x1.16	x1.13	/1.08
Shuffles left	x1.62	x1.01	x1.6	x1.26	x1.55	/1.48	x1.19	/3.44
Shuffle steps	x1.44	/4.05	x1.51	/1.19	x1.74	x1.31	x1.71	/1.35
Windshield wipers	x1.44	/1.16	x1.53	x1.27	x1.58	x1.29	x1.61	/1.1

steps, both participants showed a deceleration in errors and an acceleration in the frequency of corrects. Specifically, Tina showed this pattern for toe taps, tip steps, dig steps, and shuffles left. Berry showed this pattern for dig steps, shuffles right, shuffles left, shuffle steps, and windshield wipers. Berry's celeration of errors was x1 for toe taps and heel taps because he never made any errors during these probes.

Application

Figure 4 contains celeration collections for application probes, and Table 6 contains the values. Except for one exercise for Tina, all other application probes for the four participants showed improvements. Daisy demonstrated the steepest accelerations in the frequency of correct steps and the steepest decelerations in the frequency of incorrect steps. Andy also showed acceleration of frequency of corrects and deceleration of frequency of errors, though they were not as steep as those of the other participants. Tina had the shallowest accelerations in corrects (x1.25 for Exercise 6 and x1.08 for Exercise 8) and acceleration of errors for Exercise 6 (x1.48) yet she showed sizable decelerations in errors for Exercise 8 (/2.25). Berry also did not demonstrate robust accelerations in the frequency of corrects (x1.33 for Exercise 6 and x1.12 for Exercise 8); however, his starting frequencies were higher than those of the other participants. He did show the greatest decelerations

in errors, nonetheless (/31.7 for Exercise 6 and /4.38 for Exercise 8).

Constant-Series Control

Jack served as the constant-series control for application probes. Researchers yoked his schedule of measurement to Berry's schedule of probes because they were of a similar age, had similar histories of sports and music exposure, and no experience with dance instruction. Figure 7 shows celeration collections for Jack's constant-series probes beside those of the experimental participants, and Table 6 contains additional measures.

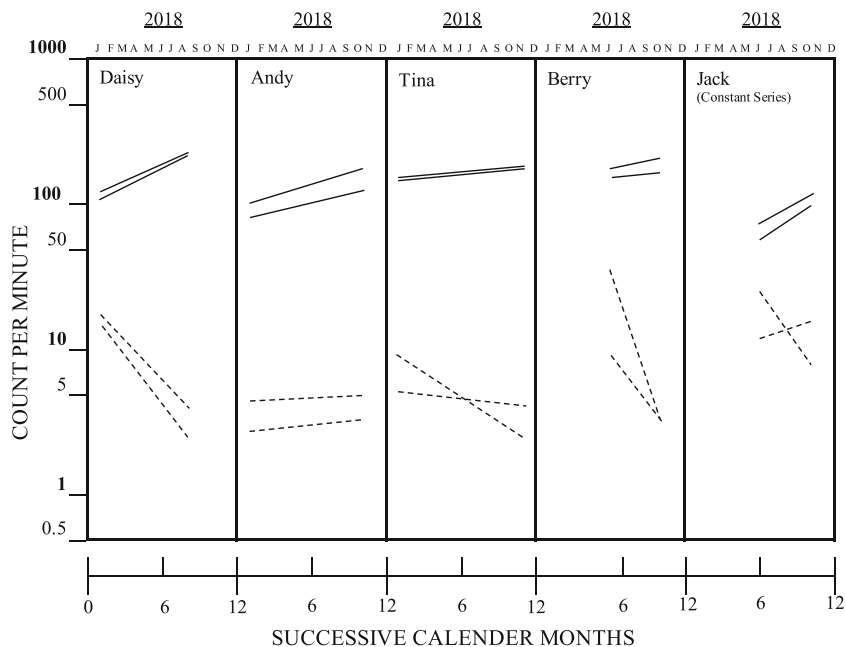
Jack showed improvements in both probes. Exercise 6 had an acceleration of x1.95 for corrects and a deceleration of x1.52 for errors, whereas Exercise 8 had an acceleration of x2.01 for corrects and a deceleration of /5.48 for errors. For both combinations, Jack showed greater accelerations in the frequency of corrects than Berry. He also showed greater deceleration of errors for Exercise 8 than Berry. Note that Berry may have experienced ceiling effects because his baseline frequencies for these exercises were in a higher cycle closer to the performance criterion, making the comparison between Jack's and Berry's celeration values difficult.

Because Tina and Jack had similar baseline frequencies, researchers compared their celerations to provide additional insight. Jack had a higher acceleration of frequency of corrects than Tina on both combinations. Both Tina and Jack had

Table 6 Celeration Values for Application Probes

Response	Daisy		Andy		Tina		Berry		Jack	
	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.	Accel.	Decel.
Exercise 6	x1.5	/4.08	x1.4	/1.26	x1.25	x1.48	x1.33	/31.7	x1.95	x1.52
Exercise 8	x1.76	/3.04	x1.32	/1.11	x1.08	/2.25	x1.12	/4.38	x2.01	/5.48

Fig. 7 Application probes celeration collections for each experimental participant and the probes of the constant-series participant (Jack) for the same combinations as application probes. Solid lines indicate the celeration of correct steps, and dashed lines indicate the celeration of errors



accelerations in errors for Exercise 6, but Tina had a lower celeration value than Jack (x1.48 vs. x1.52) and maintained fewer errors compared to Jack. Tina’s errors did not exceed 8 per minute, whereas Jack’s were as high as 20 per minute. Both had decelerations for Exercise 8. Jack’s deceleration of errors (/5.48) was steeper than Tina’s (/2.25). There is also a limitation to this comparison because Tina and Jack did not have the same schedule of probes, which also impacts celeration values. Tables 7 and 8 show several measures for Exercises 6 and 8 for experimental participants and Jack that can provide a basis for further comparison.

Despite his improvements, Jack never reached 100% accuracy on either combination. All experimental participants had greater accuracy ratios than Jack for both Exercises 6 and 8. With the exception of Tina on Exercise 6 and Andy on Exercise 8, all other first-last frequency multipliers for experimental participants showed greater deceleration of frequency of errors than Jack’s. These measures show that in the majority of cases, experimental participants gained greater degrees

of accuracy than Jack. Berry, Daisy, and Andy had greater improvement indices for Exercise 6 than Jack. However, he outperformed all experimental participants on this measure for Exercise 8.

Discussion

The main purpose of this study was to evaluate fluency outcomes as a function of building the frequencies of component responses with a tap dance training sequence. Weekly probes served to monitor potential improvements of untrained components. Results showed some facilitative effects during weekly probes, primarily for the second pair in the sequence (i.e., tip steps and dig steps). Tina also demonstrated significant improvements in celeration for the third pair (i.e., shuffles right and shuffles left) during weekly probes, suggesting some facilitative effects of training of the previous pair. The results of the retention probes also indicated that retention emerged as a function of training steps to optimal rates. Retention probes

Table 7 Comparison Metrics for Experimental Participants and Control Participant: Exercise 6

Participant	Celeration		First-Last Frequency Multiplier		Accuracy Ratio	Improvement Index
	Accel.	Decel.	Accel.	Decel.	Last Decel. to Last Accel.	Overall
Jack	x1.95	x1.52	x1.94	x1	x11	x1.28
Berry	x1.33	/31.7	x1.5	/36	x114	x42.1
Daisy	x1.5	/4.08	x1.79	/10	x100	x5.63
Andy	x1.4	/1.26	x2	/2	x40	x1.77
Tina	x1.25	x1.48	x1.68	x1	x16	/1.18

Table 8 Comparison Metrics for Experimental Participants and Control Participant: Exercise 8

Participant	Celeration		First-Last Frequency Multiplier		Accuracy Ratio	Improvement Index
	Accel.	Decel.	Accel.	Decel.	Last Decel. to Last Accel.	Overall
Jack	x2.01	/5.48	x1.6	/3	x12	x11
Berry	x1.12	/4.38	x1.31	/12	x102	x4.92
Daisy	x1.76	/3.04	x2.15	/4	x58	x5.35
Andy	x1.32	/1.11	x1.57	/2	x66	x1.47
Tina	x1.08	/2.25	x1.35	/12	x46	x2.43

showed little decay in performance after 2 weeks with no practice. Stability and endurance also seemed to have emerged for most steps across participants. General trends for stability and endurance probes showed accelerations in the frequency of corrects and decelerations of errors. Application is more difficult to interpret given the results of the control participant (i.e., Jack). Though the experimental participants all demonstrated gains in application probes, so did Jack. In most instances, he outperformed experimental participants in terms of celeration. Experimental participants, however, did outperform Jack in terms of accuracy. Anecdotally, the experimental participants also showed unmeasured improvements in the aesthetic features of the application combinations, including clarity of sound, volume of sounds, range of motion of the feet, and posture, which are important in the performing arts. Notwithstanding, there were several differences between the control participant and the experimental participants that may have presented confounds, making comparisons difficult.

There were several uncontrolled variables in the constant-series condition that impacted celeration values and presented inequities in the comparison. First, the control participant's frequencies resided in a lower cycle of the SCC than those of his experimental participant counterpart. Due to the SCC's logarithmic nature, celeration values decrease in higher cycles even if there is an equal amount of absolute change. This makes it difficult to compare growth across participants if their frequencies are not in the same cycle. Second, the control participant experienced probes within a shorter period than did all the experimental participants except his experimental counterpart, who shared his probe schedule. This also impacts celeration values on the SCC and, again, does not allow for a fair comparison with the experimental participants who did not share the control participant's probe schedule. Future studies can improve on these limitations by setting fixed probe schedules (i.e., monthly, biweekly) and ensuring that control participants and comparison experimental participants have similar baseline frequencies.

A secondary purpose of this study was to provide an example of how one can apply the PT framework to constructing, monitoring, and perfecting instructional sequences for a variety of repertoires, including those of interest in sports.

This study was a first attempt at building a training sequence for tap dancers based on a component-composite analysis. Though the study did not produce an ideal training sequence, it did yield important discoveries regarding fluency outcomes. It also provides an example of how a PT approach to training can reveal inadequacies in the instructional sequence. For example, participants took substantially longer to reach the performance criterion for one of the pairs, which highlights where a change should be made. Perhaps breaking down these steps into smaller components would have helped participants master them faster. This may have also facilitated improvements in untrained components before they reached the frequency-building phase. These are the types of speculations that are prompted when precision teachers are continuously engaged with the data of their learners. The process of pinpoint, record, change, and try again leads precision teachers to discoveries regarding the effectiveness and efficiency of their instructional arrangements. The main contribution of PT to sports and any other area is a process by which one can systematically and continuously evaluate and improve training using a sensitive dimensional unit of measurement (rate of response) and a standard real-time display (SCC) that allows for effective data-based decision making in service of learners.

Limitations aside, the present study also contributes to the literature on PT and fluency instruction in several ways. First, it adds to the small yet growing body of evidence regarding sports performance in general and dance instruction in particular. Second, it adds to the PT literature on psychomotor learning, which is limited not only in size but also in the types of behaviors examined until now. It is also among a handful of studies that have examined a repertoire of behavior that resides in the sixth cycle of the SCC. Most of the published research in PT has been with behaviors that reside within the fourth and fifth cycles, which equate to approximately 1 to 100 responses per minute. There is comparatively less research on responses such as tap dancing that can occur above 100 responses per minute or below 1 response per minute. Finally, the use of a software program to count dance steps is the first of its kind. With refinement, this tool may be useful as an electronic application for adding objective, rate-based feedback for tap dancers and instructors alike.

We encourage future research to focus on exploring the basic behaviors that make up complex gross and fine motor movements and the order of training that yields the most efficient learning and maintenance, both in sports and in other motor movement domains. PT provides a model for instructional design that focuses on building behavioral repertoires fluently from the ground up with an emphasis on continuous progress monitoring and data-based adjustments (Johnson & Street, 2013). Dance and sports training, in general, can benefit from approaching behavioral intervention from the instructional design perspective of PT. This is a prime opportunity to investigate whether PT and fluency-based instruction can impact sports performance as it has education. Though this study and the few that came before it are a start, there is much work to be done before we understand what fluency means when it comes to motor behavior and how we can capitalize on this knowledge.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee of the University of Nevada Institutional Review Board (Project 884209-1) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

- Allison, M. G., & Ayllon, T. (1980). Behavioral coaching in the development of skills in football, gymnastics, and tennis. *Journal of Applied Behavior Analysis*, 13(2), 279–314.
- Anderson, G., & Kirkpatrick, M. A. (2002). Variable effects of a behavioral treatment package on the performance of inline roller speed skaters. *Journal of Applied Behavior Analysis*, 35(2), 195–198.
- Berens, K., Boyce, T. E., Berens, N. M., Doney, J. K., & Kenzer, A. L. (2003). A technology for evaluating relations between response frequency and academic performance outcomes. *Journal of Precision Teaching and Celeration*, 19(1), 20–34.
- Berens, N. M., & Hayes, S. C. (2007). Arbitrarily applicable comparative relations: Experimental evidence for a relational operant. *Journal of Applied Behavior Analysis*, 40(1), 45–71.
- Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst*, 19(2), 163–197.
- Binder, C., Haughton, E., & Bateman, B. (2002). Fluency: Achieving true mastery in the learning process. Unpublished manuscript. *Professional Papers in Special Education*, 2–20. Retrieved from: <https://www.fluency.org/unpublished>.
- Binder, C., & Sweeney, L. (2002). Building fluent performance in a customer call center. *Performance Improvement*, 41(2), 29–37.
- Boyer, E., Miltenberger, R. G., Batsche, C., & Fogel, V. (2009). Video modeling by experts with video feedback to enhance gymnastics skills. *Journal of Applied Behavior Analysis*, 42(4), 855–860.
- Brady, K. K., & Kubina, R. M. (2010). Endurance of multiplication fact fluency for students with attention deficit hyperactivity disorder. *Behavior Modification*, 34(2), 79–93.
- Brobst, B., & Ward, P. (2002). Effects of public posting, goal setting, and oral feedback on the skills of female soccer players. *Journal of Applied Behavior Analysis*, 35(3), 247–257.
- Brosnan, J., Moeyaert, M., Newsome, K., & Healy, O. (2018). Multilevel analysis of multiple-baseline data evaluating precision teaching as an intervention for improving fluency in foundational reading skills for at risk readers. *Exceptionality*, 26(3), 137–161.
- Bucklin, B. R., Dickinson, A. M., & Brethower, D. M. (2000). A comparison of the effects of fluency training and accuracy training on application and retention. *Performance Improvement Quarterly*, 13(3), 140–163.
- Cavallini, F., & Perini, S. (2009). Comparison of teaching syllables or words on reading rate. *European Journal of Behavior Analysis*, 10(2), 255–263.
- CentralReach, LLC. (2019). CR PrecisionX [Online software]. Retrieved on May 3, 2020, from <https://app.chartlytics.com/login>
- Chapman, S. S., Ewing, C. B., & Mozzoni, M. P. (2005). Precision teaching and fluency training across cognitive, physical, and academic tasks in children with traumatic brain injury: A multiple baseline study. *Behavioral Interventions*, 20(1), 37–49.
- Chiesa, M., & Robertson, A. (2000). Precision teaching and fluency training: Making maths easier for pupils and teachers. *Educational Psychology in Practice*, 16(3), 297–310.
- Dembek, G. A., & Kubina, R. M. (2018). Talk aloud problem solving and frequency building to a performance criterion improves science reasoning. *Journal of Evidenced-Based Practices for Schools*, 16, 147–170.
- Fabrizio, M. A., Schirmer, K., King, A., Diakite, A., & Stovel, L. (2007). Precision teaching a foundational motor skill to a child with autism. *Journal of Precision Teaching and Celeration*, 23, 16–18.
- Fitterling, J. M., & Ayllon, T. (1983). Behavioral coaching in classical ballet: Enhancing skill development. *Behavior Modification*, 7(3), 345–368.
- Griffin, C. P., & Murtagh, L. (2015). Increasing the sight vocabulary and reading fluency of children requiring reading support: The use of a precision teaching approach. *Educational Psychology in Practice*, 31(2), 186–209.
- Harding, J. W., Wacker, D. P., Berg, W. K., Rick, G., & Lee, J. F. (2004). Promoting response variability and stimulus generalization in martial arts training. *Journal of Applied Behavior Analysis*, 37(2), 185–195.
- Haughton, E. (1972). Aims: Growing and sharing. In J. Jordan & L. Robbins (Eds.), *Let's try doing something else kind of thing: Behavioral principles and the exceptional child* (pp. 20–39). The Council for Exceptional Children. Arlington, VA.
- Haughton, E. C. (1980). Practicing practices: Learning by activity. *Journal of Precision Teaching*, 1(3), 3–20.
- Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). *The scientist practitioner: Research and accountability in the age of managed care* (2nd). Allyn and Bacon. Boston, MA.
- Horner, D. R., & Baer, D. M. (1978). Multiple-probe technique: A variation of the multiple baseline. *Journal of Applied Behavior Analysis*, 11(1), 189–196.
- Carl Hughes, J., Beverley, M., & Whitehead, J. (2007). Using precision teaching to increase the fluency of word reading with problem readers. *European Journal of Behavior Analysis*, 8(2), 221–238.
- Hume, K. M., & Crossman, J. (1992). Musical reinforcement of practice behaviors among competitive swimmers. *Journal of Applied Behavior Analysis*, 25(3), 665–670.

- Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, 47(11), 1475–1490.
- Johnson, K. R., & Layng, T. V. J. (1996). On terms and procedures: Fluency. *The Behavior Analyst*, 19(2), 281–288.
- Johnson, K. R., & Street, E. M. (2013). *Response to intervention and precision teaching*. The Guilford Press.
- Kladopoulos, C. N., & McComas, J. J. (2001). The effects of form training on foul-shooting performance in members of a women's college basketball team. *Journal of Applied Behavior Analysis*, 34(3), 329–332.
- Komaki, J., & Barnett, F. (1977). A behavioral approach to coaching football: Improving the play execution of the offensive backfield on a youth football team. *Journal of Applied Behavior Analysis*, 10(4), 657–664.
- Koop, S., & Martin, G. L. (1983). Evaluation of a coaching strategy to reduce swimming stroke errors with beginning age-group swimmers. *Journal of Applied Behavior Analysis*, 16(4), 447–460.
- Kostewicz, D. E., & Kubina, R. M. J. (2010). A comparison of two reading fluency methods: Repeated readings to a fluency criterion and interval sprinting. *Reading Improvement*, 47(1), 43–63.
- Kubina, R. M. (2005). The relations among fluency, rate building, and practice: A response to Doughty, Chase, and O'Shields (2004). *The Behavior Analyst*, 28(1), 73–76.
- Kubina, R. M., & Yurich, K. K. L. (2012). *The precision teaching book* (1st). Greatness Achieved Publishing. Lemont, PA.
- Lambe, D., Murphy, C., & Kelly, M. E. (2015). The impact of a precision teaching intervention on the reading fluency of typically developing children. *Behavioral Interventions*, 30, 364–377.
- Lindsley, O. R. (1992a). Precision teaching: Discoveries and effects. *Journal of Applied Behavior Analysis*, 25(1), 51–57.
- Lindsley, O. R. (1992b). *Skinner on measurement*. Behavior Research Company. Kansas City, KS.
- Lokke, G. E., Lokke, J. A., & Arntzen, E. (2008). Precision teaching, frequency-building, and ballet dancing. *Journal of Precision Teaching and Celeration*, 24, 21–27.
- Luiselli, J. K., Woods, K. E., & Reed, D. D. (2011). Review of sport performance research with youth, collegiate, and elite athletes. *Journal of Applied Behavior Analysis*, 44(4), 999–1002.
- McDowell, C., & Keenan, M. (2001). Cumulative dysfluency: Still evident in our classrooms, despite what we know. *Journal of Precision Teaching and Celeration*, 17(2), 1–6.
- McDowell, C., McIntyre, C., Bones, R., & Keenan, M. (2002). Teaching component skills to improve golf swing. *Journal of Precision Teaching and Celeration*, 18(2), 61–66.
- McTieman, A., Holloway, J., Healy, O., & Hogan, M. (2016). A randomized controlled trial of Morningside math facts curriculum on fluency, stability, endurance, and application outcomes. *Journal of Behavioral Education*, 25(1), 49–68.
- Mellalieu, S. D., Hanton, S., & O'Brien, M. (2006). The effects of goal setting on rugby performance. *Journal of Applied Behavior Analysis*, 39(2), 257–261.
- Mercer, C. D., Campbell, K. U., Miller, D. M., Mercer, K. D., & Lane, H. B. (2000). Effects of a reading fluency intervention for middle schoolers with specific learning disabilities. *Learning Disabilities Research & Practice*, 15(4), 179–189.
- Milyko, K., Berens, K., & Ghezzi, P. (2012). An investigation of rapid automatic naming as a generalized operant. *Journal of Precision Teaching and Celeration*, 28, 3–16.
- Newsome, K., Berens, K., Ghezzi, P. M., Aninao, T., & Newsome, W. D. (2014). Training relational language to improve reading comprehension. *European Journal of Behavior Analysis*, 15(2), 165–197.
- Pennypacker, H., Gutierrez, A. J., & Lindsley, O. R. (2003). *Handbook of the standard celeration chart*. Cambridge Center for Behavioral Studies. Concord, MA.
- Quinn, M., Miltenberger, R., Abreu, A., & Narozanick, T. (2017). An intervention featuring public posting and graphical feedback to enhance the performance of competitive dancers. *Behavior Analysis in Practice*, 10(1), 1–11.
- Quinn, M. J., Miltenberger, R. G., & Fogel, V. A. (2015). Using TagTeach to improve the proficiency of dance movements. *Journal of Applied Behavior Analysis*, 48(1), 11–24.
- Quinn, M. J., Sherman, J. A., Sheldon, J. B., Quinn, L. M., & Harchik, A. E. (1992). Social validation of component behaviors of following instructions, accepting criticism, and negotiating. *Journal of Applied Behavior Analysis*, 25(2), 401–413.
- Rogers, L., Hemmeter, M. L., & Wolery, M. (2010). Using a constant time delay procedure to teach foundational swimming skills to children with autism. *Topics in Early Childhood Special Education*, 30(2), 102–111.
- Schenk, M., & Miltenberger, R. (2019). A review of behavioral interventions to enhance sports performance. *Behavioral Interventions*, 34(2), 248–279.
- Schonwetter, S. W., Miltenberger, R., & Oliver, J. R. (2014). An evaluation of self-monitoring to improve swimming performance. *Behavioral Interventions*, 29, 213–224.
- Scott, D., & Scott, L. M. (1997). A performance improvement program for an international-level track and field athlete. *Journal of Applied Behavior Analysis*, 30(3), 573–575.
- Sleeman, M., Friesen, M., Tyler-Merrick, G., & Walker, L. (2019). The effects of precision teaching and self-regulated learning on early multiplication fluency. *Journal of Behavioral Education*. 1–29. <https://doi.org/10.1007/s10864-019-09360-7>.
- Smith, S. L., & Ward, P. (2006). Behavioral interventions to improve performance in collegiate football. *Journal of Applied Behavior Analysis*, 39(3), 385–391.
- Stocker, J. D., Schwartz, R., Kubina, R. M., Kostewicz, D., & Kozloff, M. (2019). Behavioral fluency and mathematics intervention research: A review of the last 20 years. *Behavioral Interventions*, 34(1), 102–117.
- Stokes, J. V., Luiselli, J. K., Reed, D. D., & Fleming, R. K. (2010). Behavioral coaching to improve offensive line pass-blocking skills of high school football athletes. *Journal of Applied Behavior Analysis*, 43(3), 463–472.
- Tai, S. S. M., & Miltenberger, R. G. (2017). Evaluating behavioral skills training to teach safe tackling skills to youth football players. *Journal of Applied Behavior Analysis*, 50(4), 849–855.
- Twarek, M., Cihon, T., & Eshleman, J. (2010). The effects of fluent levels of Big 6+6 skill elements on functional motor skills with children with autism. *Behavioral Interventions*, 25, 275–293.
- Vintere, P., & Poulson, C. L. (2010). The effects of a behavioral movement-training package on dance performance. *European Journal of Behavior Analysis*, 11(2), 151–166.
- Ward, P., & Carnes, M. (2002). Effects of posting self-set goals on collegiate football players' skill execution during practice and games. *Journal of Applied Behavior Analysis*, 35(1), 1–12.
- Weiss, M. J., Foley, K., Pearson, N., & Pahl, S. (2010). The importance of fluency outcomes in learners with autism. *The Behavior Analyst Today*, 11(4), 245–253.
- White, O. R. 1984. *Aim star wars: Setting aims that compete*. *Journal of Precision Teaching*, 5(3), 55–63.
- Ziegler, S. G. (1987). Effects of stimulus cueing on the acquisition of groundstrokes by beginning tennis players. *Journal of Applied Behavior Analysis*, 20(4), 405–411.
- Ziegler, S. G. (1994). The effects of attentional shift training on the execution of soccer skills: A preliminary investigation. *Journal of Applied Behavior Analysis*, 27(3), 545–552.