



A Precision Teaching Framework for Improving Mathematical Skills of Students with Intellectual and Developmental Disabilities

Athanasios Vostanis¹ · Ciara Padden¹ · Mecca Chiesa¹ · Konstantinos Rizos² · Peter E. Langdon³

© The Author(s) 2020

Abstract

The aim of this study was to investigate the effect of a precision teaching (PT) framework on the mathematical ability of students with intellectual and developmental disabilities. We also examined if students of moderate mathematical ability could perform as well as their peers with fewer difficulties with their math skills. Sixteen students participated and were divided into three groups. One group engaged in PT, and the other two groups functioned as comparisons. The PT group practiced six skills introduced linearly. An A-B design was used for the five component skills, and a multiple baseline across participants design was used for the composite skill (addition). The intervention led to a significant improvement in all skills, including addition, and this was associated with a large effect size; student performance met or exceeded that of their peers. Overall, the findings suggest that PT is an efficient and effective approach for teaching students with IDD.

Keywords Standard celeration chart · Autism · Mathematics · Special education · Component–composite analysis · FBPC

Introduction

Mathematics is an essential academic subject directly related to a variety of fields such as science, technology, and engineering (King et al. 2016). Along with reading and writing, mathematics completes the triad of core academic skills that are highly valued both in mainstream and special education. Despite the importance of mathematics, the performance of students with intellectual and developmental disabilities (IDDs) within this area continues to require attention. Specifically, in the United States (U.S.) in 2017, 51–69% of fourth (National Center for Educational

✉ Athanasios Vostanis
thanos.vostanis@outlook.com

Extended author information available on the last page of the article

Statistics; NCES 2017a), and eighth graders (NCES 2017b), respectively, did not meet the expected level of performance in mathematics. Similarly, 82% of students in the United Kingdom (UK) who completed Key Stage 2 and were receiving special education support, failed to meet the expected standards of performance across all three core academic skills, including mathematics (Department for Education 2017).

In response to the need for better mathematics instruction, many educators in the U.S. and the UK have turned toward evidence-based models and performance standards (Common Core Standards Initiative 2010; Department for Education 2016). While historically students with moderate or severe disabilities have accessed more functional than academic curricula (Bouck and Joshi 2015), there has been increasing emphasis on academic curricula due to ethical and legal expectations regarding students with special educational needs (Barrett et al. 1991; Department for Education 2016; U.S. Department of Education 2004). This means that students in special education who could benefit from accessing the general education curriculum should access the same evidence-based educational procedures and be evaluated in respect to the same standards as their peers in mainstream education.

To that end, various instructional models and tactics have been investigated (e.g., Copy-Cover-Compare, or Taped Problems), some of which have been shown to be effective (Grafman and Cates 2010; Kleinert et al. 2018; Kroesbergen and Van Luit 2003). Particularly regarding the teaching procedures, the following have been identified as quality indicators of successful teaching: (a) tailoring instruction to students' current skill levels, (b) providing multiple opportunities to respond on a daily basis, (c) giving immediate feedback, (d) utilizing timed practice and self-graphing, and (e) providing access to reinforcing consequences (Daly et al. 2007; Kleinert et al. 2018; Weisenburgh-Snyder et al. 2015).

One educational framework which incorporates all of these components is precision teaching (PT). Fundamentally, PT is a system for precisely defining, measuring, and analyzing behavior that can be complementary to any teaching procedure or curriculum (Kubina and Yurich 2012). However, PT has historically combined its strategies with tactics that have emerged from the field of instructional design (Johnson and Street 2012). Some notable tactics include, first, pinpointing behavior in a precise manner by combining a movement cycle and a learning channel. A movement cycle consists of an action verb and a noun (e.g., writes digit). A learning channel refers to the sensory modality in which students receive and subsequently respond to instruction, such as seeing a question and saying the answer (i.e., a See-Say learning channel set) (Haughton 1980; Lin and Kubina 2000; Lindsley 1998). A second PT tactic is the daily graphing of performance on a standard graph, the 'Standard Celeration Chart' (SCC; Calkin 2005) which students themselves are expected to use (Johnson and Street 2013). Third, component-composite analysis is an instructional tactic embedded in the framework, during which the teacher sequences the skills in a logical order from basic prerequisites (components) to more complex (composite) skills (Johnson and Street 2013). Finally, frequency building to a performance criterion (FBPC) is another instructional tactic, which involves timed practice toward a frequency aim (e.g., 70–50 correct responses per minute; Haughton 1980; Kubina and Yurich 2009) and is usually linked to fluency outcomes (Binder 1996; Johnson and Street 2012). In pursuance of clarity, we will be referring to the combination

of all these tactics as PT because it is usual for all of them to be part of the PT framework.

As PT seems to meet all the quality indicators identified in the literature, it may be a valuable approach to academic instruction for students with special educational needs. Several studies of PT have demonstrated improvements in the areas of literacy (Cavallini et al. 2010; Datchuk et al. 2015; Griffin and Murtagh 2015; Hughes, Beverley, & Whitehead, 2007), and numeracy (Chiesa & Robertson, 2000; Fitzgerald & Garcia, 2006; Greene, McTiernan, & Holloway, 2018; McTiernan, Holloway, Leonard, & Healy, 2018). However, as a recent systematic review highlighted, despite encouraging evidence, there is still a need for more research on the application of PT to students with IDD (Ramey et al., 2016).

Given the effectiveness of PT in improving academic outcomes for typically developing students, further research evaluating PT as an instructional approach for students with IDD is warranted. Therefore, the present study aimed to examine (a) whether students with IDD attending special education in the UK demonstrate improvements in mathematics, and particularly addition, following the introduction of PT; and (b) if students with IDD who have moderate mathematical ability, could equal or outperform their peers with fewer difficulties with their math skills. We first evaluated the effects of PT on five component skills within a quasi-experimental design. Next, we evaluated the effects of PT on addition skills within a multiple baseline across participants design.

Methods

Participants and Setting

Sixteen students (3 female; 13 male) participated in this study aged 7–12 years old ($M=9.8$, $SD=1.6$). Twelve had a primary diagnosis of an autism spectrum disorder (ASD), and four had a primary diagnosis of intellectual disabilities (ID; see Table 1). All participants had an Educational Health Care Plan (EHCP) in place offering information on their diagnosis, needs, and current educational provision. The language at home was Nepali for one participant (Una) and English for the rest.

The study took place at the students' school in England, which provided services for students with an EHCP aged between 3 and 19 years. The curriculum included self-help, vocational, social, and academic skills. Sessions took place in an empty 6 m × 5 m classroom, equipped with a camera, a whiteboard, two desks, three chairs, and two storage cupboards.

Inclusion Criteria

The inclusion criteria for this study were: (a) a diagnosis of intellectual and/or developmental disabilities IDD, (b) being able to complete the first 13/72 items of the Test of Early Mathematics Ability (TEMA-3; Ginsburg and Baroody, 2003), (c) being able to write numbers 0–9 and the basic mathematical symbols (e.g., +, −,

Table 1 Participants' experimental group, gender, age, primary diagnosis, and scores on the standardized assessments

Participants	Gender	Age	Primary diagnosis	VABS-II standard scores	GARS-autism index	TEMA-3 raw scores
Olaf ^a	Male	9	ASD	80	68	37
Andy	Male	9	ID	68	41	38
Tom	Male	7	ASD	97	98	43
Una	Female	8	ASD	65	66	39
John ^b	Male	7	ASD	94	81	66
Jared	Male	8	ASD	87	79	63
Connor	Male	11	ASD	102	45	62
Kyle	Male	11	ASD and ADHD	77	59	58
Emma ^c	Female	9	Variant of Jacobsen Syndrome and ID	53	41	6
Jonathan	Male	8	ASD	57	53	25
Ogden	Male	10	ASD and West Syndrome	55	68	16
Adam	Male	8	ASD	92	85	51
Cole	Male	10	ASD	72	94	25
Travis	Male	11	ID and Dyspraxia	75	59	31
Dylan	Male	12	ASD	78	55	26
Rachel	Female	11	ID	63	72	12

The first four participants comprise the PT group, the next four the WC group, and the remaining participants comprise the SC group. VABS-II TRF standard scores are presented excluding the motor domain. GARS-2 sum of standard scores and autism index are presented. TEMA-3 raw scores are presented. Weekly monitoring lasted for as long as the PT group practiced each skill

ASD, autism spectrum disorder; ADHD, attention deficit hyperactivity disorder; ID, intellectual disability

^aPrecision teaching group

^bWeekly comparison group

^cSingle-comparison group

=), (d) not exhibiting challenging behavior that could hinder engagement with the instructional procedures. Criteria (c) and (d) were determined through anecdotal observations and discussion with the teachers.

Materials

Assessment Tools

The TEMA-3 (Ginsburg & Baroody, 2003) is a 72 item measuring the general mathematical ability of an individual (Libertus, Feigenson, & Halberda, 2013). Its administration lasts approximately 40 min. It has been normed in 1219 students aged from 3 to 8 years. Its internal consistency has been reported at .90 and test-retest

reliability at .80 (Ginsburg & Baroody, 2003). This measure guided both the inclusion in the study as well as the participants' grouping.

The VABS-II TRF (Sparrow et al., 2005) is a 233 item measuring adaptive behavior. Administration lasts approximately 20 min, and standard scores are available for domain and composite scores (Sparrow et al., 2005). It has been normed on 2500 students aged 3–18 years old and can be used with individuals up to 21 years of age (Sparrow et al., 2005). Domain reliability coefficients have been reported at .90 on average, while test re-test reliability of all domains has been reported at .80 on average. Finally, the test re-test reliability of the Adaptive Behavior Composite has been reported at .91 (Sparrow et al., 2005). The VABS-II TRF provided descriptive information on participants' general abilities.

The GARS-2 is a 42-item measuring the severity of symptoms related to autism spectrum disorder (Gilliam, 2006). It has been normed on 1092 individuals with ASD aged 3–22 years old. The coefficient alpha for the four subscales has been reported at .90 on average, with test re-test reliability of the total score at .88. The GARS-2 provided descriptive information on symptoms exhibited by the participants that could be related to autism.

Classroom Materials

Ring binders were used to store participants' worksheets. A4 worksheets (using portrait format, 22 Arial font and black color) were constructed using Microsoft Office®. All worksheets presented the task (e.g., addition facts) in random order, and with more pages than the students could complete within the time provided. That way, artificial ceilings on participants' performance were avoided. Two linear graphs, a 'timings' and a 'daily graph', were also constructed to resemble the SCC. Both had blue ink and days on the x-axis but differed in terms of the axis used (i.e., equal-interval axis) and the lack of standardization. Pencils, erasers, notebooks, and digital timers were used. An A4 laminated class-shop catalog was created with pages in portrait orientation, and a 28 Times New Roman font with a picture in the middle of each page (sized 13×15 cm) showing each available item. Finally, a points board was made in portrait orientation with a Times New Roman 12 font and a 6×6 grid.

Dependent Variable

The primary skill of concern (i.e., composite skill) was single- and double-digit addition facts with a sum ≤ 20 . The skill was broken down to four smaller tasks referred to as 'slices' to optimize participants' learning. Specifically, slice 1 included all addition facts between numbers 0–3, slice 2 numbers 4–7, slice 3 numbers 8–10, and slice 4 all numbers from 0 to 10. Performance was measured by recording the correct and incorrect digits written per minute.

Research Design

A quasi-experimental (A–B) design was utilized for the five component skills, and a multiple baseline across participants design (MBL) was utilized to evaluate the effects of the intervention on the composite skill.

Participant Allocation to Experimental Conditions

TEMA-3 scores were analyzed, and participants were divided into top, moderate, and low performers before being allocated to three groups (see Table 1). Those categories reflected performance among participants in the study; not the general population. Homogeneous grouping was utilized and made possible the delivery of the same lesson to the participants receiving PT, as it was relevant for all of them. Thus, lessons were delivered on a 1:2 teacher to student ratio, while peers were receiving their regular math lesson. That way, the classroom environment was simulated to the extent possible. Participants were allocated to one of three groups based upon the assessment of their existing mathematical ability: (1) *PT group*: the first group was the ‘PT group’ ($n=4$), who were of moderate mathematical ability, relative to the participants assigned to the other two groups. The PT group received the independent variable (IV), namely daily PT practice, (2) *Weekly-comparison group*: the second group was the ‘weekly-comparison group’ (WC group; $n=4$), who were the top performers in terms of mathematical ability. The WC group did not receive the IV and were assessed once a week, but received treatment as usual (TAU) in their classroom, and (3) *Single-comparison group*: The third group was the ‘single-comparison group’ (SC group; $n=8$), who were the low performers and students who did not meet criteria for inclusion in the study. The SC group were only assessed at the beginning and end of the study and receive TAU in their classrooms. TAU consisted of 1 h of mathematics per day.

Each comparison group served different purposes: the WC group guided the frequency aim set for the PT group and allowed the monitoring of their progress across the weeks as they accessed their regular lessons. The SC group made possible the examination of variables related to internal validity such as history and maturation. Specifically, the general lack of improvement of this group confirmed that history and maturation did not affect the PT group’s performance. Una, Andy, Olaf, and Tom were chosen from two different classrooms as the PT group. Similarly, Kyle, Connor, Jared, and John were selected from two different classrooms as the WC group, while the rest of the students were included in the study as the SC group.

Table 2 All pinpointed skills, median performance of the WC group, and the performance range for the PT group

Scope and sequence	Pinpointed skills	Median performance of weekly participants	10% range created for PT participants
N/A	See-Says color name of tile on worksheet	N/A	N/A
N/A	See-Writes copies capital letter on worksheet	N/A	N/A
N/A	Free-Writes first name on worksheet	N/A	N/A
N/A	Hear-Mark circles animal when orally presented with its name	N/A	N/A
First component skill	Free-Writes tally mark on worksheet	38	42–34 marks/min
Second component skill	See-Mark circles number 0–20 among various symbols on worksheet	82	90–74 marks/min
Third component skill	See-Says number 0–20 when presented in random order on worksheet.	104	114–94 responses/min
Fourth component skill	Hear-Writes number 0–20 on worksheet when orally presented in random order.	45	50–41 digits/min
Fifth component skill	Free-Writes number from 0 to 20 in a continuous ascending stream.	69	76–62 digits/min
Composite skill	See-Writes single and double-digit addition facts with a sum ≤ 20	38	42–34 digits/min

The first four skills, which were unrelated to mathematics, did not have median scores or a range of scores as they were only used to familiarize the participants with the process before any baseline data were collected. Range is reported as maximum to minimum as the focus is on reaching the higher end of the range

Procedure

Mock Timings

To minimize the possibility of low performance due to the novelty of the PT process, all groups were familiarized with PT by practicing skills unrelated to mathematics before any baseline data were collected (see Table 2). Participants engaged in one 30-s timing per day across 4 days along with feedback, graphing their scores, evaluating their progress and receiving praise and a small token for participating (e.g., a sticker).

Component Skills

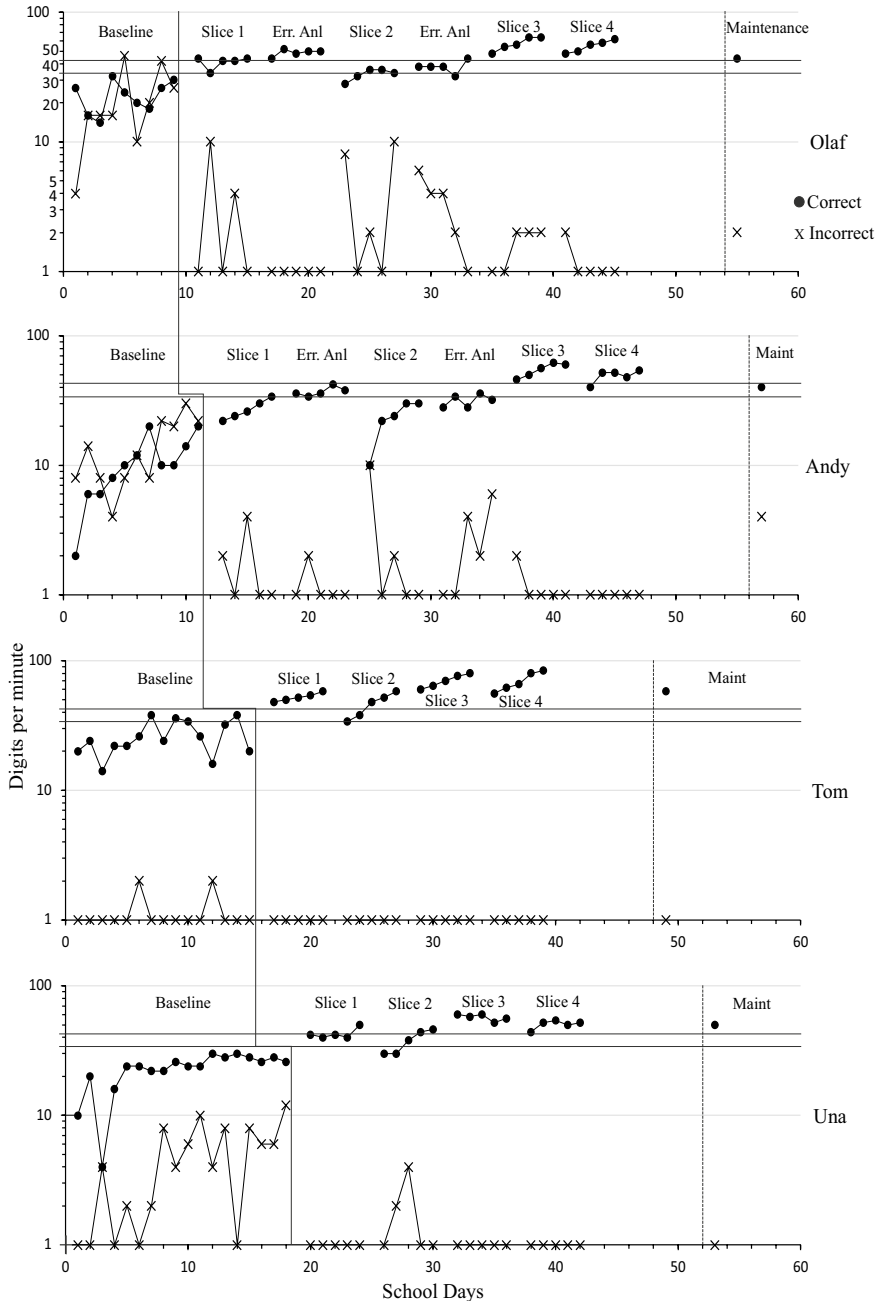
In line with component–composite analysis, the PT group received training in five component skills (see Table 2) before they practiced the composite skill (i.e., addition). When participants completed their training on all component skills, they were introduced to single-digit and double-digit addition with a sum ≤ 20 . The training was completed when participants demonstrated improvement on all skills by meeting or exceeding the median performance of the WC group. Baseline and practice conditions were similar to the ones for addition. The WC group's performance on all component skills was baselined for 3 days, while the SC group's was baselined once.

Baseline: Addition

The baseline condition included a 30-s timing on slice 4, which included all the numbers; no practice, graphing, or feedback, and 5 min of playtime at the end. For the PT group, the sequence followed the conventions of the multiple baseline design with Olaf being the first one to complete it and the rest following accordingly (see Fig. 1). WC group's performance was baselined for three days, while SC group's performance was baselined once. The median performance of the WC group, during baseline, was used to calculate a 10% target range and set the frequency aim for the PT group. Weekly assessment for the WC group ended at the same time as practice ended for the PT group.

PT and FBPC: Addition

When the baseline was completed, the PT group engaged in daily practice with a Board-Certified Behavior Analyst (BCBA) with over 7 years of experience. Practice consisted of an untimed and a timed component. Initially, participants engaged in untimed practice which included active student responding and timely feedback using similar worksheets to the ones used for the timings. Next, they engaged in FBPC (i.e., timings and performance feedback) for a maximum of five 30-s timings. During FBPC, participants were prompted to answer accurately and at their natural pace. They had a pre-specified daily frequency aim highlighted on their 'Timings' graph that they were expected to achieve through FBPC. The aim was adjusted daily



All students received PT through sliced practice. The first two students received an extra week on slices 1 and 2 where the tactic called error analysis was applied. A maintenance assessment was conducted 2 weeks after the completion of PT practice. The horizontal lines across the graph show the performance expectations set for the PT group. Zero values were transformed to a value of one to allow their graphing on the semi-log graph

Fig. 1 The PT group’s performance on See-Writes single and double-digit addition facts with a sum ≤ 20

based on the previous best score of the day so that it would be on a trajectory of at least $\times 1.30$ (i.e., 30%) increase per week. If participants reached it on the first timing, the lesson was completed for the day. If they did not meet the daily frequency aim after five timings, the next day's aim was calculated based on their best score and the study's decision protocol. After each timing, participants graphed their correct and incorrect digits on the 'Timings' graph. They visually evaluated their performance within the session and in relation to the highlighted frequency aim. Upon completion of their daily practice, they graphed the correct and incorrect digits of the best timing of the day on the 'Daily graph.' Thus, they evaluated their performance across days. Participants practiced each slice for a week although extra practice was provided, if necessary, following the decision protocol of the study.

Maintenance Assessment-Addition

Two weeks after the last participant completed the practice, an assessment was conducted to examine if participants maintained their improved performance in addition. Participants engaged in a 30-s timing on slice 4 without engaging in any prior practice, and without graphing their scores or receiving feedback. When the assessment was completed, participants were given a certificate to mark the completion of their practice.

Decision Protocol

The decision protocol was applied at the end of each week and was based on the performance (i.e., frequency of correct/incorrect digits) and learning rate (i.e., celeration) of the PT group. If participants met the frequency aim, they progressed to the next slice in the sequence. If performance was below the aim, but celeration was above $\times 1.30$ per week, they were given an extra week of the same practice. If performance was below the aim and celeration was less than $\times 1.30$ per week, error analysis was employed. Error analysis involved examining worksheets to identify which particular targets proved difficult to master and tailoring practice around those targets while worksheets for timed practice remained the same.

Duration of Sessions and Reinforcing Conditions

The maximum duration of a lesson was set to 1 h to allow comparisons between the PT, WC, and SC groups. On average, the PT group engaged in 47 min of daily practice (min: 46–max: 49) and an average of 62 sessions (min: 55–max: 66) as some students engaged in more practice following the study's protocol. The PT group received points throughout the lesson for engaging and completing worksheets and timings on a variable schedule of reinforcement (VR4). At the end of each session, points were exchanged by participants for a preferred item from the class-shop. Items included toys, building blocks, coloring materials, and the iPad®.

Interobserver Agreement

Interobserver agreement was calculated, depending on the skill, for an average of 39% of all the sessions for each participant (min: 38%–max: 39%). A BCBA with over 10 years of experience scored video recordings of the sessions. Agreement was calculated in a two-step manner. First, the correct and incorrect responses were calculated separately by dividing the smaller by the larger number and then multiplying by 100. The two percentages were added together and divided by two to produce an overall agreement for each skill. The overall average agreement was calculated by adding baseline, intervention, and post-assessment agreement for all the skills and dividing by three. Agreement was 99% (min: 97%–max: 100%) for Una, 94% (min: 87%–max: 100%) for Andy, 97% (min: 91%–max: 100%) for Olaf, and 100% (min: 99%–max: 100%) for Tom.

Procedural Fidelity

A separate checklist was created for the baseline, intervention, and final assessment phases. The baseline and final assessment fidelity checklists included the same 10 items, while the intervention checklist included 14 items. The same BCBA scored them by answering YES or NO for each item. Procedural fidelity was calculated for an average of 39% (min: 38%–max: 39%) of all the sessions for each participant. Procedural fidelity was 100% across all participants and all conditions.

Social Validity

Participants were provided with a questionnaire which included open-ended questions about different aspects of mathematics and a happy and unhappy face underneath each question. Students were prompted to circle the happy or unhappy face for each of the questions. At the end of the study, the participants were provided with the same questionnaire with additional questions relating to all aspects of their training (e.g., 'how do you feel about graphing your scores?').

Data Analysis

Precision Teaching Metrics

Data were plotted on using an online software program named PrecisionX. The program provided the SCC for visual analysis and calculated a series of behavioral measures (i.e., metrics). The primary metrics utilized were the: (a) level, (b) level change multiplier (c) celeration, (d) bounce, and (e) frequency multiplier. The level shows the average performance of the individual across time. The geometric mean was calculated as it is more appropriate for data plotted on the SCC, and it is less affected by extreme variables (Everitt and Howell 2005). The level change multiplier

produces a ratio of how much the average performance changed between the two conditions. Celeration (e.g., 5 responses per minute-per week) is a frequency derived measure quantifying the student's learning rate across time. Bounce produces a ratio which quantifies behavioral variability. Finally, the frequency multiplier provides a ratio of the change between two scores (e.g., first and last assessment). For all the ratios calculated, the multiplication (\times) or division (\div) sign were affixed to indicate an increase or decrease in performance across time (Kubina and Yurich, 2012). For this study, all ratios were transformed into percentages. For example, a $\times 2.00$ ratio equals a 100% increase in performance, while a $\div 2.00$ decrease equals a 50% reduction in performance.

Effect Size and Statistical Significance

In addition to the metrics, the Non-Overlap of All Pairs (NAP) was used to calculate PT's effect on the skills targeted (Vannest, Parker, Gonen, & Adiguzel, 2016). NAP is an appropriate effect size measure for single-case research which is correlated with R^2 (Parker & Vannest, 2009). NAP was calculated by comparing the correct digits during baseline with the correct digits across all the slices practiced. Finally, statistical significance was calculated via bootstrapping using the Simulation Modeling Analysis software retrieved from (<http://www.clinicalresearcher.org/software.htm>). Bootstrapping is a nonparametric approach to statistical inference during which variability within a sample is used to test the sampling distribution empirically. This occurs by repeatedly engaging in a random resampling, with replacement, from the sample (Lavrakas, 2008).

Results

Component Skills

Overall, the PT group made a meaningful improvement in all component skills (see Table 3). The average improvement from first-to-last assessment was 64%, (min: 47%–max: 108%), while the group's average NAP score was .98 (min: .94–max: 1.00) suggesting a strong effect. The WC group did not improve in all skills, as indicated by the range of improvement. On average, their performance increased by 24%, (min: –22%–max: 27%). The SC group increased by an average of 63% (min: 30%–max: 125%) across all component skills. The NAP was not calculated for the comparison groups as no intervention was delivered.

Composite Skill (Addition)

Olaf

Olaf's average performance with addition increased by 150% (see Fig. 1; top panel). Overall, his learning rate increased by 33% per week (min: 9%–max: 68%) and 59%

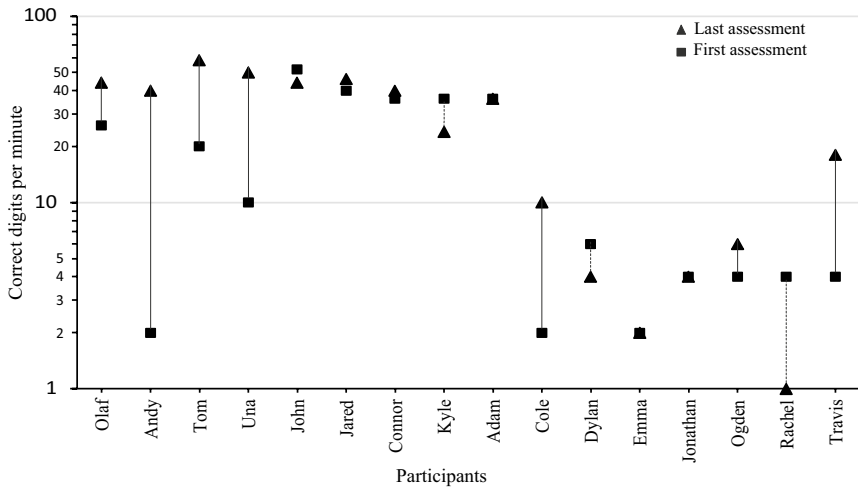
Table 3 The average performance of the three groups, on the component skills, across all phases of the study

Pinpointed skill	Experimental group	Baseline	PT or weekly monitoring	Final assessment
Free-Writes tally mark on worksheet	PT	Marks/min 33	49	47
	WC	39	48	42
	SC	23	–	29
See-Mark circles number 0–20 among various symbols on worksheet	PT	39	97	81
	WC	72	85	92
	SC	19	–	43
See-Says number 0–20 when presented in random order on worksheet	PT	Responses/min 57	89	86
	WC	102	93	98
	SC	24	–	43
Hear-Writes number 0–20 on worksheet when orally presented in random order	PT	Digits/min 34	61	49
	WC	47	48	44
	SC	13	–	17
Free-Writes number from 0 to 20 in a continuous ascending stream	PT	42	77	75
	WC	71	76	90
	SC	22	–	21

The geometric mean was used to calculate all the average scores

Some cells are empty as not all calculations were relevant for all students

PT, precision teaching group; WC, weekly-comparison group; SC, single-comparison group. Although the assessments lasted 30 s the values are being reported as frequency per minute



The lines allow a comparison of the participants' progress. The first four participants belong to the PT group, the next four to the WC group, and the last eight to the SC group. Zero values were transformed to a value of one to allow their graphing on the semi-logarithmic graph.

Fig. 2 Participants' first and last performance on the composite skill

per week on slice 4. The incorrect digits dropped from an average of 18 per minute to 0 per minute on slice 4. The average bounce of correct digits was 24%, (min: 10%–max: 40%) and 10% on slice 4. Olaf's NAP score was .98, 90% CI (.61, 1.00). Moreover, his improvement was significant ($r = .72$, $p < .02$, 95% CI). Finally, his overall improvement from first-to-last assessment was 69% (see Fig. 2).

Andy

Andy's average performance with addition increased by 444% (see Fig. 1, second panel). Overall, his learning rate increased by 90% per week (min: 25%–max: 478%) and 44% per week on slice 4. The incorrect digits dropped from an average of 12 per minute to 0 on slice 4. Andy's average bounce was 27% (min: 10%–max: 70%) and 30% on Slice 4. His NAP score was .98, 90% CI (.62, 1.00), while his improvement was significant ($r = .72$, $p < .03$, 95% CI). Improvement from first-to-last assessment was 1900%.

Tom

Tom average performance with addition increased by 176% (see Fig. 1, third panel). Overall, his learning rate increased by 90% per week (min: 38%–max: 163%) and 111% per week on slice 4. Tom had no incorrect digits recorded. His average bounce was 5%, (min: 0%–max: 10%) and 10% on slice 4. Tom's NAP score was .98, 90% CI (.64, 1.00), and his improvement was significant at ($r = .82$, $p < .01$, 95% CI). Improvement from first-to-last assessment was 190%.

Una

Una average performance with addition increased by 138% (see Fig. 1, bottom panel). Overall, her learning rate increased by 45% per week, (min: 19%–max: 138%) and 23% per week on slice 4. Una had an average of 3 incorrect digits per minute which dropped to zero on slice 4. Her bounce was an average of 17%, (min: 10%–max: 20%) across the slices, and 20% on slice 4. Finally, Una's NAP score was also .99, 90% CI (.71, 1.00), while her improvement was significant ($r = .84$, $p < .01$, 95% CI). Improvement from first-to-last assessment was 400%.

Weekly and Single Comparison Groups

The WC group demonstrated a stable performance across the weeks, with an average of zero incorrect digits. The stability of performance is evident when examining the participants' learning rate per month. John's learning rate decreased by 2% per month during the weekly probes, while his overall performance decreased by 15% from first-to-last assessment. Jared's learning rate increased by 1%, while his overall performance improved by 15%. Connor's learning rate decreased by 2%, while his overall performance improved by 11%. Kyle's learning rate decreased by 1%, while his performance improved by 33% from first-to-last assessment. For the SC group, Adam achieved a 0% change from first-to-last assessment. Cole improved by 400%, Dylan's performance decreased by 33%, while Emma and Jonathan produced a 0% change. Ogden improved by 50%, Rachel's performance decreased by 75%, and finally, Travis improved by 350%.

Discussion

PT has been demonstrated to be successful with typically developing students, but a need for further research on the use of PT with students with IDD has been indicated (Ramey et al. 2016). This study explored whether students with IDD could improve in mathematics following the introduction of PT, FBPC, and component-composite analysis, which are instructional tactics that are frequently embedded within the PT framework. A secondary aim was to examine if students with IDD who were of moderate mathematical ability, in relation to the other participants in the study, could equal or outperform their peers with fewer difficulties with their math skills.

Considering the primary aim, the results suggested that students with IDD can indeed benefit from PT. Despite the PT group's initial dysfluency in addition, their abilities improved significantly. Also, an improvement was demonstrated on all component skills. Therefore, a primary finding of this study was that daily PT practice focusing not only on quality but also the pace of responding led to important gains in all the skills. Thus, this paper provides additional evidence regarding the effectiveness of PT when used with children with IDD (Ramey et al. 2016).

Equally encouraging were the results of the maintenance assessments, where none of the PT group reverted to baseline levels (the only exception was Una on the

first component skill). This finding suggests that students who achieve higher frequencies can maintain their level of performance. Although fluency was not assessed through the test of Maintenance, Endurance, Stability, and Application (MESA), the acceptable maintenance of the skills could be linked to fluency outcomes reported in the literature (Fabrizio and Moors 2003; Weiss et al. 2008). Overall, the improvement led us to suggest that PT and its additional components are a viable and effective alternative for students with IDD, which could prove useful for teachers in special education looking to utilize evidence-based practices (Odom et al. 2005).

In regards to the second research aim, the results were also positive. The PT group reached or exceeded the average performance of their peers with fewer difficulties with their math skills. Also, none of the comparison participants ($n = 12$) managed to meet both the final performance and overall progress the PT group made in addition. This finding should be considered vital as it suggests that PT is a system which optimizes instruction better than traditional teaching practices. Therefore, PT could prove a valuable addition to current practice which should not be treated as a one-size-fits-all approach (Ledford et al. 2016).

The current study also included a series of other findings. First, some students needed more practice than others. Specifically, Olaf and Andy needed two more weeks, for the skill of addition, compared to Tom and Una. However, in the end, all four students met their frequency aim. This finding suggests that students with IDDs might be capable of making significant progress as long as they engage in enough practice. Thus, these findings have certain implications. On a theoretical level, the findings suggest it might be better for students to receive fewer parts of the curriculum so that they have enough time to master them. On a practical level, it confirms the importance of homogeneously grouping students so that the same lesson can be delivered.

Second, the participants' overall improvement demonstrated that they were not performing to the maximum of their abilities before PT was introduced. This finding may be linked to the argument made by McDowell et al. (2002) who discussed how cumulative dysfluency in prerequisite skills could significantly limit performance in more advanced skills. Despite being dysfluent in basic mathematical skills, those students had been taught addition, subtraction, and in some cases multiplication and division in their general classroom. So, the question arises as to whether they should have progressed that far in the curriculum.

Third, during PT the average frequency recorded across all the skills was 64/min (min: 10/min–max: 130/min). This finding suggests that students with IDDs might have a higher natural pace when training is appropriate. This is crucial considering that a lot of instructional tactics emphasize accuracy measures (e.g., discrete trial training; Smith 2001) which place a ceiling on students' performance.

Another current finding is related to the acceleration of the PT group' learning rate. The PT literature highlights $\times 1.30$ per week (i.e., 30%) as the absolute minimum expectation of progress (White and Haring 1980). However, more recent PT literature suggests that $\times 2.00$ (i.e., 100%) weekly expectation should be the minimum aim for teachers (Johnson and Street 2013). Although these acceleration values have been recommended for typically developing individuals, there are no official recommendations for students with IDDs. In this study, students'

celeration values exceeded $\times 1.30$ expectations, and in some cases, surpassed $\times 2.0$ expectations. This result suggests that students with IDD can potentially increase their performance by 100% or more every week. Although more data are warranted to be able to generalize such a conclusion, this finding still highlights that the potential to improve might be greater than expected for students with IDDs.

Finally, the fact that students practiced for an average of 47 min across 62 sessions demonstrated that students with IDDs can engage in rigorous practice as long as expectations are tailored to their current skill level, and certain reinforcing conditions are in place to keep them motivated.

Social Validity

Three out of the four participants said they enjoyed using a timer, having a frequency aim and being prompted to answer at their natural pace. Also, all four said they enjoyed graphing their scores and engaging in PT practice with the primary author. Similarly, 10 comparison participants said they enjoyed using the timer and responding to their natural pace. This suggests that PT seems to be an acceptable framework for the majority of students that took part in the study.

Strengths and Limitations

Some methodological limitations should be considered when interpreting the findings of this study. First and foremost, a quasi-experimental A-B design was used for all component skills, and therefore any positive results should be interpreted with caution. Due to limited resources, the multiple baseline design was utilized only for the skill of addition. Moreover, due to the small sample size, group comparisons were not possible, and a future study that includes a larger number of participants is needed to determine the efficacy of PT with students who have IDDs.

Furthermore, the allocation of participants was not randomized but rather based on their performance in the TEMA-3. There is a possibility that the results would be different if participants were randomly allocated to the groups. Finally, although precision teaching measures fluency through the MESA test, for this study the test was not employed as the aim of this project was to examine if the PT group would achieve intermediate aims based on WC group's performance.

Despite its limitations, this study also had some strengths. First and foremost, the effectiveness of PT was replicated across six skills for all four participants. In regards to the internal validity of the study, SC group's data indicated that the threat from maturation effects was minimal. Similarly, testing effects were evident in the control participants' data but not to the extent that would suggest a major problem with the study's internal validity. Moreover, this study managed to acquire these outcomes without exceeding the duration of a general lesson in mathematics that participants would normally receive. In addition, a clear functional relation was

demonstrated for the skill of addition. Finally, effect sizes indicated a strong effect of the intervention, while the results of addition were statistically significant.

Future Directions

Despite this study's encouraging results, a replication or extension with more students of different ability and other mathematical skills is suggested. It would also be worth examining whether students with IDD can reach frequencies that would allow them to pass the MESA test successfully. Such an outcome could show that fluency is a functional concept not necessarily affected by the presence of a diagnosis.

Funding No funding was received for this study.

Data Repository The standard celeration charts and additional data are available in the Figshare data repository. DOI: <https://doi.org/10.6084/m9.figshare.c.4931862.v1>.

Compliance with Ethical Standards

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

Ethical Approval Ethical approval was received from the ethics committee at the University of Kent.

Informed Consent Informed consent was obtained from the participants' parents as well as the participants' assent.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Barrett, B. H., School, F., Sopris, R. B., Binder, W. C., Cook, D. A., Engelmann, S., et al. (1991). The right to effective education. *The Behavior Analyst, 14*, 79–82. <https://doi.org/10.1007/BF03392556>.
- Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst, 19*, 163–197. <https://doi.org/10.1007/BF03393163>.
- Bouck, E. C., & Joshi, G. S. (2015). Does curriculum matter for secondary students with autism spectrum disorders: analyzing the NLTS2. *Journal of Autism and Developmental Disorders, 45*, 1204–1212. <https://doi.org/10.1007/s10803-014-2281-9>.

- Calkin, A. B. (2005). Precision teaching: The standard celeration charts. *The Behavior Analyst Today*, 6, 207–215. <https://doi.org/10.1037/h0100073>.
- Cavallini, F., Berardo, F., & Perini, S. (2010). Mental retardation and reading rate: Effects of precision teaching. *Life Span and Disability*, 13, 87–101.
- Chiesa, M., & Robertson, A. (2000). Precision teaching and fluency training: Making maths easier for pupils and teachers. *Educational Psychology in Practice*, 16, 297–310. <https://doi.org/10.1080/713666088>.
- Common Core Standards Initiative. (2010). *Common Core State Standards for mathematics*. National Governors Association Center for Best Practices, Council of Chief State School Officers. Retrieved from http://www.corestandards.org/wp-content/uploads/Math_Standards1.pdf.
- Daly, E. J., Martens, B. K., Bamett, D., Witt, J. C., Leaming, I., & Olson, S. C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology*, 36, 562–581.
- Datchuk, S. M., Kubina, R. M., & Mason, L. H. (2015). Effects of sentence instruction and frequency building to a performance criterion on elementary-aged students with behavioral concerns and EBD. *Exceptionality*, 23, 34–53. <https://doi.org/10.1080/09362835.2014.986604>.
- Department for Education. (2016). *DfE strategy 2015–2020 Word-class education and care*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/508421/DfE-strategy-narrative.pdf.
- Department for Education. (2017). *National curriculum assessments at Key Stage 2 in England, 2017 (revised)*. Retrieved from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/667372/SFR69_2017_text.pdf.
- Everitt, B., & Howell, D. (Eds.). (2005). *Encyclopedia of statistics in behavioral science: Volume 1* (Vol. 1). Chichester: Wiley.
- Fabrizio, M., & Moors, A. (2003). Evaluating mastery: Measuring instructional outcomes for children with autism. *European Journal of Behavior Analysis*, 4, 23–36. <https://doi.org/10.1080/1502149.2003.11434213>.
- Fitzgerald, D. L., & Garcia, H. I. (2006). Precision teaching in developmental mathematics: Accelerating basic skills. *Journal of Precision Teaching and Celeration*, 22, 11–28.
- Gilliam, J. E. (2006). *Gilliam autism rating scale* (2nd ed.). Austin, TX: Pro-Ed.
- Ginsburg, H. P., & Baroody, A. J. (2003). *Test of early mathematics ability* (3rd ed.). Austin, TX: Pro-Ed.
- Grafman, J. M., & Cates, G. L. (2010). The differential effects of two self-managed math instruction procedures: Cover, copy and compare versus copy, cover, and compare. *Psychology in the Schools*, 47, 153–165. <https://doi.org/10.1002/pits.20459>.
- Greene, I., McTiernan, A., & Holloway, J. (2018). Cross-age peer tutoring and fluency-based instruction to achieve fluency with mathematics computation skills: A randomized controlled trial. *Journal of Behavioral Education*, 27, 145–171. <https://doi.org/10.1007/s10864-018-9291-1>.
- 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, 186–209. <https://doi.org/10.1080/02667363.2015.1022818>.
- Haughton, E. C. (1980). Practicing practices: Learning by activity. *Precision Teaching*, 1, 3–20.
- Hughes, C. 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, 221–238.
- Johnson, K. J., & Street, E. M. (2012). From the laboratory to the field and back again: Morningside academy's 32 years of improving students' academic performance. *The Behavior Analyst Today*, 13, 20–40. <https://doi.org/10.1037/h0100715>.
- Johnson, K. J., & Street, E. M. (2013). *Response to intervention and precision teaching: Creating synergy in the classroom*. New York: Guildford Publications.
- King, S. A., Lemons, C. J., & Davidson, K. A. (2016). Math interventions for students with autism spectrum disorder: A best-evidence synthesis. *Exceptional Children*, 82, 443–462. <https://doi.org/10.1177/0014402915625066>.
- Kleinert, W. L., Coddling, R. S., Minami, T., & Gould, K. (2018). A meta-analysis of the taped problems intervention. *Journal of Behavioral Education*, 27, 53–80. <https://doi.org/10.1007/s10864-017-9284-5>.
- Kroesbergen, E. H., & Van Luit, J. E. H. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education*, 24, 97–114. <https://doi.org/10.1177/07419325030240020501>.

- Kubina, R. M., & Yurich, K. K. L. (2009). Developing behavioral fluency for students with Autism. *Intervention in School and Clinic, 44*, 131–138. <https://doi.org/10.1177/1053451208326054>.
- Kubina, R. M., & Yurich, K. K. L. (2012). *The precision teaching book* (1st ed.). Lemont, PA: Greatness Achieved.
- Lavrakas, P. J. (Ed.). (2008). *Encyclopedia of survey research methods*. Thousand Oaks, CA: Sage Publications Inc. <https://doi.org/10.4135/9781412963947>.
- Ledford, J. R., Barton, E. E., Hardy, J. K., Elam, K., Sebolt, J., Shanks, M., et al. (2016). What equivocal data from single case comparison studies reveal about evidence-based practices in early childhood special education. *Journal of Early Intervention, 38*, 79–91. <https://doi.org/10.1177/1053815116648000>.
- Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences, 25*, 126–133. <https://doi.org/10.1016/j.lindif.2013.02.001>.
- Lin, F., & Kubina, R. M. (2000). Learning channels and verbal behavior. *The Behavior Analyst Today, 5*, 1–14. <https://doi.org/10.1037/h0100134>.
- Lindsley, O. R. (1998). Learning channels next: Let's go. *Journal of Precision Teaching and Celeration, 15*, 2–4.
- Mcdowell, C., Keenan, M., & Kerr, K. P. (2002). Comparing levels of dysfluency among students with mild learning difficulties and typical students. *Journal of Precision Teaching and Celeration, 18*, 37–48.
- McTiernan, A., Holloway, J., Leonard, C., & Healy, O. (2018). Employing precision teaching, frequency-building, and the Morningside Math facts curriculum to increase fluency with addition and subtraction computation: A randomised-controlled trial. *European Journal of Behavior Analysis, 19*, 90–104. <https://doi.org/10.1080/15021149.2018.1438338>.
- National Center for Educational Statistics. (2017a). The Nation's Report Card-NAEP Mathematic's Report Card: Grade 4. Retrieved November 2, 2018, from https://www.nationsreportcard.gov/math_2017/nation/achievement?grade=4.
- National Center for Educational Statistics. (2017b). The Nation's Report Card-NAEP Mathematic's Report Card: Grade 8. Retrieved November 2, 2018, from https://www.nationsreportcard.gov/math_2017/nation/achievement?grade=8.
- Odom, S. L., Brantlinger, E., Gersten, R., Horner, R. H., Thompson, B., & Harris, K. R. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children, 71*, 137–148. <https://doi.org/10.1177/001440290507100201>.
- Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy, 40*, 357–367. <https://doi.org/10.1016/j.beth.2008.10.006>.
- Ramey, D., Lydon, S., Healy, O., McCoy, A., Holloway, J., & Mulhern, T. (2016). A systematic review of the effectiveness of precision teaching for individuals with developmental disabilities. *Review Journal of Autism and Developmental Disorders, 3*(3), 179–195. <https://doi.org/10.1007/s40489-016-0075-z>.
- Smith, T. (2001). Discrete trial training in the treatment of autism. *Focus on Autism and Other Developmental Disabilities, 16*, 86–92. <https://doi.org/10.1177/108835760101600204>.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales* (2nd ed.). Circle Pines, MN: AGS Publishing.
- U.S. Department of Education. (2004). *Individuals with Disabilities Education Act (IDEA)*. Washington, DC.
- Vannest, K. J., Parker, R. I., Gonen, O., & Adiguzel, T. (2016). Single case research: Web based calculators for SCR analysis. (Version 2.0) [Web-based application]. College Station, TX: Texas A&M University. Retrieved March 14, 2019, from <http://www.singlecaseresearch.org/calculators/nap>.
- Weisenburgh-Snyder, A. B., Malmquist, S. K., Robbins, J. K., & Lipshin, A. M. (2015). A model of MTSS: Integrating precision teaching of mathematics and a multi-level assessment system in a generative classroom. *Learning Disabilities: A Contemporary Journal, 13*, 21–41.
- Weiss, M. J., Fabrizio, M. A., & Bamond, M. (2008). Skill maintenance and frequency building: Archival data from individuals with Autism Spectrum Disorders. *Journal of Precision Teaching and Celeration, 24*, 28–37.
- White, O. R., & Haring, N. G. (1980). *Exceptional teaching* (2nd ed.). Columbus: Charles E. Merrill Publishing Co.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Affiliations

Athanasios Vostanis¹  · **Ciara Padden¹**  · **Mecca Chiesa¹** ·
Konstantinos Rizos² · **Peter E. Langdon³** 

Ciara Padden
c.m.padden@kent.ac.uk

Mecca Chiesa
m.e.chiesa@outlook.com

Konstantinos Rizos
konstantinos.rizos@forestbridgeschool.org.uk

Peter E. Langdon
peter.langdon@warwick.ac.uk

- ¹ Tizard Centre, University of Kent, Cornwallis North East, Canterbury, Kent CT2 7NF, UK
- ² Forest Bridge School, Chiltern Road, Maidenhead, Berkshire SL6 1XA, UK
- ³ Centre for Educational Development, Appraisal and Research, University of Warwick, Coventry CV4 8UW, UK