



Peer-reviewed articles and lesson plans written by teachers and researchers to inform classroom practice.

Editors

Gil Naizer
April Sanders
Laura Isbell
Tami Morton
Susan Williams

IMPACT OF SHORT-TERM INTENSE FLUENCY INSTRUCTION ON STUDENTS' READING ACHIEVEMENT: A CLASSROOM- BASED, TEACHER-INITIATED RESEARCH STUDY

Kristy DiSalle

Dorr Street Elementary School, Toledo, OH

Timothy Rasinski, Ph.D.

Kent State University, Kent, OH

Abstract

Reading fluency continues to be a critical factor in elementary students' reading development. Many students, who struggle in reading, manifest difficulties in some area of fluency. In the present study, a fourth grade teacher implemented the Fluency Development Lesson (FDL), an intensive fluency instructional routine, with her six lowest achieving students. In a twelve-week implementation of the FDL all students made significant and substantial progress in both fluency and reading comprehension. The authors make a call for more studies of intensive fluency interventions to demonstrate and confirm their effectiveness.

Keywords: Fluency, Comprehension, Reading, Struggling Readers

Introduction

Recent policy positions and scholarly research in reading education have identified reading fluency as a critical and essential reading competency that is necessary for full proficiency in reading. Both the National Reading Panel (2000) and the Common Core State Standards (2016) have noted reading fluency as foundational for reading growth and should be mastered in the elementary grades. Research has found that approximately 75% of students who struggle in high stakes tests of reading achievement demonstrate difficulty in one or more components of reading fluency (i.e., word recognition accuracy, word recognition automaticity, and reading prosody) (Valencia & Buly, 2004). Moreover, other research has shown many students beyond the elementary grades continue to struggle in reading fluency and that measures of reading fluency continue to be highly correlated with overall reading proficiency (Paige, Magpuri-Lavell, Rasinski, & Smith, 2013; Paige, Rasinski, & Magpuri-Lavell, 2012; Rasinski, Padak, McKeon, Krug, Wilfong, Friedauer, & Heim, 2005). Indeed, comprehensive and scholarly reviews of research related to fluency have concluded that it is a critical component for success in learning to read (Chard, Vaughn, & Tyler, 2002; Kuhn & Stahl, 2003; Rasinski, Reutzel, Chard, & Linan-Thompson, 2011).

Literature Review

Fluency is important because it is a prerequisite to more sophisticated levels of reading comprehension (Rasinski, 2012). Once students are able to read words in texts accurately, automatically and with expression that reflects meaning, students are more able to focus their cognitive resources on making meaning - comprehension—rather than on the more basic and foundational competencies in reading – word recognition.

Despite the scholarly work that has consistently demonstrated the relevance of reading fluency to reading achievement, there seems to be a degree of dismissiveness toward fluency within the literacy community. In the annual *What's Hot; What's Not* survey of literacy experts (e.g., Cassidy, & Grote-Garcia, 2014) reading fluency has been consistently identified as a reading competency that is *not hot* and *should not be hot*. While the *What's Hot; What's Not* survey does not speculate as to the reasons for this reaction to fluency, it may be that the negative reaction to fluency has been caused by the way in which fluency instruction has been manifested in many programmatic approaches to fluency. In these programs, fluency is measured by students' reading speed as measured by words read correctly per minute (WCPM). Research acknowledges that this is an appropriate measure of word recognition automaticity which itself is related to general reading proficiency (Fuchs, Fuchs, Hamlett, Walz, & Germann, 1993; Fuchs, Fuchs, Hosp, & Jenkins, 2001). Given the correlation between reading speed and word recognition automaticity, many fluency instructional programs as well as many well-meaning teachers have implicitly reversed the logic and made instruction that focuses on increasing reading rate as the primary method for improving fluency. We see this manifested in the regular use of *timed readings* in which students are asked to read and reread relatively short passages at an ever

quicker pace. This type of speed-oriented reading is in opposition to the authentic meaningful reading that most reading scholars feel is key to reading growth. Moreover, there is no compelling research that has demonstrated that explicit instruction in increasing students' reading speed results in improved overall reading.

Literacy scholars suggest that reading fluency is best developed through some very basic reading activities. These include word recognition instruction, listening to fluent readings of texts, wide reading, repeated reading, and assisted reading in which students read a text while simultaneously hearing the text read to them in a fluent manner (Rasinski, 1989, 2010). Individually there is a good body of research that supports wide reading, repeated reading, and assisted reading (Rasinski, Reutzel, Chard, & Linan-Thompson, 2011). However, even greater effects can be anticipated when these individual instructional approaches are combined in a synergistic and authentic manner.

The Fluency Development Lesson. The Fluency Development Lesson (FDL) (Rasinski, Padak, Linek, & Sturtevant, 1994) was developed as a fluency intervention that can be applied to large groups of typically developing elementary grade student or more intensively to smaller groups of students who have yet to achieve proficiency in fluency and who also struggle in overall reading achievement. The FDL is a daily lesson in which students are given the task of mastering to the point of fluency a new relatively short (100-200 words) text each day. The lesson takes approximately 20 minutes and can be implemented with classroom groups, small groups, or individual students. Throughout any part of the FDL there is never an explicit or implicit focus on increasing reading rate. The general daily protocol for the FDL involves the following steps:

1. In preparation for the lesson the teacher selects a text for the day. The text can be a passage from a story, an informational piece, a poem, or a song. The texts should be at or slightly above the students' instructional reading level and should be a reading with good phrasing and expression. The teacher makes two copies of the text for every student, and also makes a larger display copy for group reading.
2. Modeling Fluent Reading. The teacher introduces the display copy of the text to students and reads it to the students two to three times while students follow along silently. The teacher can read the text with various forms of expression or lack of expression.
3. Following the teacher's reading, students are led in a brief discussion of the text and the nature of the teacher's oral reading.
4. Assisted Reading. Next, the teacher and students read the display copy of the text two to three times chorally. The choral readings can change from the whole group reading the text to having different subgroups read the passage.
5. Assisted and Repeated Reading. Following the choral reading, students are divided into groups of two or three, given their individual copies of the text, and are given about five minutes to practice the text in their groups. One student reads the passage while his or her partner(s) follow along silently, provide help as needed, and

give positive feedback. Each student is given the opportunity to practice in this manner.

6. At this point students are able to read the text with some degree of fluency. In order to make the FDL an authentic activity, students are invited to perform their text for an audience. The audience can simply be other classmates, but it can also be made up of volunteer adults stationed outside the classroom, or even other classrooms of students.
7. Word Work. At the end of the performance the teacher and students select 5-10 words from the passage and engage in quick word study activities. These can include finding other words that contain a selected rhyme or word family from the passage (e.g. From the poem *Rain Rain Go Away*, other -ay words such as *day*, *play*, *stay*, and *stray* can be discovered and displayed for students to read), sorting the corpus of words in various ways, examining the morphological nature of certain words (e.g. *tract* is a morpheme in *tractor* means to pull; other words that contain the tract morpheme and that mean to pull include *distract*, *attract*, *extract*, and *contract*), and playing word games (e.g. word ladders using words from the passage). The formal FDL ends with the word study.
8. Repeated Reading. The FDL continues at home. Students take their second copy of the passage and are encouraged to read the passage to family members at home a select number of times, usually five or more.
9. Repeated Reading. A new Fluency Development Lesson is implemented the following day with a new text. However, before beginning to read the new text, the teacher leads students in reading and celebrating their mastery of texts from previous days.

Teachers employing the FDL are encouraged to vary the protocol to meet their own style of instruction and needs of the students. The key elements required in any FDL are modeling fluent reading, assisted reading, repeated reading, and word work. The key goal for any FDL is for students to master a new text (poem) with each lesson to the point of reading the text with good fluency - word recognition accuracy, automaticity, and expression.

The present action research study attempted to determine the effects of the Fluency Development Lesson when employed in a regular classroom setting with students identified as struggling in reading. The following research question guided the study: What are the effects of a regular classroom implementation of the Fluency Development Lesson on the reading achievement of fourth grade struggling readers?

Methodology

Implementing the FDL in a Real Classroom Setting. Kristy DiSalle is a fourth grade teacher who attended a series of professional development workshops by Timothy Rasinski. During his workshop Rasinski argued for the need for intensive fluency instruction for struggling readers (students who are reading at least one grade level below their assigned grade level). He described the FDL in depth and provided a simulation activity for teachers attending the workshop. Having six students in her class who were struggling in reading,

DiSalle decided to implement the FDL with them daily in addition to their regular reading curriculum. The regular reading curriculum included guided reading activities in which groups of students read, discuss, and respond to stories and informational texts, word study, and writing.

DiSalle had 6 students who were reading at a 2nd grade level, according to the reading assessment used by the school (Renaissance Learning, 2016). Kristy began using the FDL with her 6 struggling students (3 girls and 3 boys) for 4 to 5 days a week for approximately 20 minutes per lesson for a three-month period. She also added reciprocal teaching comprehension activities (Oczkus, 2010) to the FDL to create a more intense focus on comprehension. Five of the six students (Students 1 through 5 in Table 1) also received Title 1 reading services 4-5 days per week for 20 minutes. Up until the implementation of the FDL, Kristy's six struggling readers had not been making adequate progress (their reading achievement had remained stagnant from the beginning of the school year until the beginning of school until the implementation of the FDL protocol) in reading despite regular classroom reading instruction and Title I intervention for 5 students. The FDL was chosen and implemented for its potential to accelerate students' growth in reading.

Over the course of the implementation period the FDL was administered approximately 50 times to the group of students (because of absences not all students participated in all lessons). The FDL intervention was implemented at a table in the back of the classroom. The remainder of the class worked independently at their seats on vocabulary building activities during the FDL. Kristy chose a new poem for each lesson. She used seasonal themes and difficulty of the poem as chief criteria for poem selection. Poems were selected from a variety of authentic sources (e.g., Liatsos, 1995; Scholastic, 2004).

DiSalle began the FDL by having students skim and scan the daily poem to make predictions about the content and structure of the poem. Then as students progressed through the FDL, they worked to clarify words or phrases they didn't understand, formed questions about content in the selection, and created a summary of the poem. With the poem projected on the Smartboard, students read and performed the poem to the rest of class using music stands donated by the middle school band teacher. Kristy's word study activities included identifying common word families in poems and playing word games using words from the selected each poem. The home portion of the FDL involved the students reading the poem three times to family members. One parent commented, "The poems are a great tool to use at home, as we enjoy reading together. We appreciate the reading fluency lesson, as Jay is benefiting greatly. Thank you!"

Assessment Method. The STAR Reading assessment (Renaissance Learning, 2016), the reading assessment that was chosen by the school to measure students' reading achievement and progress, was used to measure students' growth in reading. It is a computer-adaptive assessment of general reading achievement and comprehension of students in grades 1 through 12. The assessment provides information on students' general performance in reading comprehension. The difficulty of items is adjusted automatically to reflect the skill level of all students, including students with special needs. Students read a

series of selections, with the length and difficulty dependent on the student's level and progress during the assessment, and answer multiple-choice comprehension questions. The STAR Reading assessment generates a Grade Equivalent (GE) score to measure proficiency level in comprehension. The National Center for Response to Intervention's analyses of the STAR Reading assessment reports validity coefficients ranging from .68 to .92 and reliability coefficients from .81 to .92 (National Center for Response to Intervention, 2016).

The Estimated Oral Reading Fluency (Est. ORF) is part of the STAR Reading assessment and is a measure of proficiency in fluency (word recognition accuracy and automaticity), foundational reading competencies that many students who struggle in reading are not proficient (Valencia & Buly, 2004). The Est. ORF was developed by linking STAR assessment data with known oral reading fluency assessments of over 12,000 students in grades 1 through 4. The Est. ORF is reported in words read correctly per minute on grade level material.

Results

Students Gain in Comprehension and Fluency. Students were tested on November 30, 2015, and March 9, 2016, to determine progress. Testing included the STAR Reading test to determine a student's growth in reading comprehension and an estimated oral reading fluency. The results of both assessments can be seen in Tables 1 and 2.

The tables provide a summary of the assessments of students. Table 1 details the results for each student while Table 2 summarizes results for all 6 students. In approximately 12 weeks of using the FDL, the students made on average slightly over a year's growth in reading comprehension. Oral reading fluency increased from 69.2 to 96.8 word read correctly per minute (WCPM), a gain of 27.6 WCPM during the implementation of the FDL.

Although only one of the six students achieved grade level proficiency in comprehension, the gains all six students made in both comprehension and fluency are remarkable in terms of their magnitude. In approximately one third of a school year, students made slightly over a year's growth in reading comprehension. Moreover, at the initial assessment students' reading comprehension was, on average, at a 2.8 grade level equivalent. This means that in roughly 3.3 years of reading instruction (Grades 1 through 4) prior to the initial assessment the six students in the present study made, on average, slightly over a half year's progress in comprehension per year of instruction. During the 12 week FDL implementation students made as much progress in comprehension as had been previously made in approximately 2 years.

Similarly, all six students in the present study have not achieved benchmark levels for oral reading fluency as set by the STAR Reading assessment. Still, the gains made by students in fluency are remarkable. Assuming 3.3 years of reading instruction prior to implementing the FDL, the six students in the present study made, on average an increase of 21 WCPM per year. In the 12 week implementation of the FDL, students in the present cohort made an

average gain of 27.6 WCPM. In other words, the students in our study using the FDL made a greater gain in a 12 weeks implementation of the FDL than students had on average made in an entire year of reading instruction.

During the period of the FDL intervention period all six of Kristy's struggling readers made significant progress in both comprehension and reading fluency. Other gains were measured by student feedback. One student commented, "FDL helps me present in front of the class. It is fun and I like the poems because they are silly. Now I read better, like I have more expression. It also helps me learn new words that I don't understand." Another student shared, "It (FDL) helps me understand words. It also helps me with my expression and to not be scared to perform in front of people."

Table 1: Student Progress Using the Fluency Development Lesson

| Date | Student | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |
|----------|---------|--------------------------------|--------------------------|-------------------|----------------------|
| 11/12/15 | 1 (boy) | 2.8 | 4.23 | 73 | 125 |
| 3/9/16 | | 3.4 | 4.62 | 93 | 139 |
| G Gain | | +0.6 | +0.39 | +20 | +14 |
| | | | | | |
| Date | Student | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |
| 11/12/15 | 2 (boy) | 2.4 | 4.25 | 63 | 125 |
| 3/9/16 | | 4.1 | 4.62 | 105 | 139 |
| Gain | | +1.7 | +0.37 | +42 | +14 |
| | | | | | |
| Date | Student | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |

| | | | | | |
|-------------|----------------|---------------------------------------|---------------------------------|--------------------------|-----------------------------|
| 11/12/15 | 2 (boy) | 2.4 | 4.25 | 63 | 125 |
| 3/9/16 | | 4.1 | 4.62 | 105 | 139 |
| Gain | | +1.7 | +0.37 | +42 | +14 |
| | | | | | |
| Date | Student | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |
| 11/12/15 | 3 (girl) | 2.4 | 4.23 | 63 | 125 |
| 3/9/16 | | 3.1 | 4.62 | 82 | 139 |
| Gain | | +0.7 | +0.39 | +19 | +14 |
| | | | | | |
| Date | Student | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |
| 11/12/15 | 4 (boy) | 2.4 | 4.23 | 63 | 125 |
| 3/9/16 | | 3.4 | 4.62 | 91 | 139 |
| Gain | | +1.0 | +0.39 | +28 | +14 |
| | | | | | |
| Date | Student | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |
| 11/12/15 | 5 (girl) | 2.9 | 4.23 | 79 | 125 |
| 3/9/16 | | 3.4 | 4.62 | 93 | 139 |

| Gain | | +0.5 | +0.39 | +14 | +14 |
|----------|----------|--------------------------------|--------------------------|-------------------|----------------------|
| | | | | | |
| Date | Student | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |
| 11/12/15 | 6 (girl) | 2.8 | 4.23 | 74 | 125 |
| 3/9/16 | | 4.6 | 4.62 | 117 | 139 |
| Gain | | +1.8 | +0.39 | +43 | +14 |

Table 2: Student Progress Summary Results

| Date | Comprehension Grade Equivalent | Expected Grade Placement | ORF Scores (WCPM) | Expected Grade 4 ORF |
|----------|--------------------------------|--------------------------|-------------------|----------------------|
| 11/12/15 | 2.6 | 4.23 | 69.2 | 125 |
| 3/9/16 | 3.7 | 4.62 | 96.8 | 139 |
| Gain | +1.1 | +0.39 | 27.6 | +14 |

Discussion and Implications

What Does This Mean? We acknowledge that this non-traditional form of research has many limitations. The number of students participating is quite low and makes it difficult to generalize to other situations. We did not have a control group against which we could compare results. Given that the research was done within an actual classroom setting meant that we could not control other instructional and other variables that may have impacted the results. We also note that prosody, an important component of reading fluency, was not assessed in the present study. Although parents and teacher observations as well as student comments noted improvements in prosody, it was not formally assessed.

On the other hand, the fact that this study was done within the context of an actual classroom also meant that it took place in an authentic school placement in which the teacher (Kristy DiSalle) had to deal with all the classroom exigencies that teachers deal with on a daily basis. She was unable to control for when students may have been ill or called out of the classroom for other reasons. Moreover, she still had to deal with issues related to her other students in her classroom during the FDL instruction. In other words, the real-life-classroom nature of the study means that the results that were observed from using the FDL can occur when teachers intentionally implement the lesson in order to meet the specific needs of students.

Despite the acknowledged limitations noted above, we do think there is much to take from the present study. Previous research has demonstrated that fluency is a critical variable for reading achievement. As reading fluency is achieved, readers are able to channel their cognitive resources to comprehension of text, the true goal of reading. Moreover, research has also shown that many students struggle to achieve appropriate levels of reading fluency. Clearly instructional interventions to improve and even accelerate students' fluency development are called for. Many current commercial approaches to fluency implicitly or explicitly tend to focus on an extraneous variable to fluency – speed of reading. These programs improve reading speed, a measure of fluency, by focusing students' attention on increasing speed. Although reading speed may increase, there is limited evidence that word recognition automaticity and reading comprehension also improve.

Rather than focus on increasing reading speed, the Fluency Development Lesson combines widely acknowledged components of exemplary reading fluency instruction to create an authentic reading activity in which students rehearse a text in order to eventually perform it for an audience. The present study found that implementation of the FDL with students who have demonstrable difficulty in reading fluency can dramatically accelerate their growth in both reading fluency and comprehension in a relatively short period of time. The present study suggests that the FDL, as well as other synergistic fluency protocols, can make a real difference in helping struggling readers move toward proficiency in both the foundational reading competencies and reading comprehension.

One study of six students is clearly not definitive. However, we hope that the present study will prompt other teachers and school staff to implement the FDL in their own classrooms and schools. The results of many small scale studies conducted in real classrooms can move the field of literacy education forward to the point where many students have new hope for gaining full proficiency in reading.

About the Authors

Kristy DiSalle. Kristy DiSalle is a 4th grade teacher at Dorr Street Elementary School in Toledo Ohio. She is dedicated to continuing her professional learning in order to provide her students with the best possible instruction. Email: kristydisalle@springfield-schools.org.

Timothy Rasinski, Ph.D. Timothy Rasinski is a professor of literacy education at Kent State University. His scholarly interests include reading fluency and word study, readers who struggle, and parental involvement. In 2010 Rasinski was elected to the International Reading Hall of Fame. Email: trasinsk@kent.edu.

References

- Cassidy, J., & Grote-Garcia, S. (2014). *What's Hot, What's Not* survey. *Reading Today*, 32 (1), 8-12.
- Chard, D.J., Vaughn, S., & Tyler, B. (2002). A synthesis of research on effective interventions for building fluency with elementary students with learning disabilities. *Journal of Learning Disabilities*, 35, 386-406.
- Common Core State Standards Initiative (2016). Common Core State Standards Initiative. Downloaded May 12, 2016 from: <http://www.corestandards.org/ELA-Literacy/>.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? *School Psychology Review*, 22, 27-48.
- Fuchs, L. S., Fuchs, D., Hosp, M., & Jenkins, J. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies in Reading*, 5, 239-256.
- Kuhn, M.R., & Stahl, S.A. (2003). Fluency: A review of developmental and remedial practices. *Journal of Educational Psychology*, 95, 3-21.
- Liatsos, S. (1995). *Poems to Count On: 32 Terrific Poems and Activities to Help Teach Concepts*. New York: Scholastic Professional.
- National Center for Response to Intervention. (2016). STAR Reading. Washington, DC: American Institutes for Research. Accessed 4/28/2016 at <http://www.rti4success.org/star-reading>.
- National Reading Panel. (2000). *Report of the National Reading Panel: Teaching children to read. Report of the subgroups*. Washington, DC: U.S. Department of Health and Human Services, National Institutes of Health.
- Oczkus, L. (2010). *Reciprocal Teaching at Work: Powerful Strategies and Lessons for Improving Reading Comprehension*. Newark, DE: International Reading Association.
- Paige, D. D., Magpuri-Lavell, T., Rasinski, T. V., & Smith, G. (2013). Interpreting the relationships among prosody, automaticity, accuracy, and silent reading comprehension in secondary students. *Journal of Literacy Research*, 46(2), 123-156.
- Paige, D. D., Rasinski, T. V., & Magpuri-Lavell, T. (2012). Is fluent, expressive reading important for high school readers? *Journal of Adolescent & Adult Literacy*, 56(1), 67-76.
- Rasinski, T. V. (1989). Fluency for everyone: Incorporating fluency in the classroom. *The Reading Teacher*, 42, 690-693.
- Rasinski, T. V. (2010). *The Fluent Reader: Oral and Silent Reading Strategies for Building Word Recognition, Fluency, and Comprehension*. New York, NY: Scholastic.

- Rasinski, T.V. (2012). Why reading fluency should be hot. *The Reading Teacher*, 65, 516-522.
- Rasinski, T., Padak, N., McKeon, C., Krug, Wilfong, L., Friedauer, J., & Heim, P. (2005). Is Reading Fluency a Key for Successful High School Reading? *Journal of Adolescent and Adult Literacy*, 49, 22-27.
- Rasinski, T. V., Padak, N. D., Linek, W. L., & Sturtevant, E. (1994). Effects of fluency development on urban second-grade readers. *Journal of Educational Research*, 87, 158–165.
- Rasinski, T. V., Reutzel, C. R., Chard, D. & Linan-Thompson, S. (2011). Reading Fluency. In M. L. Kamil, P. D. Pearson, B. Moje, & P. Afflerbach . (Eds), *Handbook of Reading Research, Volume IV* (pp. 286-319). New York: Routledge.
- Renaissance Learning. (2016). *STAR Reading*. Wisconsin Rapids, WI: Renaissance Learning. Accessed 4/28/2016 at <http://www.renaissance.com/Products/Star-Assessments/Star-Reading>.
- Scholastic. (2004). *Perfect Poems with Strategies for Building Fluency: Grades 3–4*. New York, NY: Scholastic.
- Valencia, S. W., & Buly, M. R. (2004). Behind test scores: What struggling readers really need. *The Reading Teacher* , 57, 520-531.

USING TEACHER ACTION RESEARCH IN HIGH SCHOOL CHEMISTRY TO DEVELOP NOVEL ASSESSMENT TOOLS

Chad Husting

Sycamore High School, Cincinnati, OH

Jordan Harshman, Ph.D.

University of Nebraska - Lincoln

Ellen Yezierski, Ph.D.

Miami University

Abstract The analysis of three years of reflecting on teaching practices in a blog lead a teacher to realize that assessment plays a critical role in day to day teaching and could be improved in his classroom. The literature does not provide very much on the "nuts and bolts" of using assessment data to inform teaching; however, data-driven inquiry provides a useful framework for addressing this problem. A method and tool were developed to facilitate rapid and meaningful student data collection, scoring, and instructional decision-making in a high school chemistry setting. The tool captures groups of students' knowledge, understandings, and misconceptions at the beginning of the year in an efficient and effective manner by leveraging easily accessible software (Excel). The student data are linked to tested and research-driven curriculum materials with a strong basis in chemistry education theory. The assessment methods and tool, its impetus, design considerations, functionality, and implementation are described with hope of introducing teachers to approaches that can improve how they coordinate assessment with curriculum and instruction to improve student learning.

Keywords: assessment, formative assessment, data-driven inquiry, teacher action research, collaborative action research, blogging, qualitative analysis, high school, science, chemistry

Introduction

For almost one hundred years, critical reflection on teaching has been noted as one of the most important processes for improving instruction (e.g., Dewey, 1933; Yost, Sentler, Forenza-Bailey, 2000). Recently, scholars have examined web logs or “blogs” as tools to promote and facilitate teacher reflection (Ray & Hocutt, 2006; Loving, Schroeder, Kang, Shimek & Herbert, 2007; Killeavy & Moloney, 2010; Luik, Voltri, Taimalu & Kalk, 2011). Although blogs have been shown to be good for critical events, they may not elicit deep reflection without explicit prompting (Wopereis, Sloep & Poortman, 2010). Blogging has also been used as a research tool in ethnographic fieldwork (Saka, 2008), which provides a useful framework for the analysis of teacher blogs. This study incorporates blogging and qualitative analysis within the framework of teacher (or action) research, which has been shown to be a highly effective means of promoting teacher growth and has been a common component of professional development (e.g., Lytle & Cochran-Smith, 1992).

Literature Review

In Tobias and Baffert’s work (2009), they have recently highlighted the need for elevating the professional status of science and mathematics teachers. Action research may also be a vehicle for this end, as teacher research has been viewed by Burton and Bartlett (2005) and Zeichner and Noffke (2001) as a highly effective means of collaborating with teachers to improve their professionalism and the status of the teaching profession in general. Zeichner and Noffke (2001) also found that teacher research contributes to the knowledge base. Such findings suggest that action research can produce many intellectual and social benefits to the teacher as well as their students and the academic community. The similarity between teaching and research has been called out by Huberman (1996) and Freeman (1998) and has recently framed a key work by Meijer, Oolbekkink, Meirink, and Lockhorst (2013). Meijer *et al.* (2013) describes how the linkages between teaching and research parallel key ideas in higher education first introduced by Boyer (1990) who coined “the scholarship of teaching of learning” thus expanding the traditional idea of research as only the scholarship of discovery to four areas. Coppola, Banaszak Holl, and Karbstein (2007) describe how teaching and research are integrated and informed, intentional, impermanent, and inheritable. However, when the researcher is not the classroom teacher, much can be missing from the research. Teachers have the knowledge of the students, classroom, and school environment as well as the needs of all stakeholders including themselves (Meijer, Oolbekkink, Meirink, & Lockhorst, 2013). To improve teacher learning, research quality, scholarly impact, authenticity, and impact on students, collaborative action research, in particular, is highly valued (Meijer, Oolbekkink, Meirink, & Lockhorst, 2013), in which teachers work with science education research faculty.

The study here details a collaborative action research project carried out by a high school chemistry teacher and a chemistry education research faculty member and graduate student. The study was guided by Anderson and Herr’s (1999) five validity types for practitioner research (outcome, democratic, catalytic, dialogic, and process validity) with a major emphasis on outcome validity, “whether the research undertaken leads to outcomes

for teachers and for the school.” As such, we will not only present the methods and findings of the study in a traditional manner, but we also reveal the major consequence of the knowledge generated by this collaborative endeavor. The aim of this last piece is not only to demonstrate outcome validity but also is to provide a rarely seen product in reports of action research. The description of how the teacher responded to the research findings and the product inspired by the findings are presented at the end of this article.

The impetus for the action research project theoretically lies in increasing the quality and quantity of teacher reflection (Dewey, 1986). Specifically, how employing descriptive, critical, and comparative reflection (Jay & Johnson, 2002), can help teachers identify, unpack, and overcome important classroom learning problems. The impetus for the project more practically comes from the first author’s participation in the Master Teacher Program sponsored the state of Ohio. Teachers were provided time to meet in groups throughout the year to create portfolios. The portfolios focused on evidence that examined what the teachers were doing before, during, and after a lesson to help students. It also focused on evidence that demonstrated communication with students, parents and community involvement. The following year, in attempt to continue reflecting on teaching and documenting evidence for student learning, the first author began creating blog entries focusing on the same themes from the master teacher program the previous year. Each blog entry focused on examining, constructing, and developing evidence that demonstrated student achievement in the classroom. Another major influence for teacher change and reflection in the blog was the author’s participation in the Target Inquiry Program at Miami University, Oxford Ohio (TIMU). This was an intensive two and a half year program funded by the National Science Foundation. It focused on chemistry teachers doing the scientific process and translating this into their teaching by developing and implementing inquiry labs and studying their impacts on student learning (<http://targetinquirymu.org/>).

Methodology

The research question that guided this study was: How can a teacher researcher’s qualitative analysis of blog entries be used to focus instructional improvement strategies?

Setting. This study was carried out at a suburban public school in southwest Ohio. Of the approximately 1800 students, 73% of the students are White, about 8% are Black, about 10% are Asian/Pacific Islander, about 3% are Hispanic and about 5% are two or more races. The instructor (first author) has a master’s degree and is considered “Highly Qualified” by the State of Ohio. He has 24 years of teaching experience, mostly in chemistry. Over the course of this study (3.5 years), the teacher schedule consisted of two classes of Accelerated Chemistry and three classes of Academic Chemistry. Accelerated Chemistry students were primarily tenth graders with a few eleventh graders who were considering future science classes in high school and college. The Academic Chemistry classes consisted mainly of eleventh graders with a few twelfth graders who are college bound but not necessarily considering taking science classes in their future.

Data Collection. Beginning in 2011-12, the first author began writing a weekly blog responding to these questions:

1. What am I doing to help kids achieve?
2. How do I know when they are there?
3. What is the evidence?

The blog entries were posted on <http://simpleteach.blogspot.com/> and aimed to be more than a personal reflection because they focused on the evidence used to reach conclusions about and the teacher's role in student learning. There were 97 entries (averaging 329 words each) collected between December 2012 and June 2015.

Data Analysis. The blog entries were printed and read multiple times. With collaboration from a first year college student, entries were placed into categories. The first author generated category names and descriptions to serve as codes. Each entry was coded such that it had exactly one code assigned to be most descriptive. Approximately 10% of the entries were coded by the third author and found that the descriptions were not detailed enough to attribute only code to each blog entry. As such, the first author revised the descriptions and selected 12 random blog entries. These entries along with the category names and descriptions were given to a new rater. Ten of this rater's assignment of entries to categories matched with the first author's assignments for this subset of the data corpus. The final category names, descriptions, and frequencies of occurrence in the data set are found in Appendix A. In subsequent analyses, we grouped the categories into themes. The emergent themes describing the data are presented in the Results.

Results and Discussion

The researchers grouped the codes into themes based on the Data-Driven Inquiry (DDI) framework. These themes represent the products of three of the four steps of the DDI process: assessment (collect evidence), reflection (make conclusions), and activities (take action). Although goal determination is a critical part of the DDI process, the blog entries did not focus on curriculum, which is likely why this theme did not emerge in the data analysis. Figure 1 displays a concept map that shows the connections among the blog entry categories and questions guiding the blog and how they are subordinate to major emergent themes.

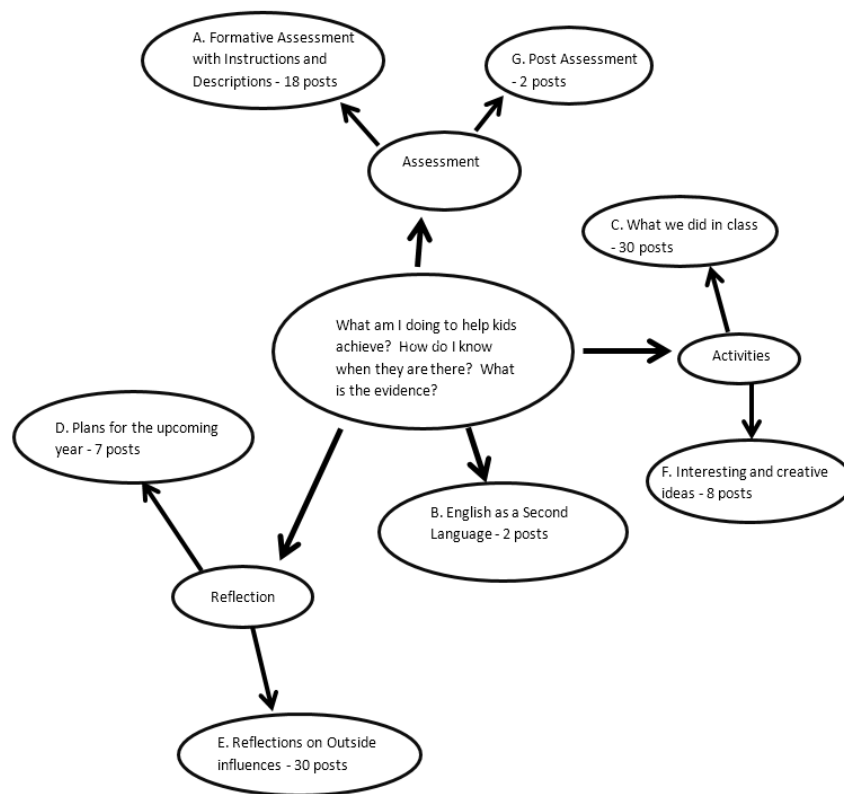


Figure 1: Concept map of blog questions, categories from coding, and emergent themes. A much more detailed concept map of block entries may be found in the supplemental information.

After examining 97 blog entries and the most frequently occurring categories, assessment emerged as a major theme. As such, the first author chose this as a focus for instructional improvement. This choice is strongly aligned with the Next Generation Science Standards, which are built upon the expectation that teachers must rely on assessment information that guides instruction (National Research Council, 2012). Specifically, he wanted to find a way to create formative and summative assessments in the most effective and efficient way possible. This was particularly timely since during the 2014-15 school year students were subjected to a significant increase in state mandated assessments. The challenge identified by the qualitative data analysis process was to improve assessment without increasing time spent on formal assessment activities. This warranted applying a new theoretical and practical lens to this initiative, and data driven inquiry served this need.

Viewing Results through the Lens of Data Driven Inquiry (DDI). Data from our own students is one of the best sources of information for us as teachers to make instructional decisions. This process is familiar to most teachers, and in a recent literature review has been described (Harshman & Yeziarski, in press). The process goes by multiple names in the

literature (U.S. Department of Education, 2008; 2011; Institute of Education Sciences, 2009; Calfee and Masuda, 1997) but will be referred to as Data Driven Inquiry (DDI). The steps are:

1. Determine goals.
2. Collect evidence.
3. Make conclusions.
4. Take Action (Harshman & Yezierski, in press).

Although there seems to be an abundance of research that shows that teachers should use data driven inquiry, the practicality of implementing the use of assessment to inform instruction on a day to day basis has not been well studied (Harshman & Yezierski, 2014, Harshman & Yezierski, in press). Although 20 out of 97 blog entries directly pertained to assessment, examination of the other categories (particularly what we did in class, plans for next year, and interesting and creative teaching ideas) in light of a DDI framework shows how assessment was pervasive in the data set. This is not surprising upon a re-examination of the questions guiding every blog entry:

1. What am I doing to help kids achieve?
2. How do I know when they are there?
3. What is the evidence?

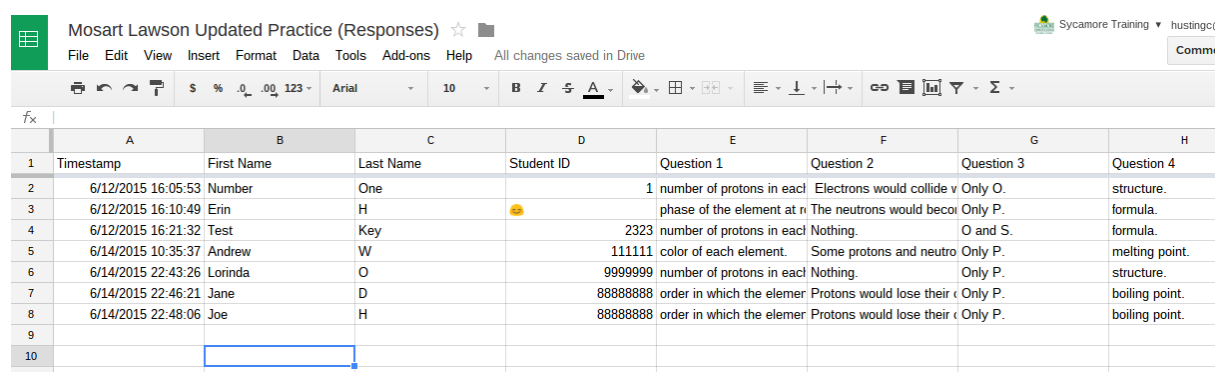
The similarities among the above questions and the components of DDI highlight how the first author has converged on assessment as a means to improve overall instruction. With the DDI process in mind, we have developed a tool to improve teaching and learning with an assessment focus.

Using Results: Developing a Tool to Meet Need Identified in Blog Analysis. The following considerations, which emerged from the analysis of the blog entries, guided the development of a tool to effectively and efficiently collect and use student data. The goal was to identify and use high quality assessments developed through research that could be delivered to students in such a way that scoring was quick and the collation/analysis of data enabled the teacher to put results to use immediately.

Assessment: First, the aim was to select multiple-choice chemistry questions that could evaluate conceptual knowledge (a higher level than just rote memorization). Questions that could provide information about a student's developmental level were particularly desirable, and it was essential that questions were aligned with the content and goals of the chemistry curriculum of the school and district. As mentioned before, it was also critical that the assessments were developed based on sound educational theory and evidence. The three tests used were the Lawson Classroom Test of Scientific Reasoning (LCTSR) (Lawson, 1978), the Chemistry 9-12 Misconception Oriented Standards Based Resource for Teachers (MOSART) from the Science Education Department of the Harvard-Smithsonian Center for

Astrophysics (Sadler, Coyle, Smith & Miller, 2006), and two-tiered chemistry questions that directly aligned with the curriculum to be taught.

Efficiency: The next step was to develop a method of delivery. Pen and paper, Scantrons and “bubbling in” answers is time consuming and cumbersome. A fast, effective, inexpensive method that would provide quick and reliable student data was needed. Permission was granted from test authors to develop electronic versions of the LCTSR and MOSART tests as long as they were secured. All of the questions from the two tests and the other questions developed were written into a single Google Form. Advantages to this method are the link can be easily and securely shared with students, it works on multiple devices, and the teacher can easily collect responses into a Google Sheet. Another advantage to the Google Form is that there is a free add-on called “Flubaroo” (<https://chrome.google.com/webstore/detail/flubaroo/mjkbmijfpphoabkogbdmdkolcnaenai?hl=en-US>). As students submit their data, the Flubaroo add-on grades it instantly. The graded sheet by Flubaroo provides a teacher with possible points, average points, and time of submission. An excerpt of student responses in the Google Sheet (Figure 2) is followed by a sample of the graded Flubaroo sheet (Figure 3).



| | A | B | C | D | E | F | G | H |
|----|--------------------|------------|-----------|------------|-----------------------------|------------------------------------|------------|----------------|
| 1 | Timestamp | First Name | Last Name | Student ID | Question 1 | Question 2 | Question 3 | Question 4 |
| 2 | 6/12/2015 16:05:53 | Erin | Number | One | 1 number of protons in each | Electrons would collide with | Only O. | structure. |
| 3 | 6/12/2015 16:10:49 | Erin | H | | phase of the element at n | The neutrons would become | Only P. | formula. |
| 4 | 6/12/2015 16:21:32 | Test | Key | 2323 | number of protons in each | Nothing. | O and S. | formula. |
| 5 | 6/14/2015 10:35:37 | Andrew | W | 111111 | color of each element. | Some protons and neutrons | Only P. | melting point. |
| 6 | 6/14/2015 22:43:26 | Lorinda | O | 9999999 | number of protons in each | Nothing. | Only P. | structure. |
| 7 | 6/14/2015 22:46:21 | Jane | D | 88888888 | order in which the elements | Protons would lose their electrons | Only P. | boiling point. |
| 8 | 6/14/2015 22:48:06 | Joe | H | 88888888 | order in which the elements | Protons would lose their electrons | Only P. | boiling point. |
| 9 | | | | | | | | |
| 10 | | | | | | | | |

Figure 2. Raw response data from two tests collected with a Google Form and displayed via Google Sheet.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N |
|----|---------------------------------|-------------------|------------------|-------------------|---------------------|----------------|------------------------|-----------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| 1 | Summary: | | | | | | | | | | | | | |
| 2 | Points Possible | 71 | | | | | | | | | | | | |
| 3 | Average Points | 16.4 | | | | | | | | | | | | |
| 4 | Counted Submissions | 5 | | | | | | | | | | | | |
| 5 | Number of Low Scoring Questions | 68 | | | | | | | | | | | | |
| 6 | | | | | | | | | | | | | | |
| 7 | Submission Time | First Name | Last Name | Student ID | Total Points | Percent | Times Submitted | Emailed Grade? | Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Question 6 |
| 8 | 6/12/2015 | Erin | H | 111111 | 16 | 22.53% | 1 | | 1 | 0 | 0 | 0 | 0 | 0 |
| 9 | 6/12/2015 | Andrew | W | 999999 | 18 | 25.35% | 1 | | 0 | 0 | 0 | 0 | 1 | 0 |
| 10 | 6/14/2015 | Lorinda | O | 888888 | 14 | 19.71% | 1 | | 0 | 0 | 0 | 0 | 0 | 0 |
| 11 | 6/14/2015 | Jane | D | | 19 | 26.76% | 1 | | 1 | 1 | 0 | 0 | 0 | 0 |
| 12 | 6/14/2015 | | | | 15 | 21.12% | 2 | | 0 | 0 | 0 | 0 | 0 | 0 |
| 13 | | | | | | | | | | | | | | |
| 14 | | | | | | | | | 40 | 20 | 0 | 20 | | |
| 15 | | | | | | | | | | | | | | |

Figure 3. Student responses and percentage correct as graded with a key by Fluberoo in the same Google Sheet as Figure 2.

The aforementioned online data collection process reduces time filling out Scantrons, grading them, and preparing to analyze the data by organizing it in a spreadsheet. It is possible to have the graded data before students leave the classroom.

Action. Although the speedy data collection improves efficiency in collecting the data, there needed to be a method to simplify data analysis. Two formatted Google Sheets were developed to automate data analysis. Raw data are cut, copied and pasted in one of the formatted sheets and the graded results from Flubaroo in the other. What makes these new sheets different is the conditional formatting. Figure 4 shows the Raw Data (similar to the information in Figure 2) that has been pasted into the new formatted sheet.

Practice Analysis with Links ☆

File Edit View Insert Format Data Tools Add-ons Help All changes saved in Drive

fx

| | A | B | C | D | E | F | G | H | I |
|----|-----------------|------------|----------------|------------|--------------------------------|-------------------------|--------------------------------|-------------------------|--------------------------------|
| 1 | | | | | MOSART | MOSART | MOSART | MOSART | MOSART |
| 2 | | | | | Periodic Table | Nuclear | Periodic Table | Organic | Reaction Rates |
| 3 | Timestamp | First Name | Last Name | Student ID | Question 1 | Question 2 | Question 3 | Question 4 | Question 5 |
| 4 | 6/12/2015 16:05 | Number | One | 1 | number of proton | Electrons would | Only O. | structure. | increase the cha |
| 5 | 6/12/2015 16:10 | Erin | Husting | | phase of the eler | The neutrons wo | Only P. | formula. | make it possible |
| 6 | 6/12/2015 16:21 | Test | Key | 2323 | number of proton | Nothing | O and S. | formula. | increase the cha |
| 7 | 6/14/2015 10:35 | Andrew | Wheatley | 111111 | color of each ele | Some protons ar | Only P. | melting point. | make it possible |
| 8 | 6/14/2015 22:43 | Lorinda | Ottaway | 9999999 | number of proton | Nothing. | Only P. | structure. | make it possible |
| 9 | 6/14/2015 22:46 | L | Ottaway | 88888888 | order in which th | Protons would lo | Only P. | boiling point. | make it possible |
| 10 | 6/14/2015 22:48 | L | Ottaway | 88888888 | order in which th | Protons would lo | Only P. | boiling point. | make it possible |
| 11 | 6/27/2015 11:05 | MOSART | Misconceptions | 222 | | | Only O. | structure. | make it possible |

Figure 4. Graded responses from Figure 2 subjected to conditional formatting in Google Sheet based on these criteria: Green = correct; yellow = popular misconception; white = incorrect.

The sheet shown in Figure 4 is conditionally formatted in a way that allows the teacher to quickly get a sense of multiple aspects of student performances to make teaching decisions based on the data. First, each question column labels the test type. The question is assigned to the type of assessment it came from such as “MOSART,” “LAWSON” or “Gen Chem.” The next cell down is a word or words that identifies the targeted content topic of that question. If you look carefully, you will notice that each of these words are links. There is a web site (<https://sites.google.com/site/simpleinqchem/home>) that has a page for each topic. The page provides inquiry activities that have been developed through researched chemistry education principles and teacher tested as discussed earlier. Some of these activities include Target Inquiry from Grand Valley State University and Miami University (TIMU), Process Oriented Guided Inquiry Learning (POGIL, 2012), Modeling and Inquiry labs from reputable sources. The goal was to only use materials that were vetted by educators and researchers. As in Figure 4, cells with student answers automatically fill with colors based on particular criteria. If it is white, the answer is wrong. If it is yellow, this answer has been identified by the authors of the MOSART test to be a popular misconception. If the answer is green, the response is correct.

To compile individual students’ results to examine question performance across an entire class, a Flubaroo formatted sheet was constructed using conditional formatting to color code levels of performance by question. The Flubaroo formatted graded sheet (Figure 5) has the question type, content, and link at the top of each question column. It also has a color-coded percentage correct at the bottom. If 75% or greater of the students in the class correctly answered the item, the percentage cell is green. If 40 to 74% of the students score correctly on the item, the cell with the percentage is yellow, and anything below 40% is red.

| | | | | | | | | | | | | |
|-----------------------------|--|----------------|------------|----------------|------------|----------------|------------|------------|------------|---------------|-------------|--|
| | | MOSART | MOSART | MOSART | MOSART | MOSART | MOSART | MOSART | MOSART | MOSART | MOSART | |
| | | Periodic Table | Nuclear | Periodic Table | Organic | Reaction Rates | Nuclear | Nuclear | Energy | Atomic Theory | Reactions | |
| Feedback for S FOR FLUBAROO | | Question 1 | Question 2 | Question 3 | Question 4 | Question 5 | Question 6 | Question 7 | Question 8 | Question 9 | Question 10 | |
| | | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | |
| | | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | |
| | | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | |
| | | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | |
| | | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | |
| Per Questions-> | | 100 | 20 | 0 | 20 | 20 | 40 | 40 | 20 | 20 | 0 | |

Figure 5. Conditional formatting by item to examine whole-class performance on MOSART items.

Implications

The assessment tool was implemented in the first author's class in fall 2015. Students took the assessment at the beginning of the year. Since students have experienced visiting a doctor, this analogy was useful in conveying the meaning of a "diagnostic" tool. The doctor or nurse first takes their temperature, blood pressure, height, and weight. These measurements are not necessarily seen as a test. One cannot really "pass" or "fail" any of these measurements, but they serve as diagnostic tools to guide the doctor in providing the best care possible. For the assessment, if everyone does well on the questions about the periodic table, the instructor should respect the students and spend less time addressing what they already know. Setting the proper tone and culture helped students take the assessment seriously.

Once students took the assessment, it was graded with Flubaroo and pasted into formatted google sheets. The DDI framework guided how formatted Flubaroo results were used. First, data from the Lawson test were examined. Heterogeneous student groups were established from their Lawson scores. It has been shown that once a student reaches a developmental milestone, the chances of them changing greatly within a school year are slim (Lawson, 1978). However, constant exposure to the other levels of reasoning might be helpful.

Next, the instructor focused on data from the MOSART and the Gen Chem questions. He spent limited class time on topics in which 75% or more of the class already demonstrated mastery. More instructional time was spent on the yellow questions, ones for which 40-74% of the students earned correct answers. These cutoffs are somewhat arbitrary; however, the green threshold (75%) reflects the percentage of students who demonstrate competency when the first author typically moves on in the curriculum. For anything in red, where less than 40% of the class scored correct answers, the instructor assumes that students know little to nothing and starts with basic ideas. It is encouraged that should any other teacher use this tool, s/he consider percentages that best meet the needs of his/her students.

An example is the instructor response to question 9 on the pre assessment. It is clear from the initial raw data that the majority of students do not know basic information about

isotopes. Only thirteen percent of students from all three classes answered correctly. The analysis shows that about half had no idea and half had popular misconceptions. This indicated to the instructor that the content, including common misconceptions, needed to be addressed. As part of the lesson on isotopes, the instructor used a POGIL activity. Students had to not only answer the questions but had verbally respond to instructor “checkpoints.”

Next the formatted raw data, particularly the MOSART questions, were examined to use the incorrect answer choices to guide instruction. Popular ideas aligned with misconceptions (formatted in yellow) should be treated differently than incorrect ones (formatted in white). As an example, there is a question that asks students if the distances change between the three atoms of a water molecule when water goes from a solid to a liquid. The correct answer is that the distance does not change. The answer that is a popular misconception is that the distance between the atoms in the molecule gets larger. In the commentary section of the MOSART, test authors note that the relationship between macroscopic changes and microscopic particles is evaluated by this item (Sadler et al., 2006). Clicking on the link “Molecular Shapes” at the top of the column for this question links to a site with several possible activities to teach molecular shapes. The action of the instructor should be to pick an activity that stresses the particulate nature of matter. Not only can the instructor clearly define goals and align them with the curriculum, assessments can be identified and added to the links.

When students are not meeting the objectives of the unit, the initial assessment data can be used on a student-by-student basis. For example, the Lawson test provides information about a student’s scientific reasoning ability. If a student is struggling with proportional reasoning in stoichiometry, the instructor may choose to narrow the goal and address some basic skill development around proportional reasoning. Other features of items may provide insights. The MOSART test, along with some of the author-developed general chemistry items, addresses specific levels of chemical knowledge (macroscopic, particulate, or symbolic). As such, the responses to the initial assessment questions may help the teacher identify problematic domains for the students. Since some questions stress one level more than another, the instructor may wish to identify different activities that address the same chemistry topic but emphasize specific levels.

The DDI process requires a careful examination of evidence of student knowledge, making a reasonable decision based on that evidence, and then carrying out instruction that is supported by current chemistry educational theory to be effective. The tool employs these processes and aims to synthesize high-quality assessments and curriculum materials into actions that respond to high-quality student data. We have evidenced the effectiveness of these novel tools by tracking how responses to available questions change from pre (first semester) to post (second semester) administrations. Incorporating the feedback from his own blogs and novel assessment tools, the teacher observed improvements with all but two of his students (Figure 6, left). These differences were statistically significant [$t = 12.3(44)$, $p < 0.001$, $d = 1.8$] and the change in quartiles are shown below (Figure 6, right). These results

indicate that students improved their scores dramatically on account of the instruction received, which contained the novel tools discussed here.

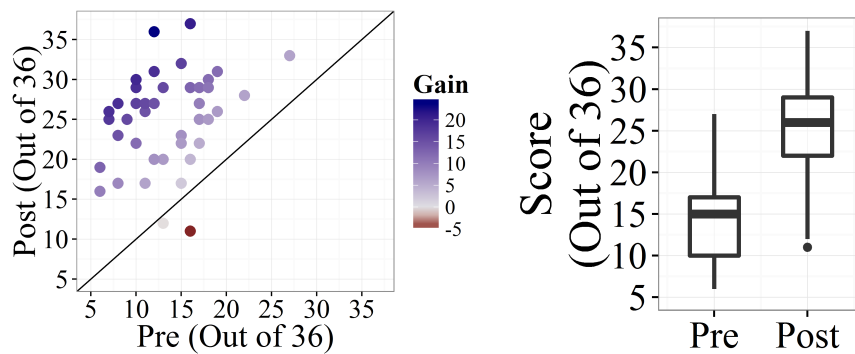


Figure 6: Pre and post scores for each individual student (left) and aggregated boxplots (right) demonstrate the large gains made by students during the year the tool was implemented.

Conclusion

Today's classroom teacher faces an avalanche of activity the minute they walk into the building. Students ask for help, daily emails, unexpected interruptions, labs that need to be prepared, meetings with parents, phone calls, endless grading of papers, unfunded state and federal mandates all add up to the point where, literally, it is now not uncommon to hear the words, "I don't even have time to use the bathroom." In the chorus of chaos, the simple act of documenting reflection forces a teacher to provide a small sliver of information about what is really important. Furthermore, systematic and valid research methods help to mine the blog information to help focus the teacher on what really matters when trying to help improve student learning.

We examined almost 100 blog entries by a single educator that always focused on the same three questions that dealt with student learning and evidence for it. Although several themes emerged from analyzing the blog entries, assessment was a central theme. Based on these findings and a DDI framework, a novel tool was created to collect data to inform and guide daily instruction in the most efficient and effective way possible. Using the tool helps and respects students' prior knowledge and leverages readily available technology. The research and practical outcomes presented here demonstrate how a classroom teacher and a researcher can effectively collaborate to combine teacher's knowledge of the students, classroom, and school environment during professional development to study and improve instruction while maintaining research quality and authenticity. Further research would be greatly beneficial in detailing the effectiveness of this particular strategy in specific topics.

As our scope was to introduce the idea of developing a tool that can be used to generally inform activities, studies that use this strategy and measure outcomes on specific topics would yield valuable insights on the pragmatic challenges, effectiveness, and tips on how to implement at a topic-specific level. As mentioned previously, these day-to-day details are scarcely found in relevant literature but crucial for successful adaptation of strategies such as the one proposed here.

Rather than ending the discussion here, we invite readers to use social media to view the ongoing project, examine its progress in real time, and converse with the first author/practitioner. The first author's progress can be tracked on his blog at <https://www.chemedx.org/blogs/chad-husting>. Additionally, inquiries are welcome from educators who wish to adopt the curricular activities embedded in the tool and/or discuss employing DDI strategies their classrooms.

Acknowledgments

We thank Claire Husting and Susan Husting for assistance with coding, and MOSART and Lawson test authors for granting permission for use. This material is based upon work supported by the National Science Foundation under Grant No. DRL-1118749.

About the Authors

Chad Husting teaches in a diverse suburban high school of about 1800 students in a suburb of Cincinnati, Ohio. Mr. Husting is a member of the American Association of Chemistry Teachers and has written for *Solutions*. He has a blog you can follow at <https://www.chemedx.org/blogs/chad-husting>. Email: hustingc@sycamoreschools.org

Jordan Harshman, Ph.D., is a postdoctoral research associate in the Department of Chemistry at the University of Nebraska – Lincoln. Dr. Harshman has worked with professional development for instructors both in his graduate experiences at Miami University and his current postdoctoral position.

Ellen Yeziarski, Ph.D., is a professor of chemistry at Miami University and director of the NSF-funded professional development and research project, Target Inquiry at Miami University. Dr. Yeziarski's research group investigates chemistry instruction in a variety of learning environments. Email: yeziers@miamioh.edu.

References

- Anderson, G. L., & Herr, K. (1999). The new paradigm wars: Is there room for rigorous practitioner knowledge in schools and universities?. *Educational researcher*, 28(5), 12-40.
- Bartlett, S., & Burton, D. (2006). Practitioner research or descriptions of classroom practice? A discussion of teachers investigating their classrooms. *Educational Action Research*, 14(3), 395-405.
- Boyer, E. L. (1990). *Scholarship revisited*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.
- Calfee, R. C., & Masuda, W. V. (1997). Classroom assessment as inquiry. In Phye, G. D. *Handbook of classroom assessment. Learning, adjustment, and achievement*. San Diego: Academic Press.
- Coppola, B. P., Banaszak Holl, M. M., & Karbstein, K. (2007). Closing the gap between interdisciplinary research and disciplinary teaching. *ACS Chemical Biology*, 2(8), 518-520.
- Dewey, J. (1933) *How We Think: A Restatement of the Relations of Reflective Thinking to the Educative Process* (2nd Revised Edition), Boston: D.C. Heath.
- Dewey, J. (1986, September). Experience and education. In *The Educational Forum* (Vol. 50, No. 3, pp. 241-252). Taylor & Francis Group.
- Freeman, D. (1998). *Doing teacher research: From inquiry to understanding*. Pacific Grove: Heinle & Heinle.
- Harshman, J., & Yeziarski, E. (2015). Guiding teaching with assessments: High school chemistry teachers' use of data-driven inquiry. *Chem. Educ. Res. Pract.* 16, 93-103.
- Harshman, J., & Yeziarski, E. (in press). Assessment data-driven inquiry: A review of how to use assessment results to inform chemistry teaching. *Science Educator*.
- Huberman, M. (1996). Focus on Research: Moving Mainstream: Taking a Closer Look at Teacher Research. *Language Arts*, 73(2), 124-140.
- Institute of Education Sciences, Hamilton, L., Halverson, R., Jackson, S., Manidnach, E., Supovitz, J., & Wayman, J. (2009). *Using student achievement data to support instructional decision making (NCEE 2009-4067)*. Washington, DC: National Center for Educational Evaluation and Regional Assistance, U.S. Department of Education.
- Jay, J. K., & Johnson, K. L. (2002). Capturing complexity: A typology of reflective practice for teacher education. *Teaching and Teacher Education*, 18(1), 73-85.
- Killeavy, M., & Moloney, A. (2010). Reflection in a social space: Can blogging support reflective practice for beginning teachers? *Teaching and Teacher Education*, 26(4), 1070-1076.
- Lawson, A. E. (1978). *Journal of Research in Science Teaching*, Vol. 15, 1, 11-24.
- Loving, C. C., Schroeder, C., Kang, R., Shimek, C., & Herbert, B. (2007). Blogs: Enhancing links in a professional learning community of science and mathematics teachers. *Contemporary Issues in Technology and Teacher Education*, 7(3), 178-198.
- Luik, P., Voltri, O., Taimalu, M., & Kalk, K. (2011). On the use of student teacher blogs during teaching practice. *Procedia-Social and Behavioral Sciences*, 11, 165-169.

- Lytle, S. L. & Cochran-Smith, M. (1992). Teacher Research as a Way of Knowing. *Harvard Educational Review* 62 (4), Harvard University.
- Meijer, P. C., Oolbekkink, H. W., Meirink, J. A., & Lockhorst, D. (2013). Teacher research in secondary education: Effects on teachers' professional and school development, and issues of quality. *International Journal of Educational Research*, 57, 39-50.
- National Research Council (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- POGIL Activities for High School Chemistry (2012). Trout, L., ed. Batavia, IL: Flinn Scientific.
- Ray, B. B., & Hocutt, M. M. (2006). Teacher-created, teacher-centered weblogs: Perceptions and practices. *Journal of Computing in Teacher Education*, 23(1), 11-18.
- Sadler, P. M., Coyle, H. P., Cook-Smith, N., & Miller, J. L. (2006). MOSART: Misconceptions-oriented standards-based assessment resources for teachers. *Cambridge, MA: Harvard College*. Retrieved from <https://www.cfa.harvard.edu/smgphp/mosart/>.
- Saka, E. (2008, May). Blogging as a research tool for ethnographic fieldwork. In *EASA Media Anthropology Network e-seminar*.
- Tobias, S., & Baffert, A. (2010). *Science Teaching as a Profession: Why it Isn't, How it Could Be*. Arlington, VA: NSTA Press.
- U.S. Department of Education, Means, B., Chen, E., DeBarger, A., & Padilla, C. (2011). *Teachers' ability to use data to inform instruction: Challenges and supports*. Office of Planning, Evaluation and Policy Development. Washington, DC.
- U.S. Department of Education, Gallagher, L., Means, B., & Padilla, C. (2008). *Teachers' use of student data systems to improve instruction, 2005 to 2007*. Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Services. Washington, DC.
- Wopereis, I. G., Sloep, P. B., & Poortman, S. H. (2010). Weblogs as instruments for reflection on action in teacher education. *Interactive Learning Environments*, 18(3), 245-261.
- Yost, D. S., Sentner, S. M., & Forlenza-Bailey, A. (2000). An examination of the construct of critical reflection: Implications for teacher education programming in the 21st century. *Journal of Teacher Education*, 51(1), 39-39.
- Zeichner, K. M., & Noffke, S. E. (2001). Practitioner research. *Handbook of Research on Teaching*, 4, 298-330.

Appendix A: Final Category Names, Descriptions, and Frequencies of the Occurrence in the Data Set

| | Name | Description | Frequency |
|---|---|---|------------------|
| A | Formative assessment with instructions and descriptions | This category attempts to describe some type of formative assessment and then provides instructions on how to carry this out in the classroom. Often this is done at the beginning of a unit or assignment and usually not for a grade. It is then used to inform my instruction. | 18 |
| B | English as a second language | ESL teaching techniques can help all students. These techniques use other methods besides of instruction and assessment besides words. It is helpful with students who either struggle with English or have been raised speaking a foreign language. | 2 |
| C | What we did in class | These are actual labs and activities we are doing or are going to try to do in class. Most of these center around inquiry. A theme with these labs is that many are performance assessment. Students ultimately should be able to predict an outcome or some type of end measurement. It is not something they could look up on Google. | 30 |
| D | Plans for next year | After reflecting and having a set of experiences, these are plans and “big ideas” that I hope to guide my instruction in the future current year or following year. I have had a set of experiences and am trying to look at guiding principles to guide my future teaching. | 7 |
| E | Reflections on outside influences | These entries are not about classroom incidences. These are about events that have happened outside the classroom (classes I am taking, books I have read, talks I have heard). Furthermore, I have reflected on these events and this could influence my teaching. | 30 |
| F | Interesting and creative teaching ideas: | These are ideas that I have gotten from outside sources. I hope to use or have used that seem like student-centered creative ideas. They focus on hands on manipulatives that students can do as labs, projects or in some cases, manipulatives that I can use as a type of assessment. | 8 |
| G | Post assessment | This is an assessment that is at the end of an activity and is similar to, but not as detailed, as a formal summative assessment. It is fast, simple and at the end of an activity. | 2 |

DATA ON THE DATA: A METHOD FOR IMPROVING THE FIDELITY OF OFFICE DISCIPLINE REFERRAL COMPLETION

Paul Gavoni

Indian River State College

W. Alex Edmonds

Nova Southeastern University

Tom D. Kennedy

Nova Southeastern University

Tom Gollery

Southeastern University

Abstract As behavior problems increase within public schools, attempts at intervention are most effective when accurate data are available for planning and assessment. The current investigation examined the relationship between predictive variables impacting the completion of office discipline referrals (ODRs); moreover, this investigation sought to demonstrate the effectiveness of a coding system developed by the researchers as a method of measuring and improving the fidelity of the completion of ODRs in a school utilizing School Wide Positive Behavior Support (SWPBS). These findings shed light on the “coding” process as a potentially viable data source.

Keywords: behavior management, discipline, referrals, positive behavior support, office discipline referrals

Introduction

Safe and orderly schools and classrooms provide structure that is critical to improving student achievement. Student misbehavior is disruptive to the learning environment and impacts student performance as critical instruction time provided by teachers is redirected toward behavior management (Musti-Rao & Haydon, 2011). Wreting (2010) suggested “there may be no greater hurdle in public schools today than that presented by students who exhibit challenging behavior” (p. 48).

As a result, the need for school-wide proactive and systematic behavior management systems such as school wide positive behavior support (SWPBS) has emerged (Childs, Kincaid, George, & Gage, 2015). Thus, as part of a continued improvement process for promoting a positive learning environment, schools must establish consistent practices of collecting and monitoring data. As such, accurate office discipline referrals (ODR) data is a key component to the successful application of SWPBS interventions (Kincaid, Childs, Wallace, & Blasem 2007).

Theoretical Perspective

Constructs fundamental to School Wide Positive Behavior Support (SWPBS) and that are known to influence a teacher’s completion of a formal written office discipline referral (ODR) are grounded in the theory of applied behavior analysis (ABA; Skinner, 1953). Behavior analysis is used to improve socially significant behaviors (Morris, Smith, & Altus, 2005).

SWPBS is systematic framework for teaching and reinforcing adaptive behavior within the school setting (Flannery, Fenning, Kato, & McIntosh, 2014). Within the SWPBS system, evidenced-based strategies and interventions are selected based on student outcomes within a three-tiered framework to problem solve and reduce the frequency of maladaptive behavior (Bradshaw, Mitchell, & Leaf, 2010).

Literature Review

Office Discipline Referral. Office discipline referrals (ODRs) are a typical source for measurement of school-wide behavior and the impact of primary interventions in schools utilizing SWPBS (Bradshaw et al., 2010). ODRs are used to report student violations to administration (Irvin, Horner, Ingram, Todd, Sugai, Sampson, & Boland, 2006). As part of the SWPBS discipline procedures, teachers are required to document an ODR for each instance a student is removed from the classroom as a result of previously defined misbehavior that has become unmanageable in the classroom, or what SWPBS terms office managed behavior (Irvin et al., 2006).

Researchers have found that students who receive an ODR lose approximately 45 minutes of instruction per referral (Muscott, Mann, & Lebrun, 2008). However, research has shown

that schools that use ODRs to implement SWPBS interventions consistently lower student misbehavior (Childs, Kincaid, George, & Gage, 2015). Thus, ODR data can play a critical role in identification and remediation of issues which impacts student achievement.

Some researchers recognize ODRs as an efficient and increasingly standard source of data collection for monitoring SWPBS (Flannery, Fenning, Kato, & Bohanan, 2011). However, other researchers have cited a variety of limitations that result either in the underestimation or overestimation of remediation (i.e. ineffective data entry, inconsistent ODR submission, teacher bias and/or tolerance of misbehavior, over-reporting of minorities; etc.) and suggest additional data sources should be used (Childs, Kincaid, George, & Gage, 2015; Martinez, McMahon, & Tregor, 2015; Rusby, Taylor, & Foster, 2007; Sugai, Sprague, Horner, Walker, & Hill, 2000).

School Wide Positive Behavior Support. School wide positive behavior support (SWPBS) intervention program puts into place a support system to aid behavioral change of students who display negative behaviors. There is a wealth of research that shows the use of SWPBS has a positive impact on reducing problem behavior in the classroom and in increasing student achievement (McIntosh, Bennett, & Price 2011).

There is evidence that supports the validity of Benchmarks of Quality (BOQ) as a fidelity measure based on a SWPBS teams estimation of implementation (Childs, George, & Kincaid, 2011), a primary measure for assessing SWPBS and the impact on student behavior continues to be the use of ODRs. However, while touted as an efficient data source, Sugai et al. (2000) cautioned that the accuracy of the ODR collection and monitoring systems is critical to the application of ODRs as an effective measure for informing school-based intervention.

Research Questions

The purpose of the study was to search for workable solutions to more accurately measure student behavioral issues and the impact of interventions within the school setting. In addition, this investigation sought to examine the relationship between predictive variables impacting the completion of office discipline referrals. The following research questions drive the study:

1. What is the frequency of teachers who formally document office discipline referral forms?
2. What is the average ratio or percentage of teacher submissions of a behavioral infraction code to actual ODR written documentation?
3. To what degree does one's teaching experience, training hours in behavioral modification and ease of coding (covariates) correlate with the teachers' documentation of student's misbehavior in written form using the office discipline referral form?

4. Which covariates (teaching experience, training hours and/or ease of coding) is statistically significant in predicting a participant's likelihood of formally documenting an ODR.

Methodology

Design. A participatory action-research approach was utilized for this study. This type of design provides the framework for applying a systematic approach in order to study a problem when seeking to augment performance within the context of specific educational settings. However, participatory action research is small scale and not generalizable even though it is often used by educators to identify a problem, collect data, analyze the data, and develop an action plan to solve the problem (Stringer, 2013).

Setting and Participants. The study took place in a mid-sized central Florida Title I elementary school in a predominantly low-income area. The participants were 40 female classroom teachers. The average age of participants was 38.4 years (Range: 22–63), the average years teaching was 10.2 (Range: 1–43), the average hours of university coursework enrolled in by participants that pertained to behavior management, 96.0 (Range: 0–384), and the average hours of training with which participants were credited post-university graduation pertaining to behavior management was 23.6 (Range: 3–150).

Training Received. At the beginning and middle of the school year, staff (teachers, office secretaries, counselors, school psychologist, and administrators) received training about both 1) behavioral management and 2) how to request administrator assistance. Calling for an administrator was common practice within the targeted school and most elementary schools within the district when teachers had student behavioral problems that could not be handled. During the behavioral management portion, staff received information about the importance of effective classroom management as well as the range of differentiated strategies that could be used for responding to student misbehavior, such as praising students on task and ignoring inconsequential behavior (Reinke, Herman, & Sprick, 2011).

During the session where teachers learned to call for administrator assistance, they learned about the SWPBS plan, as well as: (a) school-wide expectations, (b) rules, (c) reward systems, and (d) discipline referral process and procedures. Teachers were taught the only purpose of “calling a code” was to obtain assistance with managing severe or repeated instances of disruptive behavior, or to obtain assistance with medical emergencies; it was not to be used as a threat to the student (e.g., “Stop that or else I will call a code”). In addition, teachers were taught the coding procedures developed by the investigators and told to request assistance from administrators when behavioral problems escalated to office managed behaviors by using the following three codes:

1. Code 1: Continuous Aggression or Self-Injury or High Intensity Property Destruction. Identified an emergency situation where the student presented an immediate physical danger to self or others.

2. Code 2: Continuous High Magnitude Disruption or Low Intensity Property Destruction. Identified a situation where a student was continuously disrupting the classroom environment. Lower magnitude behaviors were identified as talking out while high magnitude behaviors was yelling and physical contact.
3. Code 3: Out of Assigned Area. Identified a situation where a student intentionally left their assigned area or had taken flight from the school.

Collecting Data. The office discipline referrals (ODRs) were the unit of analysis to determine the predictive ability of variables influencing a teacher's completion of the ODR forms. A structured survey recording approach was used to obtain data from the ODRs and was examined using the three predictor variables: (a) years teaching, (b) codes; code/ODR ratio, and (c) training hours focused on student behavior management and discipline processes.

The codes, talked about in the training section above, were only necessary for the office to know how to manage the "call for help." All codes were recorded (i.e., time, location, code, student, staff calling the code, staff responding, and the staff logging the data) by a designated dispatcher into an Excel database. The dispatcher, who was the front desk clerk who routinely communicated requests for assistance from the teacher to the administrator, was trained to obtain verbal confirmation that an administrator was responding.

To ensure fidelity with the coding procedures, the primary investigator monitored the radio, verified the code was logged once assistance was requested, provided constructive feedback to correct errors, and verbally praised the dispatcher for implementing the procedure correctly. This process was repeated for the backup dispatcher, and then faded to random checks twice weekly. Since office managed behavior required the completion of an ODR as a critical element of SWPBS, each code called was required to be accompanied by a referral. In addition, teachers were responsible for contacting the parent/guardian and documenting the contact on their classroom parent contact log as part of the school's and district's expectation.

Data Analysis. To review and analyze data, the investigators used existing ODR data from the district database input by teachers, code data input into the school's Excel program by the front desk clerk, and data from the survey distributed to teachers at an all-staff meeting by the primary investigator. The ODR data source utilized the Skyward System, a licensed private software, to examine data. The Skyward discipline database was customized to allow schools the ability to capture and graph aggregate data necessary for assessing and problem solving within a SWPBS structure. ODR data aggregated within the Skyward System allowed users to track major behavior infractions and resulting discipline referrals weekly, monthly, and yearly by individual students, targeted groups, or school-wide.

The first two research questions were addressed by an array of descriptive statistics for comparative purposes. Measures of central tendency, variability, and percentages were utilized to illustrate the nature and distribution. Research question three was analyzed by the Pearson Product-Moment Correlation Coefficient (r) employed to determine the strength of relationship between the predictor variables and the criterion variable. The fourth question, which focused on the predictive abilities of the predictors on the criterion, was determined by a multiple regression.

Results

Research Q1 and Q2. The ratio of ODR written documentation instances per participant registering of behavioral infraction codes was 34%. In other words, approximately one third of teachers calling in codes to the office also completed the ODR. The range of ratios with the 40 participants was 0-100%, with the most occurring at 0% ($n = 14$; 35% of the total participant group). Moreover, the variability of code registration data to subsequent written ODR data among the 40 participants was statistically significant ($t = 5.45$; $p < .000$).

Research Q3. The correlation coefficients for the predictor variables of “Years Teaching” and “Hours of Training” represent very weak, inverse relationships with the criterion variable (Written ODR). Moreover, both correlations were not statistically significant ($p = .229$; $p = .402$). The relationship between codes formally registered by participants and the criterion variable of written ODR is moderately high ($r = .535$), and statistically significant ($p < .000$).

Table 1: Matrix of Relationships Among Three Variables

| Correlations | | ODR write-up | Teaching years | Codes registered |
|------------------------|----------|-----------------|-------------------|---------------------|
| Pearson correlation | ODR | 1.000 | -.121 | .535 |
| | Teach | -.121 | 1.000 | -.106 |
| | Codes | .535 | -.106 | 1.000 |
| | Training | -.041 | .178 | .034 |
| Sig. (1-tailed) | ODR | . | .229 | .000 |
| | Teach | .229 | . | .258 |
| | Codes | .000 | .258 | |
| | Training | .402 | .136 | .417 |

In general, when two variables under investigation yield correlation coefficients of .70 and beyond, it is thought to complicate the precise measurement and predictive power of each individual variable in the Multiple Linear Regression analysis. Collinearity statistics (VIF Mean = 1.032; Tolerance Mean = .969) affirm the avoidance of collinearity issues that may have affected the current model of multiple predictors (Menard, 1995). Table 1 shows the correlation coefficient for the three variables under investigation.

Research Q4. Participant formal registration of a behavioral infraction code was the most robust and only statistically significant predictor of the likelihood of formal written ODR documentation ($B = .072$; $t = 3.76$; $p < .001$). The other two predictor variables (Years of Teaching; Training Hours) were not statistically significant predictors of the likelihood of participant written documentation of an ODR ($B = -.019$; $t = -.390$; $p = .699$ and $B = -.006$; $t = -.341$; $p = .735$).

The focus of this investigation examined the predictive abilities of three predictor variables thought to impact a teacher's likelihood of documenting an office discipline referral (ODR) in written form. The results revealed that follow-through in the written documentation process were largely inconsistent. Moreover, the variability of participant ratio of coding to written ODR was statistically significant ($p < .000$).

The consequence of this lack of action on the part of many participants is, at the very least, counter-productive to the process of initiating meaningful intervention as a result of unreliable data. It was the researchers' contention that the closer the ratio of formal written ODRs is to the registered behavioral infraction codes, the greater the probability that timely and effectual remediation of respective issues of misconduct will be enacted.

However, participants manifested a ratio of slightly over one third (34%). The data shows the need for formalized follow-up training on the importance of teachers' consistent use of formal written ODR documentation in the wake of registering a behavioral infraction code to enhance the fidelity between the two actions. Ideally, the ratio should be as close to 100% as possible if meaningful intervention and positive change is to take place in the redirection and amelioration of maladaptive behavior.

To test the ratio of ODRs to codes, a combination of Binary Logistic Regression and ROC Curve statistical techniques was used, in a post hoc fashion, to assess the predictive prowess of the ratio with regard to the likelihood of a teacher's registration of a formally written ODR. Results from this exploratory post hoc analysis confirmed the notion that the ratio of registered infraction codes to actual written ODR may represent an even more critical datum than the actual registration of the code itself. This information is critical for administrative staff to understand the relationship between these variables and ultimately how it affects the classroom dynamics. The "Ratio" is a statistically significant predictor of a teacher's probability of formally writing the ODR ($p < .001$; $R^2 = .81$). From an "Odds Ratio"

or “Exp(B) perspective with regard to the Binary Logistic regression analysis, a one percentage increase in the ratio percentage increases the probability that a teacher will formally write an ODR by 10%. This in-turn increases the probability that the behavioral incident will be addressed.

The Code/Written ODR ratio represented the only statistically significant predictor of a teacher’s likelihood to register a written ODR ($p < .002$) when compared together in the regression model (model = $p < .000$; $R^2 = .828$; Hosmer & Lemeshow = .723).

Receiver Operation Characteristics (ROC) Curve Analysis. The ROC curve analysis is performed when a researcher wishes to heighten the “sensitivity” and “specificity” of prediction. In the current investigation, ROC curve analysis confirmed the findings of the binary logistic analysis, further asserting the predictive prowess of the ratio of code registration and written ODR with regard to a teacher’s likelihood of committing to the writing of a formal ODR (Area Under the Curve/AUC= .973; $p < .000$ -all other predictor variables AUC ranging from .515 to .615).

Discussion and Implications

The major area of emphasis of the current study involved the identification of predictor variables (covariates) that might accurately predict the likelihood of a teacher completing formal written documentation of an ODR. It was the researchers’ contention that once variables are identified that may serve as accurate predictors of the written ODR considerable effort could then be devoted to professional development activity. Years of teaching, hours of post-university training in the area of behavior management, and teacher registration of formal behavioral infraction codes were selected as the prime predictor variables (covariates) in the study. An additional variable thought to be a possible predictor of written documentation of an ODR was the age of the teacher. However, to ensure clarity and uniqueness of prediction in the study, we chose to omit the variable “age” due to its high level of relationship to the variable of years in the teaching profession ($r > .70$).

Results of the regression analysis conducted on the three selected predictor variables highlighted the superior predictive ability of a teacher’s registration of a behavioral infraction code ($B = .072$; $t = 3.762$; $p < .001$). The relationship between the coding and the actual result of a written ODR was positive, approaching “strong” ($r = .54$, $p < .001$). Years of teaching and hours of post-university training included in the study for predictive purposes manifested a very weak relationship with the written ODR, thus were not very effectual as predictors of a teacher’s likelihood of producing a written ODR. The Multiple Linear Regression analysis results, along with relevant descriptive data obtained in the current study, strengthen the case for the importance of the role that the actual registering of behavioral infraction codes plays in the series of events that precede actual intervention and, in turn, bring clarity in defining the process itself. It would appear, from the results of the current study, that professional development energies would be best dedicated to instructing and nurturing teacher awareness and discipline in committing to greater levels of

fidelity in registering behavioral infraction codes and following that action closely with written documentation of an ODR.

Limitations

There were several limitations inherent in this investigation. One limitation concerned self-reporting in relation to college credits and training hours earned in classroom or behavior management. If participants did not provide accurate information in regard to their training history, outcomes may not be reflective of the actual relation between the covariates. To minimize this threat, teachers were provided their documented training hours logged in the districts data warehouse. Additionally, the quality of professional development and its design – such as, job-embedded, conducted over a period of time, and including practice, observations, and feedback – or “one-shot, sit-and-get” presentations – were not part of the study. Not all “training” for teachers on behavior management techniques is equally useful.

Another limitation deals with historic school goals in relation to reducing misbehavior that may have impacted the results. The school of study had been provided additional behavioral supports like the development and training in calling codes that likely impacted the fidelity of SWPBS implementation. In addition, teachers were aware that ODRs were closely measured through the coding process, which may have impacted teacher behavior in regard to ODR documentation.

Conclusion

Given the increase in evidenced-based practices within public education over the last decade, valuable discipline data continue to be aggregated and analyzed to support student needs and growth. As a result, the effectiveness of SWPBS, when implemented with fidelity (McIntosh, Bennett, & Price 2011), cannot be repudiated as objective data are used to drive decision-making processes (Bradshaw et al., 2010); however, additional research is needed to increase the variety of accessible measures used within the SWPBS systems. Martinez, McMahon, and Treger (2015) recommended increased research to examine the differences in ODR rates by teacher as related to teacher training and other contextual influences. This investigation, while seeking to bridge the gap in literature, has highlighted a process beyond the BOQ (Childs, George, & Kincaid, 2011) as a measure of the fidelity regarding the documentation of ODRs. Considering the importance of data assessment and evaluation measures to district-based administrators, principals, teachers, and parents, increasing the breadth of data resources offers an opportunity for schools and districts to continue refining their processes by embedding the coding process utilized within this investigation. While research exists that demonstrates the modest validity and reliability of the ODR process (Irvin et al., 2006), a question arises as to whether this process is consistently the best measure to support the SWPBS process. Use of the coding process has the potential to increase the reliability of ODR data. As of 2016, no research has investigated predictors of ODR completion or demonstrated a potential source of data and measure of fidelity as reflected in this study through the coding process.

Lessons Learned. After the investigation, following the premise of action research, code data was used by the assistant principal to measure ODR documentation and prompt teachers to completion when necessary. Specifically, teachers whose ODR data paralleled the code data were sent emails praising them for inputting the ODR; teachers who had a code recorded but did not have a corresponding referral were prompted through an email to input an ODR. This process resulted in a 100% completion of ODRs as measured by the recorded codes. Although the immediate impact on the completion of ODRs is outstanding, to ascertain the long-term implications of this, a follow-up data analysis is required.

The current investigation was a preliminary endeavor to augment the existing body of SWPBS literature for the purpose of informing assessment and evaluation methods of practitioners who implement SWPBS. Research on school-wide behavior management has demonstrated that students who receive an ODR lose approximately 45 minutes of instruction per referral (Muscott, Mann, & Lebrun, 2008). The correlation between lost academic time and behavioral issues deserves increased focus in an environment where the need for reducing the achievement gap is at the vanguard of educational initiatives. Since, ideally, the ratio between ODRs and coding process demonstrated in this investigation should approximate 100% as much as possible, use of codes within schools where teachers call for assistance or “student pick-ups” when office managed behavior occurs may provide a very efficient yet highly meaningful data source to support the SWPBS process.

As a result of the findings and feedback model, we have developed the key “take-aways” that will be used in the school:

1. Teachers will be trained to call codes for students who meet office managed behavior.
2. Office staff or those who acknowledge calls for assistance will be trained to log data in a database.
3. Administrators or designees will be taught to compare the code data with ODR data.
4. Administrators or designees will be taught to praise or prompt ODR input based on the ODR to code ratio.

Future Research. Future research efforts relevant to the topic of the current study should be engaged in further development of a comprehensive profile of behavioral infraction coding and written documentation of an ODR. One suggestion for continued research efforts might center on a broader, more stratified sample of participants. The current study was comprised of participants teaching at Title I schools. Students enrolled at Title I schools, on average, tend to manifest more behavioral infractions than regular general education environments and are not reflective of a “normal” distribution.

Another research path that may be taken in light of the importance of behavioral infraction coding could involve the establishment of internal reliability benchmarks and measures of

the accuracy of the coding itself. The magnitude and variability of coding is an area that was not addressed in the current study, but warrants attention from researchers in the time ahead.

From a phenomenological perspective, future research might address a possible change in participant behavior in the coding/ODR process related to formal observation by researchers. Is reactivity a consideration in establishing the validity of participant response, and if so, does it tend to increase or decrease participant response in the coding/ODR process?

Future investigators might seek to use qualitative data by interviewing those teachers who did write formal discipline referrals and those teachers who “called the code” but who did not write formal discipline referrals. This insightful qualitative data may potentially help researchers make greater sense of their quantitative data.

Lastly, and perhaps the most meaningful of research paths that might be taken on the topic of the current study would involve an investigation of predictor variables that may have a mediating or moderating effect in regression analysis upon the likelihood of a teacher’s engaging in the written documentation of an ODR. The focus of the current investigation was the initial identification of predictor variables that might serve as accurate predictors of that likelihood. However, it would be helpful to determine if a specific predictor variable or variety of variables have been playing a significant role in the written documentation of an ODR or the decision not to do so.

About the Authors

Paul F. Gavoni, EdD, BCABA, is an expert in human performance and organizational leadership who has over 20 years of experience in education and human services primarily focused on bringing out the best in people. As an Adjunct Professor at Indian River State College and Florida Atlantic University, Paul is passionate about engaging and empowering his students by helping them develop, analyze, and successfully apply knowledge. Paul holds a Doctorate of Education with a concentration in Organizational Leadership from NSU, a Specialist of Education with a concentration in Educational Leadership from NSU, and a Masters of Social Work with a concentration in Children and Families from Barry University. Beyond his work in education and human services, Paul is also a highly respected striking coach in combat sports. In 1992, Paul began boxing in South Florida and went on to win a Florida Golden Gloves Heavyweight Title in 1998. Since then, Coach "Paulie Gloves," as he is known in the MMA community, has trained many champions and UFC vets using technologies rooted in the behavioral sciences. A featured coach in the book [*Beast: Blood, Struggle, and Dreams at the Heart of Mixed Martial*](#)

[Arts](#), Coach Paulie is also an author who has written for online magazines such as [Scifighting](#), [Last Word on Sports](#), [MMASucka](#), and [Bloody Elbow](#).

W. Alex Edmonds, PhD, BCB, is currently an associate professor of research at Nova Southeastern University in Davie, Florida. He graduated from Florida State University and received his doctoral degree in Educational Psychology with a minor in Statistics and Measurement. Over the years, Dr. Edmonds has applied his knowledge of research design, measurement, and assessment in both field and laboratory examinations. He has published extensively in a variety of areas such as psychophysiology and educational psychology. His primary interest revolves around applying unique methodological and statistical techniques as a means to exploring the relationship between emotions, physiology and performance in a variety of domains. He also has over 10 years of experience in applying biofeedback in field and research applications. While in graduate school, he conducted his field work with the track and field team at Florida State and started using biofeedback for research and practice during this time. He has utilized biofeedback extensively with various types of athletes for performance enhancement, as well as stress-regulation techniques for individuals with type two diabetes and pain management for patients suffering from chronic pain. Dr. Edmonds is certified through the Biofeedback Certification International Alliance in general biofeedback.

Tom D. Kennedy, PhD, BCB, is currently an associate professor of research at Nova Southeastern University in Davie, Florida. He received his PhD in Counseling Psychology from the University of Miami and his MA in Clinical Psychology from Southern Methodist University. He has been a faculty member at NSU for five years where he has developed and taught research and statistics courses. He is the Institutional Review Board representative for the School of Education. His clinical experience consists of providing neuropsychological assessments, behavioral medicine interventions, and group therapy in inpatient and outpatient settings, including the following: the University of Texas Southwestern Medical Center, University of Miami Mailman Center for Child Development, Jackson Memorial Hospital, University of Miami Institute for Individual and Family Therapy, and the Dallas County Jail. His research interests include two convergent tracks, one focusing on at-risk children and adolescents and the other exploring biofeedback and other complementary and alternative medicine interventions. He has been the recipient of multiple grants and provides evaluation and grant writing services for various organizations in south Florida. He has published and presented in the areas of juvenile crime, psychophysiology, and research methodology. Dr. Kennedy is a licensed psychologist PY 8307 and is certified through the Biofeedback Certification International Alliance in general biofeedback and maintains a small private practice for children and adolescence.

Tom Gollery, EdD received an MA Degree in Counseling from Appalachian State University, and an EdD emphasizing both special education and educational leadership from the University of Florida. With more than 25 years as a school administrator, Tom is currently a professor of graduate studies in the College of Education at Southeastern University. Tom primarily functions as a quantitative research methodologist for the doctoral program in the College of Education at Southeastern. His research interests include: assessment of student achievement; educational program evaluation; institutional effectiveness; and interdisciplinary research.

References

- Bradshaw, C., Mitchell, M., & Leaf, P. (2010). Examining the effects of school wide positive behavioral intervention and supports on student outcomes results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Interventions*, 12(3), 133–149.
- Childs, K., George, H. & Kincaid, D. (2011). Stability in Variant Administration Methods of the School-Wide PBS Benchmarks of Quality (BoQ). *OSEP Technical Assistance Center on Positive Behavioral Interventions*. Retrieved from www.pbis.org.
- Childs, K., Kincaid, D., George, H., & Gage, N. (2015). The relationship between school-wide implementation of positive behavior intervention and supports and student outcomes. *Journal of Positive Interventions*, 11(1). Retrieved from <http://dx.doi.org/doi:10.1177/1098300715590398>.
- Flannery, K., Fenning, P., Kato, M., & Bohanan, H. (2011). A Descriptive Study of Office Disciplinary Referrals in High Schools. *Journal of Emotional and Behavioral Disorders*, 21(2) 138–149.
- Flannery, K., Fenning, P., Kato, M., & McIntosh, K. (2014). Effects of school-wide positive behavioral interventions and supports and fidelity of implementation on problem behavior in high schools. *School Psychology Quarterly*, 29(2), 111–124.
- Irvin, L., Horner, R., Ingram, K., Todd, A., Sugai, G., Sampson, N., & Boland, J. (2006). Using office discipline referral data for decision making about student behavior in elementary and middle schools: An empirical evaluation of validity. *Journal of Positive Behavior Interventions*, 8(1), 10–23.
- Kincaid, D., Childs, K., Wallace, F., & Blase, F. (2007). Identifying barriers and facilitators in implementing school-wide positive behavior support. *Journal of Positive Behavior Interventions*, 9(3), 174–184.
- Martinez, A., McMahon, S., Tregor, S. (2015). Individual and school-level predictors of student office disciplinary referrals *Journal of Emotional and Behavioral Disorders* 1-12. Retrieved from <http://dx.doi.org/1063426615588289>.
- McIntosh, K., Bennett, J., & Price, K. (2011). Evaluation of social and academic effects of school-wide Positive behaviour support in a Canadian school district. *Exceptionality Education International*, 21, 46-60. Retrieved from <http://ir.lib.uwo.ca/eei/vol21/iss1/5>.
- Menard, S. W. (1995). *Applied logistic regression analysis*. Thousand Oaks, CA: Sage Publications.
- Morris, E., Smith, N., & Altus, D. (2005). B. F. Skinner's contribution to applied behavior analysis. *The Behavior Analyst*, 28(2), 99–131.
- Muscott, H., Mann, E., & Lebrun, M. (2008). Positive behavioral interventions and supports in New Hampshire: Effects of large-scale implementation of schoolwide positive behavior support on student discipline and academic achievement. *Journal of Positive Behavior Interventions*, 10(3), 190–205.
- Musi-Rao, S. & Haydon, T. (2011). Strategies to increase behavior-specific teacher praise in an inclusive environment. *Intervention in School and Clinic*, 47(2), 91–97.
- Reinke, W., Herman, K., & Sprick, R. (2011). *Motivational interviewing for effective classroom management: The classroom check-up*. New York, NY: Guilford.
- Rusby, J., Taylor, T., & Foster, E. (2007). A descriptive study of school discipline referrals in first grade.

Psychology in the Schools, 44(4), 333–350.

Skinner, B. (1938). *The behavior of organisms: An experimental analysis*. New York, NY: Appleton-Century.

Stringer, E. T. (2013). *Action research* (4th ed). Thousand Oaks, CA: Sage.

Sugai, G., Sprague, J., Horner, J., Walker, R., & Hill, M. (2000). Preventing school violence: The use of office discipline referrals to assess and monitor school-wide discipline interventions. *Journal of Emotional and Behavioral Disorders*, 8(2), 94-101.

Westling, D. L. (2010). Teachers and challenging behavior: Knowledge, views, and practices. *Remedial and Special Education*, 31(1), 48–63.

AN ACTION RESEARCH ON ENCOURAGING STUDENTS TO SPEAK IN SECOND LANGUAGE EDUCATION: A LEARNER-CENTERED APPROACH

Vahid Rahmani Doqaruni

Ferdowsi University of Mashhad

Abstract

Although the ability to speak to people in their language or in a language that both speakers can understand is definitely the aim of many second language learners, some of the students keep silent all the time in class. This article reports on an action research focusing on increasing the English as a Foreign Language (EFL) students' confidence in speaking in an Iranian context. Ten male university students from different majors participated in this study. The method used was the learner-centered approach to increase students' performance in terms of pair work and group work. The present study used teacher observation as the sole elicitation instrument in order to gain a rich understanding of the participants' confidence development. The findings suggest that although teacher-centered instruction, typically utilizing little interaction, is the most common in EFL classrooms, an instructional methodology stressing peer collaboration as a tool for increasing the ability of the students to speak is likely to result in higher confidence.

Keywords: Action Research; Student Confidence; Speaking Skill; Learner-Centered Approach; Teacher Observation; Peer Collaboration

Introduction

Speaking skills are an important part of the curriculum in language teaching and the ability to speak in a foreign language is at the very heart of what it means to be able to use a foreign language (Biggs & Moore, 1993; Liu, 2001; Tsou, 2005; Tsui, 1992; Van den Branden, Bygate, and Norris, 2009). Being able to speak to friends, colleagues, visitors and even strangers, in their language or in a language which both speakers can understand, is surely the goal of very many learners (Luoma, 2004). In addition, the linkage between students' classroom oral participation and their academic achievement is undeniable. Studies have

shown that when students participate actively in class, their academic achievement seems to be higher than that of those who are passive in class (Krupa-Kwiatkowski, 1998). The importance of the ability or perception of ability to speak should not therefore be underestimated by either teacher or pupil (Turner, 2010).

In my teaching experience, however, some of the students keep silent all the time in class. They do not want to speak English. Even when they know the answer to a simple question, they hesitate to open their mouths. Although they clearly know that they should speak English well in order to make themselves qualified for the current competitive society, they are indifferent about talking in English. Similarly, by way of interviews, observations, journals, and surveys, research has revealed that Second Language (L2) learners often seem passive and reticent in language classrooms. Encouraging students to talk in a language classroom is thus a problem that many language teachers face (Tsui, 1996; White & Lightbown, 1984). With the advent of globalization, however, there is a pressing need for English as a Foreign Language (EFL) teachers to help reticent students develop the skills, and confidence needed to take an active role in oral classroom lessons (Liu & Jackson, 2009). In this way, students should be encouraged to work cooperatively so that they could have opportunities to progress in achieving their academic objectives (Dailey, 2009; Tong, 2010). Only in this situation, they can share ideas and structures with each other.

Researchers have discovered various reasons for this reluctance to speak in L2 classroom situations such as the following: fear of losing face; low proficiency in the L2; previous negative experiences with speaking in class; cultural beliefs about appropriate behavior in classroom contexts; habits; personality; and lack of confidence. Since few, if any, studies have previously provided valuable information about confidence in an Iranian L2 context, the aim in this paper is to look at this affective variable through doing an action research and find out whether encouraging students to collaborate with their peers in extra speaking activities incorporated into the classroom leads to a more active role on the part of learners and increases their self-confidence.

Why choose action research for the present study? The first thought that comes to mind when one hears the term "research" is that the researcher should spend months or even years going through many difficult steps in a process to create a scholarly piece of work. However, many teachers are so concerned with their routine activities that they are left with almost no time for research whatsoever (Salmani Nodoushan, 2009). In this way, many teachers consider research an extra burden in their regular schedule unless the word research is used in a new sense. Here is where action research plays an important role. It is more practical and user-friendly than research defined in traditional senses and does not necessitate some essential features of the conventional research such as a sizable sample of participants or time limitations (Wallace, 1998). Given that I sought to inform my own teaching practice with the research—particularly with my specific classroom context—it was clear that an action research project would be the most suitable for my purposes.

Literature Review

Confidence, motivation, and language ability are often treated as distinct but related learning dimensions in the field of second language (Clement & Kruidenier, 1985). The literature widely holds that these concepts are directly related and impact each other and that if one of the factors increases or decreases, the others will follow in a direct relationship. Yashima, Zenuk-Nishide, and Shimizu (2004) claim "considering that students need to communicate in order to improve communicative skills and gain confidence, the researchers hope to postulate a circular and interactive model to show the dynamics of interest, motivation, learning, confidence, and communication" (p. 144). Thus, few, if any, attempts have been made to explore confidence in isolation and this concept has largely been regarded as a corollary of other studies dealing with other affective variables such as anxiety or motivation. However, these studies are important since they have identified an association between self-confidence in language ability and other language-related phenomena.

An important strand of research in second language acquisition has investigated the relationship between reticence and confidence. Riasati (2014), in his recent review of literature on reticence, argues that reticence has a detrimental effect on the L2 learning process. He believes that this is why many different researchers have made use of various ways of identifying students' reasons of reticence. In order to understand why some language learners choose to remain silent in language classrooms, these researchers have employed interviews, observations, and journals. Dwyer and Heller-Murphy (1996), for example, conducted interviews with six Japanese students to identify sources of reticence among these learners. It was found that the students were reticent due to several reasons, including lack of self-confidence. Similar findings were achieved by Li and Liu (2011) who investigated the issue of reticence in China. They assert that many Chinese learners prefer to remain silent in language learning contexts. A range of factors is identified as leading to reticence among which is lack of self-confidence. However, they believe that "through more personal engagement with class activities, it is expected that both students' confidence and their ability to speak English will be improved" (Li & Liu, 2011, p. 961).

Many previous studies have shown that there is a direct relationship between students' confidence and their speaking behavior in L2. For example, Lai (1994) attempted to identify Hong Kong secondary students' level of confidence in using English and the factors leading to different confidence levels in oral participation in classrooms. The findings show that most of the subjects "felt a lack of confidence in using English as a means of communication in the classroom" (Lai, 1994, p. 122). In another study, MacIntyre, Clement, Dornyei, and Noels (1998) suggest that self-confidence significantly contributes to the learner's willingness to communicate in a foreign language. According to them, affective factors such as motivation, personality, intergroup climate, and self-confidence underlie willingness to communicate, and the factor of self-confidence including overall self-confidence in L2 and situational self-confidence in communication play an important role in determining the

learner's willingness to communicate. Yashima et al. (2004) cite a study of high school students who traveled abroad to study English. Some students were not ready to communicate due to some factors, including lack of L2 confidence, and found themselves in an endless cycle: needing to communicate with native speakers to gain L2 confidence, but due to a lack of confidence, unable to initiate interactions. Cao (2011), in a recent study, investigated the willingness to communicate of a group of international students studying EAP (English for Academic Purposes) during one academic year at a New Zealand university. She suggests that three dimensions overlap and interrelate in order to inhibit or facilitate learners' willingness to communicate: First, individual characteristics, including self-confidence; second, classroom environmental conditions, such as topic; and third, linguistic factors. Thus an *ecological* perspective is drawn, whereby an individual is seen as interacting in a complex fashion with his/her environment.

Due to its negative effects, some researchers in their studies attempted to propose solutions for students' lack of confidence in L2 classrooms. A study conducted by Burden (2004) reveals that almost 70% of 289 university freshmen surveyed felt unconfident speaking English. Burden (2004) thus suggested teachers use *cooperative* as opposed to *competitive* goal structures as a means of creating interdependencies between learners to increase their self-confidence. Ewald (2007) reports that only about half of the students in her study claimed to experience a relative level of confidence in their upper-level classes. She then suggested that teachers work actively to build upper-level students' confidence through encouraging them. In her words, "convince them that the challenge of learning to use a foreign language is not outside their grasp. Assure them that mistakes are normal and expected and that even through flawed participation they learn" (Ewald, 2007, p. 134).

Problem Statement

The problem I identified in my classroom was that most of my students were not active enough during the class and did not speak most of the time. Following informal talks to my students, I discovered that most of them had problems with speaking English. In order to investigate this problem, I found it necessary to determine what speaking activities I currently use and the amount of interactions and language my students generate. For that, I made use of audio recordings of my class (three sessions) as tools for aiding my research. The first issue to be investigated was the type of speaking activities I used in my class and the students' response to those activities. In addition, I wanted to know about teacher and student talking time and the amount of my students' speaking. Listening to the recordings, I found out that my teaching was not as effective as I hoped it to be. The evidence for that was that my students were silent during speaking activities most of the time, and it seemed that they did not have enough confidence for speaking. Investigating students' attitudes toward learning English, I asked the question: which skill do you want to improve the most? I found that most of them wanted to improve speaking, but they were not brave enough to express their ideas. In other words, they wanted to speak, but they lacked *confidence* to speak. Regarding the activities, just a small portion of my teaching was devoted to this skill

because I was only teaching based on the book I had to cover in that semester in which all the skills were worked on simultaneously and no especial attention was paid to oral productive skills.

While language teachers often teach based on informal analyses of their learners' needs (Tarone & Yule, 1989), however, I thought that this informal analysis was not enough to confirm that students lacked confidence in order to speak. So I decided to investigate students' attitudes toward speaking skill through collecting data from a need analysis questionnaire (see Appendix A) I had designed for this purpose. The needs analysis revealed that most of the students were really interested in speaking, but they had low confidence in speaking English. Further, the data showed that the students attributed their inability to speak English confidently to the lack of speaking experience and/or opportunities to engage in L2 conversation inside the classroom. Thus, the key research question that forms the basis of the present action research is as the following:

Does the EFL students' confidence in speaking in the classroom enhance as a result of providing students with additional speaking activities?

Methodology

Participants. Participants involved in this study were 10 university students at B.A. level from different fields of study. They were all males and had participated in a general English classroom in a private English language institute in Iran. The students had an intermediate level of English, as determined by their TOEFL (Test of English as a Foreign Language) test taken by the institute. They had two 90-minute language sessions per week over a 12-week semester.

Context. The participants in this study had a reasonable knowledge of English grammar but were reticent to speak or produce the target language. According to Jahangard (2007), one of the main reasons that contributes to the Iranian students' inability to speak English is that students' aural and oral skills are not emphasized in Iranian prescribed EFL textbooks, especially at high school. These skills are not tested in the university entrance examination, as well as in the final exams during the three years of senior high school and one year of pre-university education. Teachers put much less emphasis, if any, on oral drills, listening and speaking abilities than on reading, writing, grammar and vocabulary. The main focus is to make students pass tests and exams, and because productive abilities of students are not tested, most teachers then skip the oral drills in the prescribed books. Thus, the students are the products of a deficient educational system that has not allowed for active participation of the students in the classroom.

Material. Many previous studies have shown that L2 students' lack of confidence is attributed to their lack of speaking practice (Benson, 1991; Biggs, 1994; Schneider, 2001).

For example, Kubo (2009) claims that the lack of the opportunity to practice speaking is one of the main factors that result in lack of confidence in students. Since I wanted to increase the students' confidence in speaking, they had to have more opportunities to practice speaking together with their friends. In addition, many researchers so far have proposed that students' confidence is enhanced through providing them with non-threatening activities. Using the learner-centered approach, the non-threatening activities in this study were one of the main factors to encourage students to participate in the class. This approach is a subtype of the humanistic approach, which is defined by Richards (2002, p. 13) as:

- a) the development of human values
- b) growth in self-awareness and in the understanding of other
- c) sensitivity to human feelings and emotions
- d) active student involvement in learning and in the way learning takes place.

Since I wanted to promote students' desire to express their ideas and to reduce their anxiety toward speaking, it seems that all of these principles are relevant to this project. In other words, to promote students' confidence in speaking English, non-threatening activities can be seen as a tool to promote students to be aware of their performance and feelings (Legutke & Thomas, 1991). Thus, I made use of one of the authorized books on speaking on the market in which activities were based on real-world events. *Speaking Extra* (Gammidge, 2004) is a resource book containing materials for supplementary classroom work that could be photocopied. The book helps the learners "to speak with confidence to carry out the most basic social transactions" (Gammidge, 2004, p. 7).

Procedure. I started incorporating extra speaking activities into the classroom from week 3 to the end of the semester (covering 20 sessions during 10 weeks). Twenty minutes of each session was devoted to supplementary speaking activities. The ordinary technique used in the book to engage students in speaking was "interview". In this technique, one learner was usually asked questions by one or several interviewers. As Meng (2009) asserts, "this strategy [technique] is useful for keeping a conversation going and is a worthwhile speaking activity" (p. 220).

Data Type. The present study used teacher observation as the sole elicitation instrument in order to gain a rich understanding of the participants' confidence development and of the impact of the extracurricular speaking activities on that development. In other words, the data source was my own notes based on weekly record of the most/least active, silent, confident students doing the extracurricular speaking activities during 10 weeks (from week 3 to week 12). The data also included thoughts that I recorded while preparing students for speaking activities, during class, and immediately after class, as well as my reactions to conversations with the students during class.

Results

I often went around the class whilst students were engaged in work and had a look at their interview sheets and spoke to individual students about their perceived progress, and any difficulties that they had encountered. I recorded everything that happened in my classroom, of course from my point of view, and my own feelings and thoughts about students' oral participation in my personal notes. In order to get a better idea of the students' progress, I have chosen three excerpts from my notes from three different weeks (week 5, week 8, and week 11, respectively) in a timeline. The following observation excerpt is from one of my early notes and represents an ordinary scene of my classroom situation three weeks after starting the study:

The case is often like this: two or three pairs of students are very active in doing the extra speaking activities and actively participate in their interview tasks. They are more talkative. The others speak softly as they are whispering to each other. They can perform well if I ask them to engage more actively in the activities, ask and answer more questions or state personal ideas. If not asked, they remain almost silent. The situation is even worse when they are asked to work in groups. I feel I need to make an effort...I feel that I need to work harder to get them feel relax, and get them to speak.

As the term progressed and the students gained more exposure to spoken English and became more used to interactive classroom activities, many changed from being quiet to being (more) active, a tendency not only reported by the students themselves, but also reported in one of my middle notes:

While doing the activities, students gradually decrease dependence on me and some of them have even developed as autonomous speakers. In the classroom setting, the students are brave enough to ask both their friends and me... I think that students tend to be surer of themselves after passing the first few weeks and their participation is a lot better. Many students are now more involved in the pair/group work activities and speak out...I noted that some of them called on their classmates for help when they were unsure of what to say. This group effort could be linked to recognizing group values which promote group solidarity and helpfulness...they obviously show willingness to talk and are not afraid of embarrassment anymore.

From the eighth week on, however, almost all the students in each pair/group tended to be active in extra speaking activities and no one was quiet, especially during pair work. In one of my late entries, I have particularly commented that no student was reticent during pair work:

It seems that the students are satisfied with their pair work activities. All of them actively participate in their interview activities and exchange ideas... I think the students appeared to be the most nervous for the first few weeks of the term. But as time went on, they became less and less nervous. I don't think the students are nervous during pair work because nobody else can hear them... In these last weeks of the term, I think most of them are somehow confident and relaxed... It seems that the students' attitudes to speaking have

changed. They enjoy participating in the speaking activities and are also enthusiastic and very willing to participate in these activities... Working with their friends is fun... The feeling of enjoyment and enthusiasm has motivated them and provided a positive attitude toward speaking English.

Discussion

I believe that curriculum would be better developed if teachers acted as action researchers. Elyildirim and Ashton (2006) support that action research can improve the current teaching situation in terms of boosting teachers' professional development, teacher training and presenting to an institution evidence of the need for change. This was my own experience after I undertook this action research project. At first, I tried to get information about the problems of the course in terms of students' feelings toward language learning, especially speaking. I did it through a need analysis questionnaire I had designed and speaking with my students informally. After that, I understood that students would like to speak, but they lacked confidence in speaking English. Thus, it made me think of how to provide interacting and participatory activities, materials and the opportunity to practice speaking.

The learner-centered paradigm in language teaching curriculum emphasizes on an interactive, participatory and collaborative approach to speaking. In my project, the interactive approach was used to increase students' performance in terms of pair work and group work. They had more opportunities to practice speaking due to some additional activities, because the main aspects of activities focused on an interactive approach. I used the participatory approach with students because it was very important to me to comprehend students' responses in the activities. I acted as teacher, helper, encourager and facilitator to help students to reach their goals and develop their confidence. In my opinion, the best ways to know what learners need are close interaction and participation with them as much as possible. As a result, teachers who do action research in the class with learners are a significant factor in developing language curriculum, which is one of the main factors to promote students' confidence in language learning, especially speaking.

I tried to practice as a teacher and a researcher simultaneously. If I know about classroom and educational information, I will have the potential to reflect on and analyze my own teaching in real situations, which are very important for teachers who have more opportunities to create curriculum by themselves. In addition, as a teacher and researcher, I gained significant benefits from students' feedback to improve my teaching. I could understand students' responses and their behavior in terms of passive and active learners after I had a chance to talk informally and get feedback from them. Although it was very difficult for me to consider students' responses to my teaching, it was worth doing as I discovered in my action research study.

Conclusion

This research project combined two areas in L2 teaching profession – students' confidence in speaking English in EFL classrooms and teachers' practical, classroom-based action research – both of which have been recognized "as not only being of importance, but also as areas in which there is still much progress to be made" (Curtis, 2001, p. 69). In this action research project, extra speaking activities were incorporated into the classroom to increase the students' performance in terms of pair/group work. Although teacher-centered instruction, typically utilizing little interaction in English, is the most common in EFL classrooms, this study showed that an instructional methodology stressing peer collaboration as a tool for increasing the ability of the students to speak is likely to result in confidence. The findings of this study thus suggest that L2 instructors should seek ways to include students' collaboration in the subject language as part of their curriculum design to help them gain confidence in speaking L2.

About the Author

Vahid Rahmani Doqaruni holds a B.A. in English Language and Literature and an M.A. in Teaching English as a Foreign Language (TEFL). He is currently a Ph.D. candidate in TEFL at Ferdowsi University of Mashhad and a lecturer in English Language Teaching at the University of Gonabad, Iran. He has published several papers on second language education research in national and international journals. His major research interests are: action research, teacher education, second language writing, mixed methods research and task-based language teaching. Email: rahmanidoqaruni@mail.um.ac.ir

Reference List

- Benson, M. (1991). Attitudes and motivation towards English: A survey of Japanese freshmen. *RELC Journal*, 22(1), 34-48.
- Biggs, J. (1994). From theory to practice: A cognitive systems approach. *Higher Education Research and Development*, 12, 73-86.
- Biggs, J., & Moore, P. J. (1993). *The process of learning*. Sydney: Prentice Hall of Australia.
- Burden, P. (2004). The teacher as facilitator: Reducing anxiety in the EFL university classroom. *JALT Hokkaido Journal*, 8, 3-18.
- Cao, Y. (2011). Investigating situational willingness to communicate within second language classrooms from an ecological perspective. *System*, 39 (4), 468-479.
- Clement, R., & Kruidenier, B. G. (1985). Aptitude, attitude and motivation in second language proficiency: A test of Clement's model. *Journal of Language and Social Psychology*, 4, 273-291.
- Curtis, A. (2001). Hong Kong secondary school teachers' first experiences of action research. *The PAC Journal*, 1(1), 65-78.
- Dailey, M. I. (2009). Acting out: A one-year drama class to increase participation. *The Language Teacher*, 33(12), 7-11.
- Dwyer, E., & Heller-Murphy, A. (1996). Japanese learners in speaking classes. *Edinburgh Working Papers in Applied Linguistics*, 7, 46-55.
- Elyildirim, S., & Ashton, S. (2006). Creating positive attitudes towards English as a foreign language. *English Teaching Forum*, 44(4), 2-11.
- Ewald, J. D. (2007). Foreign language learning anxiety in upper-level classes: Involving students as researchers. *Foreign Language Annals*, 40 (1), 122-142.
- Gammidge, M. (2004). *Speaking extra*. Cambridge: Cambridge University Press.
- Jahangard, A. (2007). Evaluation of EFL materials taught at Iranian public high schools. *The Asian EFL Journal*, 9(2), 130-150.
- Krupa-Kwiatkowski, M. (1998). You shouldn't have brought me here!: Interaction strategies in the silent period of an inner-direct second language learner. *Research on Language and Social Interaction*, 31(2), 133-175.
- Kubo, M. (2009). Extensive pair taping for college students in Japan: Action research in confidence and fluency building. *Accents Asia*, 3(1), 36-68.
- Lai, C. (1994). Communication failure in the language classroom: An exploration of causes. *RELC Journal*, 25, 99-129.
- Legutke, M., & Thomas, H. (1991). *Process and experience in the language classroom*. London: Longman.
- Li, H., & Liu, Y. (2011). A brief study of reticence in ESL class. *Theory and Practice in Language Studies*, 1, 961-965.
- Liu, J. (2001). *Asian students' classroom communication patterns in American classrooms*. Westport, CT: Greenwood Publishing.

- Liu, M., & Jackson, J. (2009). Reticence in Chinese EFL students at varied proficiency levels. *TESL Canada Journal*, 26(2), 65-81.
- Luoma, S. (2004). *Assessing speaking*. Cambridge: Cambridge University Press.
- MacIntyre, P. D., Clement, R., Dornyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal*, 82, 545-562.
- Meng, F. (2009). Encourage learners in the large class to speak English in group work. *English Language Teaching*, 2(3), 219-224.
- Nunan, D. (1998). *Syllabus design*. Oxford: Oxford University Press.
- Riasati, M. J. (2014). Causes of reticence: Engendering willingness to speak in language classrooms. *International Journal of Research Studies in Language Learning*, 3, 115-122.
- Richards, J. C. (2002). 30 years of TEFL/ TESL: A personal reflection. *RELC Journal*, 33(2), 1-35.
- Salmani Nodoushan, M. A. (2009). Improving learning and teaching through action research. *The Modern Journal of Applied Linguistics*, 1(4), 211-222.
- Schneider, P. H. (2001). Pair taping: Increasing motivation and achievement with a fluency practice. *TESL-EJ*, 5(2), 1-32.
- Tarone, E., & Yule, G., (1989). *Focus on the language learner: Approaches to identifying and meeting the needs of second language learners*. Oxford: Oxford University Press.
- Tong, J. (2010). Some observations of students' reticent and participatory behavior in Hong Kong English classrooms. *Electronic Journal of Foreign Language Teaching*, 7(2), 239-254.
- Tsou, W. (2005). Improving speaking skills through instruction in oral classroom participation. *Foreign Language Annals*, 38(1), 46-55.
- Tsui, B. M. (1992). Classroom discourse analysis in ESL teacher education. *ILEJ*, 9, 81-96.
- Tsui, B. M. (1996). Reticence and anxiety in second language learning. In K. M. Bailey & D. Nunan (Eds.), *Voices from the language classroom: Qualitative research in second language education* (pp. 145-167). Cambridge: Cambridge University Press.
- Turner, J. (2010). The impact of pupil-set targets on achievement in speaking: An action research project involving a mixed-sex Year 9 German class. *Journal of Trainee Teacher Educational Research*, 1, 1-34.
- Van den Branden, K., Bygate, M., & Norris, J. M. (Eds.) (2009). *Task-based language teaching: A reader*. Amsterdam: John Benjamin Publishing.
- Wallace, M. (1998). *Action research for language teachers*. Cambridge: Cambridge University Press.
- White, J., & Lightbown, P. (1984). Asking and answering in ESL classes. *The Canadian Modern Language Review*, 40(2), 228-244.
- Yashima, T., Zenuk-Nishide, L., & Shimizu, K. (2004). The influence of attitudes and affect on willingness to communicate and second language communication. *Language Learning*, 54 (1), 119-152.

Appendix A: A Sample of Needs Analysis Questionnaire (Adapted from Nunan, 1998)

Age:

Field of Study:

- 1) Do you like to learn English by: Reading Writing Listening Speaking
- 2) Do you like to: Study grammar Learn new words Practice conversation
- 3) Do you like to learn English by: Cassettes Games Conversation Studying English books
Watching T.V.
- 4) Do you like to learn by talking to friends in English? Yes No
- 5) Do you like to learn English words by: Seeing them? Using them?
- 6) Do you like to learn English with the whole class? Yes No
- 7) Do you like to learn English by talking in pairs? Yes No
- 8) Do you speak English out of the class? Yes No
- 9) Do you feel happy when you speak English? Yes No
- 10) Do you look for chances to speak English? Yes No
- 11) Do you have enough confidence to speak to a group of people in English? Yes No
- 12) Do you feel relaxed when you speak in English? Yes No
- 13) How do you learn best? Alone Pairs Small group Class Outside class
- 14) What do you feel are the most important things for you to learn in the:

Short term:

Long term: