**Literacy Studies:** Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education 25

Xinghua Liu
Michael Hebert
Rui A. Alves *Editors*

# The Hitchhiker's Guide to Writing Research

## A Festschrift for Steve Graham

Springer

# Literacy Studies

## Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education

Volume 25

While language defines humanity, literacy defines civilization. Understandably, illiteracy or difficulties in acquiring literacy skills have become a major concern of our technological society. A conservative estimate of the prevalence of literacy problems would put the figure at more than a billion people in the world. Because of the seriousness of the problem, research in literacy acquisition and its breakdown is pursued with enormous vigor and persistence by experts from diverse backgrounds such as cognitive psychology, neuroscience, linguistics and education. This, of course, has resulted in a plethora of data, and consequently it has become difficult to integrate this abundance of information into a coherent body because of the artificial barriers that exist among different professional specialties.

The purpose of this series is to bring together the available research studies into a coherent body of knowledge. Publications in this series are of interest to educators, clinicians and research scientists in the above-mentioned specialties.

Some of the titles suitable for the Series are: fMRI, brain imaging techniques and reading skills, orthography and literacy; and research based techniques for improving decoding, vocabulary, spelling, and comprehension skills.

Book proposals for this series may be submitted to the Publishing Editor: Marianna Georgouli; *Springer; Van Godewijckstraat 30;3300 AA Dordrecht; The Netherlands*; e-mail: Marianna.Georgouli@springer.com

Xinghua Liu • Michael Hebert • Rui A. Alves
Editors

# The Hitchhiker's Guide to Writing Research

## A Festschrift for Steve Graham

Springer

*Editors*
Xinghua Liu
School of Foreign Languages, Yang Yong Man Building, Shanghai Jiao Tong University
Shanghai, China

Rui A. Alves
Faculdade de Psicologia e de Ciências de Educação, Universidade do Porto
Porto, Portugal

Michael Hebert
School of Education
University of California
Irvine, CA, USA

ISSN 2214-000X ISSN 2214-0018 (electronic)
Literacy Studies
ISBN 978-3-031-36471-6 ISBN 978-3-031-36472-3 (eBook)
https://doi.org/10.1007/978-3-031-36472-3



This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Paper in this product is recyclable.

# Preface

This volume is intended as a tribute to Professor Steve Graham, a truly exceptional person on both professional and personal levels. Since receiving his PhD in 1978, his career has spanned 45 years as of the writing of this chapter, and he has established himself as one of the most productive researchers in educational psychology. As of July 2022, Steve has solely authored and coauthored 322 journal papers, 138 book chapters, 11 white papers of national significance, and 4 monographs; coedited 36 journal special issues and 22 book volumes. Impressive, not only in its sheer number, but crucially on its specific impact in writing research and science in general. Google scholar counts nearly 70k citations and an h index of 146. In the last six years, Graham's work received an average of about 5k citations yearly. Or, if we put it daily, each single day Graham's works are credited 14 times. Besides advancing so successfully his own career, Steve has been contributing enormously to the service of the community by serving as the editor for journals like *Exceptional Children*, *Contemporary Educational Psychology*, *Journal of Writing Research*, *Focus on Exceptional Children*, and *Journal of Educational Psychology*. His incredible impact on the field of writing research has been well recognized worldwide by 30 or so prizes, among which are *Thorndike Career Award from Division 15 of the American Psychological Association*, *Sylvia Scribner Award from Division C of the American Educational Research Association*, *John S. Nesbit Fellowship from the British Educational Research Association*, *Exemplary Research in Teaching and Teacher Education from Division K of the American Educational Research Award*, *Career Research Award from the International Council for Exceptional Children*, just to name a few. He was also elected to the Reading Hall of Fame for 2018. Steve's impact on writing research has been so large that his standing in the field is that of a living legend. There is no doubt that he will continue to have a profound influence on the field.

Apart from celebrating his significant academic contributions, on a personal level, we are equally moved to honor Steve for the love, inspiration, and friendship that he has bestowed upon us, his students, and colleagues. Steve is a kind, gentle, thoughtful, ingenious, supportive, and cooperative person, endowed with a great sense of humor and someone who is always fun to be with. Steve's mentorship of

novice researchers has also been impactful. The editors of this book, as well as most of the chapter authors, have worked with Steve on different research projects as PhD students and/or co-investigators. Other chapters have been written by colleagues who are themselves distinguished scholars in the field. The contributors to this volume have all been enriched by Steve in personal and scholarly ways. And this personal and scholarly influence is evident in the reflections, accompanying full-length research chapters. These short reflections clearly show the way Steve has touched and influenced their lives as a scholar, colleague, and friend.

A note about our title, "The Hitchhiker's Guide to Writing Research: A Festschrift for Steve Graham." For those that know Steve, this title is fitting on multiple levels. First and foremost, Steve spent a good deal of time in his life hitchhiking and is known to regale friends with his stories of hitchhiking his way through the country. His stories sound a lot like Jack Kerouac without the psychedelics. In fact, Steve even wrote about his hitchhiking adventures in one of his reflections (Graham, 2021), so our inclusion of hitchhiking in the title is pretty obvious. However, if you read Steve's work carefully, you'll also note that he is also a big fan of science fiction and fantasy writing. Sometimes he's obvious about this in the titles of his work (e.g., Wrath of Kahn), and other times, he more subtly uses pseudonyms for his participants that you might recognize from a favorite novel or two. In person, he's also apt to remind you that the answer is 42. (What was the question again, Steve?) Therefore, our title is an obvious homage to one of Steve's favorite books, "The Hitchhiker's Guide to the Galaxy." Finally, and potentially the best reason for the title, Steve has traveled through almost every corner of the writing research landscape through his many reviews and meta-analyses. Much like how Arthur Dent and Lord Prefect navigated their way through the galaxy in the "Hitchhiker's Guide to the Galaxy," Steve has hitched rides with writing researchers around the world and through history, and he's been gracious enough to share what he's learned in his own version of a guide to the galaxy of writing research—his life's work.

Of course, his life's work is also the reason we're honoring Steve with this book, and the reason for our subtitle. A festschrift is a collection of writings in honor of a scholar, and we can think of no better way to characterize this book. As a collection of works from a variety of colleagues and former students around the world, this book was not designed with the theme or destination in mind. Instead, our readers are invited to hitch rides with each of our authors and experience some of the landscape of the writing research world with them. Just like Steve Graham the Hitchhiker.

To help guide you in your journey through this book, we have divided the book into seven parts. Part I opens the book with Karen Harris' touching recall of how she met Steve Graham and what they have been researching. This very first chapter also contains a short reflection by their final doctoral student, April Camping.

Part II deals with experimenting new and expanding existing writing models, and it contains three papers. John R. Hayes proposes a process model for how writers express their emotions through word choice in written text. Building on Steve Graham's writer(s)-within writing model, Clarence Ng and Peter Renshaw, in the second paper in this part, reconceptualize the personal and social realms in terms of Vygotsky's notion of perezhivanie, and A. Angelique Aitken in the third paper

expands the model by viewing it through a motivational lens of Bandura's social cognitive motivational theory.

Part III is devoted to a variety of topics on writing instructions at school levels. Jill Fitzgerald, Jackie Eunjung Relyea, Jeff Elmore, and James S. Kim investigated the extent of first-grade children's academic vocabulary usage in writing. Amy Gillespie Rouse, Ashley Sandoval, and Murphy K. Young examined students' knowledge of the writing process and their knowledge of how writing can be used for different purposes in STEM subject areas. Amanda L. Lindner, Kausalai Wijekumar, and Debra McKeown analyzed the spelling errors of fifth-grade students in the United States. Michel Fayol, Bernard Slusarczyk, Virginia Berninger, and Pascal Bressoux studied the relationship between French spelling and French written compositions both within and across grade levels. Bruce Saddler situated sentence combining within the writing process and reviewed theoretical benefits and practical applications of sentence combining.

Part IV has three papers focusing on writing instruction for students with learning disabilities. Linda H. Mason and Jenna Basile's systematic review examined the effectiveness and treatment acceptability of empirical intervention research in summary and quick writing instruction for students with high-incidence disabilities. Sharlene A. Kiuhara, Joel R. Levin, Malynda Tolbert, Megan Erickson, and Kenny Kruse developed a writing-to-learn mathematics intervention that incorporated the six-stages of Self-Regulated Strategy Development (SRSD) and studied the effectiveness of this intervention for students with a mathematics disability in their understanding of fractions as numbers and their argumentative writing and mathematical reasoning. Naomi Weintraub investigated the distribution of handwriting difficulties (HD) manifestations among higher-education students, and the relationship between language, working memory as well as perceptual- and grapho-motor skills and handwriting legibility and speed. Amber B. Ray reviewed studies using SRSD to teach writing to secondary students with and at-risk for learning disabilities.

Part V revolves around teacher practice and professional development. By presenting results from a small-scale teacher development for writing instruction project which was reduced from the originally elegantly perceived professional development in evidence-based writing instruction, Erin FitzPatrick, Debra McKeown, Megan C. Brown, and Nicole Patton-Terry discussed the challenges of conducting rigorous scientific research in high-poverty, high-mobility urban settings. Alyson A. Collins, Stephen Ciullo, and Micheal P. Sandbank used generalizability theory to investigate the dependability of scores from an observation tool designed to measure writing instruction in upper elementary classrooms. Gary A. Troia examined the relationships between factors like teachers' preparation for writing instruction, their perceived competence to teach writing, and their knowledge about writing-related concepts and their instructional practices. Tien Ping Hsiang conducted an interview-based qualitative study to reveal how Chinese language teachers from nine private schools in Macao, China, taught Chinese reading and writing in primary grades.

Part VI includes three papers on writing practices by adult learners. Charles A MacArthur reviewed research relevant to recent efforts to reform community college developmental writing programs and recommended more research on improved

pedagogical methods. Teresa Jacques, Ana P. Azzam, Francisca Costa, and Rui A. Alves manipulated a disclosure topic and pronoun perspective in expressive writing tasks produced by undergraduates and examined the corresponding effects of this manipulation on the linguistic and emotional content used during writing. Situated within the Systemic Functional Linguistics paradigm, Xinghua Liu and Xuan Jiang analyzed the Attitude items in Chinese and American university students' written feedback to their counterparts' English writing in a telecollaborative written peer feedback project.

In Part VII, the paper from Tanya Santangelo, Michael A. Hebert, and Pamela Shanahan Bazis concludes the book by reviewing Steve Graham's systematic reviews and meta-analyses and highlighting how Graham's work might be used to inform future writing research.

The various topics covered in this book testify the breadth of Steve's research and his influence in the field. He is active as ever and continues to generate new ideas and profound insights. We wish him peace and prosperity for years to come.

We are pleased to see the successful completion of this project. The project was initiated with the concept of giving Steve a surprise by without him knowing. So, for a span of three years, through a pandemic, and among some thirty colleagues across four continents, this project has been a closely guarded secret. By taking this opportunity, we would like to express our profound gratitude to all our chapter authors who accepted our invitation and contributed to this volume. We wholeheartedly thank them for their support and patience, without which, this project would not be possible. We are also thankful for the guidance and support that the editorial team at Springer, Helen van der Stelt, and Enayathullah M have provided. They have kindly offered to extend the deadline for many times. We also want to extend our sincere thanks to R. Malatesha (Malt) Joshi for his consultancy and encouragement during the whole process of this project. We are convinced that, thanks to the themes covered and the excellent quality of the scholarship, this volume will be of interest to both experts in the field and postgraduate students.

We also note that it was not possible to include all of the important colleagues and students Steve has ever worked with in this volume, and we apologize to those we missed. Perhaps we will be able to add a second volume to this collection in the near future. Until then, we look forward to hitchhiking with you through your work.

Shanghai, China — Xinghua Liu

Irvine, CA, USA — Michael Hebert

Porto, Portugal — Rui A. Alves

May 2023

# Personal Reflections

## Clarence Ng

It was during the summer of 2014 that I met Professor Steve Graham and his wife Professor Karen Harris the first time at the Brisbane campus of Australian Catholic University (ACU). They paid their first visit as the university's international professorial fellows affiliated with the Institute for Learning Sciences and Teacher Education. Since then, Steve has become a collaborator and a mentor in research.

As a collaborator, Steve introduced me to the world of writing research and invited me to work with him on different projects, including writing-reading connections and writing motivation. During his visit in 2014, Steve gave a talk to our faculty members on the effects of writing and writing instruction on reading based on his meta-analytic studies. Walking him back to the office after this event, I raised the question about the effect of reading on writing and the possibility of examining this complex connection. Steve responded enthusiastically and, almost immediately, set up a research team to take on this challenging work. Two weeks later, Steve came up with a dataset containing over 10,000 publications for a meta-analytic study examining the effects of reading on writing. I counted myself lucky to have the opportunity to work with him, alongside others, in this challenging study and to observe firsthand his leadership and commitment to high quality research work. In 2016, I visited Steve and others at the Tempe campus of Arizona State University. It was during this brief visit that I had the chance to deepen my collaboration with Steve and Karen, setting up plans for research and writing collaborations. As a collaborator, Steve is always reliable, committed, and generous with his time, support, and contribution. He has never turned down a project idea I raised with him. I have benefited profoundly from collaborating with Steve.

As a mentor, Steve showed me how to do meta-analyses, which Karen describes as "Steve's toy." Prior to that, I had no working knowledge about this complex analytical skill. Steve gratefully and willingly shared his understanding and experiences. Benefited personally from Steve's sharing, I asked him whether it was possible for him to conduct a seminar on meta-analysis during one of his subsequent

visits to ACU. As usual, Steve responded positively. His seminar was well-received and many of us here at ACU learnt much from his teaching on meta-analyses. As a mentor, Steve has shown me two significant attributes of a successful academic – passion and commitment. Steve is passionate about writing research. Anything to do with writing research will interest him. He is open to new ideas and is always prepared to have a discussion. When I talked with him about the Vygotskian concept of perezhivanie, he shared his encounter with sociocultural theories and how he considered them in his theoretical model of writer(s)-within-community model of writing. Steve is a committed researcher. He always leads by example. He works so fast and effectively that I sometimes think he may have come from another planet. On one occasion when we were working on a writing motivation survey, he came to my office and made a passing comment that "you are too slow" in a lighthearted manner. Steve is so committed to his work that he does not need coffee to keep awake (but lemon does). He seldom goes out for lunch in order to avoid breaking his work rhythm during the day – but had no hesitation in following me to a sushi restaurant in the Brisbane CBD to devour kingfish sashimi, his favorite.

I am grateful for the opportunities to work with Steve in the past seven years. Steve has made significant and unique contribution to writing research. His work on writing strategies, writing instruction, and writing intervention is influential and will continue to guide future research. His recent work on writer(s)-within-community model of writing is a theoretical mine rich with research gems, awaiting extraction.

## A. Angelique Aitken & Ashley Barkel

In a world where academics are swimming in data, typing their fingers to the bone, and are so busy, we met a superhero whose power was making time. Dr. Steve Graham, who was known by the alias Super Graham, as was printed on the uniform that he hid under a Mickey Mouse collared shirt, blue jeans, and Chuck Taylors, had the power of giving you as much unrushed, undivided time as you could imagine—whenever you needed it. "How about now?" could be his catchphrase. Something magical happens every time you meet with Super Graham; you become inspired. You become inspired to approach and think about questions differently—inspired to connect things like noisy air conditioners to how you are making sense of variance in educational research—inspired to keep exploring how much fun there is to be had uncovering the mysteries of the writing universe. His genuine nature, passion for discovery, and sense of humor are spellbinding. There is no superhero we admire or strive to be like more than Super Graham.

We had the privilege of being part of the Graham-Harris writing community at Arizona State University (ASU). Angelique and Ashley started their doctoral studies the same year. Angelique's primary advisor was Steve, and Ashley's was Karen (APA Writing Wonder Woman, but that is a story for another day). However, as both Steve and Karen say, regardless of your advisor, you get a 2-for-1, meaning we all

worked closely as a team throughout our time at ASU. As same-year sisters in the Harris-Graham writing community, we quickly became close friends and shared "Steve Stories," such as when the cashier at Sacks told Steve they didn't have any more lemons, or us wondering if iced tea was what fueled his super-hero status. Was his flip phone really a clever disguise for a high-tech device for creating meta-analyses in an instant? We talked and laughed about our Steve Stories because we liked and admired him. His zest (pun intended) for life and affable personality made him a fun mentor, even if he jokingly encouraged borderline workaholic behavior.

While we sat cramped in our shared office writing away, we discovered that Super Graham had a shadow side. Angelique remembers Ashley whispering tentatively, "have you ever noticed that Steve will take your pen and not give it back?" Ashley's brown eyes were slightly widened, as she looked at me, in part, expectantly and yet unsure of my response. Peering around her monitor, Angelique locked eyes with Ashley when it clicked! There was no black hole vortex of missing socks and writing utensils; Steve had been hoarding my pens and pencils! Once we confirmed this with others, we started calling him out. Such as when Steve grabbed a pen from Ashley's hand, and she asked if he was going to give it back. To which he replied, it was the tax we all owed him! We played along for a while, laughing at each pen-stealing moment, until one national conference where we visited all the vendor booths collecting at least one of every writing utensil we could find. We presented the dozens and dozens of pens, pencils, and highlighters to Steve in a terracotta flowerpot, which read, "Steve's Stolen Pens." Did the pen-stealing cease? Of course not, but at least we had a way to redirect his behavior. He said that students' pens just wrote better. Oh, Super Graham.

Angelique frequently talks about how much she appreciated Steve as an advisor and how he has served as a great model for mentoring students. For instance, one afternoon, while hanging out with some cohort mates, the topic of imposter syndrome came up. Others shared how they hoped their advisors (and the rest of the world) wouldn't find out that they didn't know how to do X, Y, or Z. Angelique stayed quiet during this part of the conversation because she realized that her doctoral experience was different from that of her friends. She felt appropriately guided in her academic writing as well as her meta-analytic, and intervention research skills. More importantly, anytime she had a question about something she didn't understand, she had felt completely comfortable asking Steve. One time, after missing a couple of multiple regression classes due to knee surgery, Angelique confessed that she was completely lost. Steve stopped what he was doing and explained the difference between standardized and nonstandardized beta coefficients. In hindsight, it seems like such a small question, but it untangled something that could have continued to affect Angelique's understanding of statistics. Of greater significance, was the realization that Steve cared about her, her success, and would always treat her with kindness and respect. What a Super Graham!

While kindness and respect are all well and good, we can't close this reflection on such sentimentality; so, let's talk about pie, Super Graham's kryptonite. Although his mental acuity and finger-pecking typing never slowed, he was always weakened by the presence of pie. Once, some of the team, including Xing Hua Liu, one of the

editors of this book, convened on a Saturday afternoon to work on a meta-analysis. Angelique brought a cherry crumb pie, one that she and Steve shared an affinity for. At one point, each of us took a small slice of this enormous, 12-inch pie. Delicious! As the afternoon wore on, Steve kept snagging another small piece here and there, but near the end, he was eating straight from the pie pan! At the end of the day, he had eaten 90% of the pie. He was kind enough to leave about three cherries and some soggy crust.

Oh, Steve, how we love to tease one another. Thank you for not only brightening our minds but brightening our lives. While we recognize the academic contributions you and Karen have left on the world, we wanted to focus on how you and Karen didn't rescue any of us from three-headed monsters, but your mentorship helped mold the next generation of writing researchers as they developed their own superpowers.

## Jill Fitzgerald

I first met Steve in the 1980s. We were among a group of writing researchers assembled for the purpose of briefing one of the federal organizations. I don't remember the date, place, or even which organization held the meeting, but I do remember the lasting impression I had of Steve from that day. Steve delivered a review of writing research. I well remember thinking, "Good gracious. He's been meticulous in method and analysis. And he's organized it all so clearly." When he addressed questions, it felt like he was "holding audience" with the group around him. He spoke with confidence, conviction, and even passion about the topic. Yes, I was impressed—so much so that I can still call up images from that day.

Steve is one of those rare individuals who has made contributions to the fields of both reading *and* writing, for both typically developing individuals *and* struggling readers and writers, and most especially for children with disabilities. Over the course of his professional lifetime, Steve's work has impacted classroom practice, he has been a policy influencer for decades, and he's provided amazing service to the field. It's impossible to portray the breadth and depth of Steve's professional work in 1,000 words (the limit for this reflection). But I want to at least stake a few claims.

### *Research That Impacts Classroom Practice*

Perhaps the most notable characteristic of Steve's research is that his interests always squarely focus on enhancing children's literacy learning through instructional methods that are fundamentally useful and understandable to teachers. I would wager that his influence on special education research and practice in particular is singular among special education scholars. His work has always been

prescient—forward-leaning, novel, and unique among peers. As an example, he has published over 30 studies examining the effectiveness of the Self-Regulated Strategy Development (SRSD) model, an instructional model designed to improve a wide range of students' writing from primary grades through high school.

## *Distinguished Scholarship*

It would be difficult to imagine that many scholars could match Steve's prolific level of publication. Steve has published over 400 books, chapters, and journal articles. Since 1997, Steve has been ranked among the top three education researchers for the number of articles published in top-tier educational psychology journals. In 2010, he was ranked among the three most productive scholars in educational psychology. Importantly, the quality of Steve's research is unsurpassable. His journal articles appear in top-tier journals with very high standards for rigor. Moreover, the number of Steve's honors and awards is unsurpassed in the literacy field. Among the many on the list are Fellowship in the American Educational Research Association and membership in the Reading Hall of Fame.

## *Leadership and Impact on Policy*

Steve is a literacy advocate who takes his findings to the public, and his influence on national policies has been far reaching. Most notably, he has energetically promoted the need for writing instruction at all levels of schooling. As a leading expert on reading and writing instruction, he has accomplished over 45 television/radio/web news presentations and interviews in the last decade alone. He is the "go-to" person for key national panels on writing. For instance, Steve was commissioned by the Carnegie Corporation to produce a series of reports on writing research and instruction, and he has participated on advisory panels for UNESCO, the National Institute of Health, the National Writing Project, and the Common Core curriculum.

## *Service to the Field*

The level of Steve's service to the literacy field is unrivaled. He has accomplished over 250 forms of professional service during his career, ranging from mentoring other literacy and special education professors to serving on literacy grantees' advisory boards to advising state departments on writing issues to service on national and international association committees. He has also served as editor for no fewer than five premier journals, a remarkable achievement, and he has been guest editor for 23 special issues in 11 different journals.

### *Last Thoughts*

As I write today, I think about Steve's personal qualities. The Steve Graham I know might be characterized as "determined" and "principled," but also warm and "down-to-earth." He's determined to learn, to grow into something new, to lead, and to help students and teachers and policy makers. And in the face of challenging times, Steve always, I mean always, weathers on. He perseveres and never complains (at least not out loud outside of his home). And he always has that glass of iced tea in his hand and a smile on his face as if he's ready to "talk shop," or write something he's thinking about, or joke with a friend, or discuss movies.

Some years ago, at a Pacific Coast Research Conference, Steve and Karen invited me to have dinner with them. What I remember most about the evening isn't the food. It's the companionship and the stories about Steve that some of the former students around the dinner table told me. I heard that students like to play "little tricks" on Steve. One former student told how the night before a meeting with Steve, the student snuck into Steve's office and removed all the pens and pencils from the pen/pencil holder on Steve's desk. The day of the meeting, there sat the empty pen/pencil holder. Steve kept fumbling, looking all over for a pen, which evidently led to some chuckling on the part of the culprit and compatriots. Students said that Steve always took the "little tricks" in good humor. The "little tricks" stories said something to me about Steve's character. How many professors have that kind of relationship with their students?

Today, because of Steve's professional influence and because of his friendship, I'm a more knowledgeable teacher, a more knowledgeable researcher, a better editor, a better reader/writer and movie watcher, and hopefully a better person and friend.

## Gillespie Rouse

I have spent months considering how I could summarize the immeasurable and impactful ways Steve Graham has influenced my life as a scholar, and more simply, as a person. I am not sure I have come to a solution, but I *am* certain I will never be able to adequately capture my gratitude for Steve's mentorship. This reflection is my best attempt.

I remember meeting Steve – quite by chance – because the professor I was scheduled to visit was out of town during my trip to Vanderbilt. I was nervous and uncertain about my chances of working with such an esteemed expert in the field of special education. Steve's office was piled high with stacks of papers (that I would later find out were for the meta-analyses he would teach me to conduct). He hung his Converse-clad feet over the side of his chair and had a conversation with me like we had known each other for years. He was unassuming and kind.

There was no pretense, no intimidation factor, and I felt completely at ease. I'm sure we discussed research and teaching that afternoon, but all I can remember is that we bonded over the rivalry of our favorite college basketball teams, mine was UVa and his was UNC. I left our meeting ecstatic that Steve Graham actually wanted to work with me, and I felt confident that Vanderbilt was the place I wanted to be. And so began Steve's dedication to my future success.

When I reflect on the time I spent at Vanderbilt, there are many memories of long days, countless hours reading and writing, and stressful deadlines. But, when I think specifically of the time I spent working with Steve and our research team, I remember mostly laughter and fun.

Steve has that way about research, and I think more broadly, about life. He made learning fun, which made handling all of the new theories, methods, data, and academic writing seem more manageable. And Steve always prioritized who I was as a person, not as a doctoral student, before all else. He always reminded me of the restaurants I should check out before traveling to any conference, and when I was on the job market, he made sure we discussed quality of life concerns just as much as we discussed my future program of research. Life outside of my PhD program was tumultuous and emotional, and some of the hardest I've endured, but my time spent working with Steve was supportive, steady, and meaningful. Steve offered an apprenticeship of sorts, never asking me to do more than I could, but always pushing me to reach my potential. He also always checked in with me, never asking too much but just enough to make sure I was OK. His dedication to me and my success was selfless and unwavering.

Now, more than a decade later, although our interactions are fewer than those doctoral days, I still have the same unwavering support from Steve. Whenever I have a question or need research or career advice, Steve (who I firmly believe may be the busiest person in the world) makes time to take my call. When he answers, his complete focus is on me and supporting my future success, just as it has been since that first meeting so many years ago. I've never experienced the level of dedication and professional support Steve has provided me. Perhaps the most impressive aspect of it all, is that I know Steve is devoted to supporting so many others, who I know get his same undivided, selfless devotion to their success.

Given this time to reflect, I see not only how Steve has changed my path, bolstered my skills, and supported my career, but I see the ways in which Steve's support shines through in my own teaching and mentoring. I speak frequently to my own doctoral students about Steve and the ways in which he did, and still does, support me. I try to emulate the type of apprenticeship he created for me, with appropriate support and challenge for the students who I mentor. I also try to incorporate humor and lightheartedness into my teaching, research, and mentoring efforts because Steve showed me that creating a comfortable space to learn and grow matters more than many of the other aspects by which we measure growth and success.

I continue to marvel at how Steve appears to seamlessly balance being a world-renowned scholar and a down-to-earth supporter, being a mentor and also a friend, being generous and accessible and still highly productive and efficient. And although I know Steve must be proud of his own accomplishments, his impact on so many researchers, teachers, and students is his greatest gift. I am incredibly grateful to have received this gift through Steve's mentoring, support, and collaboration. Thank you doesn't feel like enough, but thank you, Steve.

**Kausalai Wijekumar**


BY THE TIME I ADD AN INTRO, A FEW PICTURES, AND ENDING, IT WILL LOOK LIKE A GRAHAM MASTERPIECE!


THIS REFLECTION IS ABOUT THE INADEQUACY OF TRADITIONAL WRITING TO CONVEY MEANING!! BY ABANDONING REPRESENTATIONALISM, I'M FREE TO EXPRESS MYSELF WITH PURE FORM!


GRAHAM IS THE LEGENDARY TYPE! AS CALVIN WOULD SAY, "ANY IDIOT CAN BE FAMOUS."


SO LET'S BEGIN WITH A GLASS OF ICED TEA!! ICE AND TEA...


HOW ABOUT A 100 LEMON SLICES TOO!


AND QUITE A HANDFUL OF PEANUT M&M'S... AFTER A FEW HOURS AND A FEW BOWLS, I'M SO OVERSTIMULATED I CAN'T SIT STILL OR EVEN THINK STRAIGHT!


In all seriousness, I cannot express in words what Steve means to my family and me. He constantly lifts me up when I am down, promotes all my crazy ideas, is precise in his scaffolding, brings wisdom and brilliant insight to solving problems, and cheerfully brings humor into all situations. It is an absolute honor and privilege to call him friend.

## Michel Fayol

I haven't had the opportunity to meet Steve Graham in person often. At most two or three times. And besides, my practice of oral English would probably not have allowed me to have fruitful exchanges with him. It is therefore essentially through reading his articles and his meta-analyses that I became familiar with his conceptions and working methods. And I am infinitely grateful to him for all his writings. If I had been younger, I would have gladly asked for the possibility of going to his university to work for a while at his side. It is impossible to go back, unfortunately.

As researchers and pedagogues concerned with improving the performance of our pupils and students in the production of written texts, we owe a lot to Steve Graham. Firstly, for having conducted experimental research to better understand and better intervene with adults and students in order to improve written verbal production. Secondly, for having carried out or commissioned meta-analyses providing the scientific community with assessments of research achievements and

highlighting gaps or weaknesses of previous work. For all this and for his personal qualities, may he be thanked and honored.

## Bruce Saddler

The politician John Crosby defined mentoring as "a brain to pick, an ear to listen, and a push in the right direction." The best mentors may share guidance, motivation, emotional support, and role modeling with a mentee. They might also assist with exploring careers, setting goals, developing contacts, and identifying resources. The best mentoring roles continue over time, changing with the needs of the mentee. I share these ideas because they describe exactly my relationship with Steve Graham.

I met Steve when I was a classroom special education teacher struggling with teaching writing to my young students. One day I came across a flyer for a seminar offered by Steve titled "I wrote 10,000 words today." I attended hoping to learn the mystical arts of proper writing instruction and ended up with much more than that. At some point during his presentation, Steve advertised a doctoral program at his university and mentioned openings for new students. I remember imagining how wonderful an experience I would have learning more about writing from someone like Steve who seemed so supportive and kind. I applied and was accepted, thus beginning a journey into academe which has led to a personally fulfilling and successful life as a scholar, teacher, and mentor to other teachers and professors.

I can say unreservedly that without Steve my accomplishments would not have happened. From my first doctoral level seminar with him through the end of my career, Steve has been a role model for me; a brain, ear, and push. He literally taught me the craft of academic writing through the development of my dissertation and coauthored works with him. His revisions of my thoughts helped me grow as a writer, and his publications served as powerful models. But more than his ability as a writer, I appreciated how he interacted with students and colleagues. The relationships he formed and the lives he enhanced. I never met a man who has been so universally admired.

So Steve, as you enter retirement, may I humbly say thank you and may you be proud of the work you have done, the person you are, and the difference you have made.

## Linda H. Mason

### *Do You Have a Car?*

In 2000, I was ending my second year of doctoral studies at the University of Maryland and seeking a graduate research assistantship. It was time to embrace full time PhD study and to make the shift from being a special education teacher to

being an academic. Steve Graham had a position. This would not be the first time I asked Steve for funding. Drs. Graham, Karen Harris, and Debbie Speece had funded my masters' degree eight years earlier through a personnel preparation grant. This time was different. I was applying for a project coordinator position for the Center for Accelerating Student Learning (CASL) at the University of Maryland. Needless to say, I was nervous about the interview. Steve asked one question: "Do you have a car?" I do not remember my response, but it must have indicated that I did. I got the job!

On reflection, there is power in Steve's loaded question, and in the implications for what I was going to learn through teaching, research, collaboration, and friendship with Steve and Karen. First, however, I need to take a step back, as what I learned as a master's student informed my career as a special educator – specifically, the value of excellence in teaching. I learned the importance of systematic and explicit instruction and self-regulated learning. I learned this from my two special education methods teachers, Steve Graham and Karen Harris. What a foundation for a future researcher and teacher! Most importantly, this informed my practice as an elementary special education teacher. I applied what I learned, and my students made growth. Self-regulated strategy development (SRSD) was working for my students, and I wanted to learn more.

Did I have a car? Obviously, I could not move forward in doctoral studies without this foundational piece of equipment. What was I to learn? First, how to establish a solid foundation for conducting research – not just any research, but research held to the highest methodological standard; research grounded in theory. I learned that research is not easy, and that much can go wrong, but also that research is truly rewarding if one works hard at it. I also learned that scholarship is demanding work, and that one needs to develop a strategy for being successful. Steve's writing "hangar" appears to be always filled with planes ready for takeoff. Steve modeled and practiced how to be a productive scholar. I continue to be impressed at all that he has achieved. But, this is not surprising as he has always encouraged me, and others, to keep moving forward; "there is still a lot to be done!"

The most important gift that Steve and Karen have given me is the gift of collaboration and friendship. It has been a most enjoyable ride these past 30 years. Glad I had the car!

## Sharlene Kiuhara

### *"An Iced Tea with Extra Lemons on the Side" Retrospective*

As the corresponding author and former post-doctoral fellow of Dr. Graham, I, along with my coauthors, are honored to be included in the *Festschrift* to celebrate and recognize Dr. Graham's career-long scientific study of how writing can be used to support learning in various content areas such as mathematics, his development

of writing interventions for children with and without learning disabilities, and his extensive contributions to the field of special education and educational psychology. To commemorate this auspicious occasion, we dedicate our chapter to Professor Steve Graham.

I first met Dr. Graham in 2004 as a doctoral student in Special Education at the University of Utah when I interviewed him for one of my doctoral seminar courses. The assignment was to interview a distinguished scholar in my area of research interest, which was writing strategies instruction for secondary students with learning disabilities. At that time, writing scholars who were interested in understanding the composing processes of secondary students with and at-risk for disabilities were few and far between. Dr. Graham shared a vision of writing research that would involve an interdisciplinary and international view for advancing scientific inquiry in writing theory and pedagogy. Implementing changes in antiquated policy and practices were signaled in a series of reports published in the 2000s by the National Commission on Writing in America's Schools and Colleges (NCOW: 2003, 2006). Dr. Graham served on my supervisory committee and was instrumental in guiding my future research agendas.

During my doctoral program, Dr. Graham and I collaborated on two studies that were influential in expanding the evidence base for middle and high school students with disabilities (Kiuhara et al., 2009; Kiuhara et al., 2012). These studies set the stage for my subsequent work in the professional writing development of classroom teachers (Gillespie & Kiuhara, 2017), explicit writing strategies instruction (Kiuhara et al., 2012), and examining explicit writing instruction in learning mathematics (Hacker et al., 2019; Kiuhara et al., 2020).

During my time as a post-doctoral fellow under Dr. Graham's mentorship, I worked closely with him and his advisees (i.e., Michael Hebert, Amy Gillespie Rouse, and Paul Morphy who have become prominent writing scholars and friends). We worked together on a project that surveyed high school language arts, social studies, science, and mathematics teachers, aimed at understanding how they used writing to support students' content learning (Gillespie et al., 2014). With the addition of mathematics teachers to the participant pool, we learned that the mathematics teachers employed writing activities to develop students' procedural and conceptual understanding of mathematics. Notetaking was encouraged to promote students' procedural understanding, and extended writing was used as a tool for problem-solving and developing their conceptual understanding. Because students with and at risk for learning disabilities were not keeping up with their same-aged peers in mathematics, this study laid the groundwork for one of my own research agendas, which was to understand whether explicit writing strategies instruction could be a way for these students to thrive in their procedural and conceptual understanding of mathematics.

Finally, I was fortunate to work with Dr. Graham on a meta-analysis that extended Bangert-Drowns et al.'s (2004) seminal work on examining the effects of different writing activities on learning for students in Grades 1–12. Graham et al. (2020) found that writing can enhance learning in science, social studies, and mathematics, with the average weighted effect size of g = .30, which did not significantly differ

across content, grade, or features of instruction, writing activity, assessment, and study quality indicators.

Writing intervention research focusing on students with and at risk for learning disabilities continues to expand nationally and internationally because of Dr. Graham's mentorship and generosity of time. I am grateful to him for paving the way for current and future scholars. Over the past 20 years, writing has become less of a neglected "R" thanks to Dr. Graham's yeoman efforts.

## Naomi Weintraub

I first met Steve in a course I took at the University of Maryland (UMD). I was an occupational therapist who had just relocated to Maryland from Israel because of my husband's training at the NIH. I intended to pursue my PhD studies in occupational therapy at Boston University, but while waiting for my husband to finish his training, I thought I would take transferrable courses at UMD. The first course I took was taught by Steve. This course was one of the best courses I ever took. Steve found the correct balance between direct instruction and engaging the students in discussion, integrating his own educational experience with theory and research evidence. After this course, I decided to stay at UMD and do my PhD in special education, focusing on learning disabilities.

When I started my doctorate studies, I was working at the school system in Maryland, with students with learning disabilities. Most of my work-load included students with handwriting difficulties. At the time (early 1990s), this was a relatively new area of practice in education and occupational therapy. While looking for research in the area, I noticed that most of the publications were from the early 1980s, and many were written by Steve. At the end of one of the classes I approached Steve, and told him that I noticed that he used to be interested in the area of handwriting, but that I hadn't been able to find recent studies. Steve answered that he is still interested in handwriting research, but has not had a student who wanted to focus on this topic. I was very excited, and asked if he would be willing to be my advisor, and that we would focus on handwriting. I was delighted that Steve agreed!!!

Being Steve, with his thorough and systematic way of working, he suggested that we start with reviewing the literature: "… but let's make it wide. Let's summarize the literature from 1980, and not only about school age children." And so, I found myself involved in conducting a review of over 200 manuscripts together with Steve. This turned out to be a very important experience. Much of my in-depth knowledge in the area of handwriting came from reading the articles for the review. Moreover, Steve taught me how to write a systematic review, which over the years was very valuable.

Another priceless skill I have learned from Steve was "Time Budgeting." In one of the meetings I had with Steve while working as his research assistant, I suggested to Steve that we write an article on a certain topic. Steve said he thought it was a

good topic and that we should start writing in 6 months from the meeting. I must have looked very puzzled, because Steve noticed, and explained. He said that in the next 2 months he would be working on a certain article, and then, in the following 2 months, he would be working on a second article, and so on. “Therefore,”, he added, “in 6 months we can start working on the article you suggested.” I still was not convinced, and then Steve asked me if I had heard of “time budgeting.”. Obviously, my answer was “no.” Steve explained his strategy of time budgeting, carefully planning his work load to coincide with his calendar. This explanation was so logical – it actually changed my entire planning strategy both in planning my own family's schedule as well as work. In fact, in all my courses I teach students what time-budgeting is. I do that because I believe this is one of the most important survival strategies for busy people.

Finally, an anecdote from my meetings with Steve, which I vividly remember. Sitting in Steve’s office discussing my thesis or work that Steve wanted me to do – and Steve had a box of paper clips next to him. Every meeting he would sit with a box of paper clips on the desk. Each time he would take a paper clip in his hand, open it, play around with it and put it down. By the end of the meeting, Steve created his own “art.” This is what it often looked like.

In summary, often, the best things in life happen when they are not planned. This was my experience in being Steve’s PhD student. In fact, Steven was my academic life-time mentor without knowing it. So much of what I know, and how to do things, I’ve learned from the years I spent at UMD under Steve’s mentorship.

So, dear Steve, words cannot express how much I appreciate all that you have taught me and have done for my academic career.

Cheers.
Naomi Weintraub

## Amber B. Ray

### *We Met in A Bar*

If you ask Steve how I became his doctoral student, he will tell you we met in a bar. I was a high school special education teacher living in Portland, Oregon, at the time. My friend, who was also a teacher, was interested in getting his doctorate and had been in touch with Steve. My friend told me Steve and Karen were going to be in town for a conference, and he was going to meet them at a restaurant downtown Portland to discuss Arizona State University's (ASU) PhD program. My friend asked me if I would come along.

So, I went downtown with him and went to the bar. Soon Steve and Karen arrived. We all sat down and enjoyed drinks, Steve of course had an iced tea with many, many lemons. I had planned to just be there to enjoy a cocktail and good conversation. However, as the evening went on, I became more and more interested in the PhD program at ASU and the opportunity of learning from Steve and Karen. Steve reassured me I would be a perfect candidate, and as you know, he can be quite convincing.

As I thanked them when we got up to go, in true Portland fashion, it began to rain. Steve and Karen had walked to the restaurant, so I offered them a ride back to their hotel. In this brief ride, I learned how friendly and down-to-earth they were as I drove them in my car filled with Special Olympics sports equipment, teaching supplies, and much more. By the end of the evening, not only was I excited about this new possibility, but was also assured that he would be the ideal advisor for me. I was ready to go home and complete my application to the doctoral program.

And that is exactly what I did. I applied to PhD program at ASU and moved to Tempe a few months later. Throughout the years the story of how I met Steve gets increasingly embellished. Steve is a fantastic storyteller and has the ability capture the interest of any audience. Steve enjoys telling people that we met in a bar. He even tells that story when serving as a job reference for me.

Steve has had a major impact on the trajectory of my career and life. I had the pleasure of Steve being my advisor, which allowed me to conduct research with both Steve and Karen throughout my time at ASU. My life has never been the same after having the privilege of being Steve's doctoral student. He is such a prolific writer and is always working on so many research projects at a time, but he always makes time to meet with me, both now and then. During my doctoral program, many of those meetings occurred while walking across campus to get iced tea and cookies. Even years after graduating, Steve still makes time to meet with me whether it is via Zoom or at a conference. I am forever grateful for his willingness to let me work on many projects and co-author articles with him. I have learned so much and continue to learn about the field of writing through him. I truly believe I am the scholar I am today because of Steve's example of what it means to be a great researcher.

Beyond his brilliance, Steve is an incredibly humble and kind person. He knows how to make anyone feel welcome. While at ASU, I enjoyed when Steve and Karen

would host dinners at their house with groups of doctoral students. Even now, I always look forward to attending Steve and Karen's "family" dinners at conferences where I get to spend time with all the other great researchers who also had the pleasure of having them as their advisors and mentors. This lineage of doctoral students is quite remarkable. Even as a full time professional, Steve is my go-to. He is very thoughtful and takes the time to help me network and introduces me to other researchers while at conferences. He truly cares and takes steps to make sure I am successful and able to help carry on his legacy.

I could never have asked for a better advisor; from Steve's educational expertise to his sense of humor, I feel my career and my life will always be a reflection of Steve's influence. I have the utmost respect, admiration, and gratefulness for Steve, a great man, an accomplished researcher, and the best mentor anyone could have. Words can never completely convey how much I care and how much Steve means to me. I am proud to be able to say Steve has changed my life and career and I will always be extremely thankful. I look forward to the next time I get to head to the bar and once again be inspired by Steve.

## Debra McKeown

In my first year as a doctoral student, I made a blatant error recording student responses on an assessment. This resulted in, as the project director, gently shared, "just an empty cell in the table, not the end of the world," but I was mortified. Shortly, thereafter, when Steve Graham and I were working on a meta-analysis for elementary writing interventions, I found a few discrepancies and incorrect calculations in the studies. We were discussing just how many things can go wrong in a study. "*It's just error*," Steve said nonchalantly walking out of his office.

*JUST error*? Error was like variability and was, as Mark Wolery so eloquently expressed, "the sucker we're trying to get our hands around and choke; it is the enemy, one of the demons that torture our souls." To Steve, the black Chuck Taylor wearing, basketball playing, writing researcher, error was "*just error*." Across the years as I learned more from Steve about conducting research rigorously – with unimpeachable design and implementation – and analyzing results meticulously, he still always faced the inevitable imperfections with the same unbothered acceptance. *It's just error*.

When we read the research of others, it can read like nothing ever goes wrong, everyone cooperates, and things go just as planned. This is what we believe when we first design our own studies. Then, when things do not go as planned, it can lead to feeling like we are not good enough to be researchers. Steve offered perspective.

The thing with Steve is he is a dispassionate observer of the data. He does not try to make the data something they are not. He has objectivity because of his acceptance of reality and that is powerful. It is the closest thing we can get to being part of a study and remain unbiased. Whatever the data say, they say, and we learn from it. I am not suggesting other scientists influence studies in undue ways, though I can

see it may sound like that. Steve just does not care about which way the data fall – he just wants to see what happens as a result of a well-run study. Either way, there is a lesson, and it is the lesson that drives him.

"*It's just error*" is not dismissive; error is requisite. There will always be error, and there is a humility in making that known, accepted, and centered. On the surface, it may come off as a flippant or careless remark. Yet, what it contains is a state of mind that privileges objectivity and sanity over unattainable perfection. And it is the core to ethics; there is little to compromise one's ethics if you are willing to make a friend of the error. One may think this approach is incompatible with being a great scientist, but I believe the opposite to be true. By recognizing that research designs can be compromised in the unreliable context of public schools, mistakes will be made, and that this process of science that we have all committed our lives to is handled by humans in human systems, Steve framed science as what it is – a process, not a god.

Anyone who conducts research in schools learns quickly that there is no perfectly controlled experiment. We cannot control all the variables that impact student behavior and learning. This can be anxiety-producing. Yet, the reality is, research conducted in a lab is quite different from the chaotic, unpredictable context of schools. So, researchers trained to conduct high quality intervention research in school settings must navigate as much of the variability as possible and also accept that errors will still happen. Steve is good at acceptance, and he is good at modeling acceptance…of error in other people's studies and error in his own studies.

Steve has shaped the world of writing in innumerable ways – through his questions, research, teaching, editorships, speaking engagements, and of course, his writings. And along the way, Steve made doing research feel achievable for so many, myself included. What he left me personally is the objective lens that lets me recognize when something is important and when "*it's just error*."

## Alyson A. Collins

### *Working Hard or Hardly Working?: Reflections on Our Research Collaborations with Steve Graham*

Many people ask about how it is to work with Steve Graham. I always answer, "Steve is the greatest mentor and colleague. He is genuine and truly loves what does. Steve is the absolute best!" When I decided to return to graduate school, I had a teacher mentor who suggested that if I was applying to Vanderbilt, I should look at the work of Steve Graham. This was my first introduction to Steve's research, and it came at a time when I was seeking more resources to help my students with writing. Curious, I reached out to Steve Graham and Karen Harris and inquired about the PhD program. It was clear from our first conversation that Steve and Karen were amazing advisors who greatly invested in their students. Learning from them was a

once in a lifetime opportunity, and I was humbled for them to welcome me into their academic family.

Although Steve moved to Arizona State University after my first year (and I was devastated when he left), he continued to informally mentor me throughout my graduate studies. After I interviewed for my current position at Texas State University, my colleague Stephen Ciullo asked if I was interested in conducting more research in the area of writing. We soon scheduled a meeting with Steve Graham to discuss the possibility of running the study we reported in the current chapter. Little did we know the study would lead to three federally funded grants to date, and a number of writing research projects together. The project in the current chapter is particularly memorable because it was the beginning of what would become the best professional team that I have ever been a part of—one I am grateful for every day.

Over the years, Steve Graham has taught me three things that I will forever remember. First, Steve recommends to find something you are passionate about and figure out a way to make it your career. Many people would say Steve Graham never stops working. However, if you know him, you also know he does not work because he has to. He does not work because he is fearful of falling behind or not being good enough. Steve Graham works because he genuinely loves writing research and everything that encompasses it.

Another attribute Steve Graham has shared is the gratification he finds in brainstorming new ideas. He is open and accepting of academic discourse and dialogue that considers different viewpoints, and in fact he welcomes it. Moreover, Steve has a wide knowledge base of different types of research methods as well as an excellent memory of prior research. To this end, he strategically applies the most rigorous methods in his research. His ability to explain the rationale behind specific methodological procedures, while considering different approaches makes him an incredible teacher and researcher.

Finally, Steve Graham has showed me how to make the work fun through meaningful partnerships. As a team, we laugh and joke around with each other. (However, Steve will also remind you to refocus and get back down to work!) After our weekly Zoom meetings, I have had people comment about how it did not seem like a "work call" because of how much laughter occurred. Steve not only adds humor to his work, but he also invests in getting to know his colleagues as individuals. He cares about who they are and engages in conversations to truly know them. Steve delights in his collaborations and his collegial friendships in the field.

## Stephen Ciullo

Steve Graham has been a teacher, colleague, and mentor to countless teachers and scholars. One of the most important experiences of my career happened in 2016 when I realized that Steve saw potential in me as a mentee and as a future colleague. The partnership that Steve, Alyson Collins, and I developed has been fulfilling and

fun. The studies that we are conducting have the potential to provide teachers with new tools to effectively teach writing to students who receive special education services. Additionally, the countless laughs we share while working (sometimes at my expense, especially when I order too much food) are also irreplaceable.

I've learned countless important lessons from Steve. However, two specific lessons will have a lasting impression on me. First, Steve has an insatiable intellectual curiosity that drives him to conduct important research. While his scholarly productivity is beyond impressive, his attention to detail on each project allows him to effectively communicate the findings to the broadest possible audience. His commitment to research quality taught me that every study or article must have implications that can benefit teachers and students. In essence, every endeavor in our career should be devoted to empowering students to become better writers.

Second, Steve taught me that working hard and having fun can (and should) occur simultaneously. Steve Graham has incredible concentration and persistence with any project that he commits to. However, Steve's tenacity doesn't interfere with his sense of humor (such as him reminding me that I can't keep eating as many calories as I get older and that "Freddy" constantly needs to be fed), or his love for traveling and hiking. In sum, you can make a difference as an academic, while also living a healthy and exciting life if you truly enjoy what you do, and if your goals are motivated by helping others.

I am thankful that Steve Graham took a chance on a "reading guy." I look forward to our continued collaborations, sharing exciting findings with teachers, and many more laughs in the years to come!

## From Both of Us

We each wrote separate tributes and reflections about our collaborations with Steve Graham. Interestingly (although not surprising), the same themes emerged around his intellectual curiosity, strong work ethic, endless energy, and sense of humor. There is no doubt our lives are forever better because we have had the great fortune and privilege of learning from one of the most prolific educational researchers in the world. Steve Graham, thank you for being a part of our dream team!

## Gary A. Troia

As some readers of this volume will know, Steve Graham was my doctoral degree advisor and mentor while I was a student at the University of Maryland in the 1990s. I had come to the special education PhD program at UMCP by way of a serendipitous admissions hiatus in the communication sciences and disorders doctoral program on campus. I had been ready to dive headlong into advanced research preparation after completing a master's thesis on phonological processing in

typically developing young children and working several years in schools and private practice (and relishing the diagnostic and assessment portions of my work that are so much like the research enterprise), but I was disheartened when my alma mater informed me the program that I had applied to was temporarily closed for new admissions. So, I quickly revised and diverted my application materials to the most obvious alternative, special education—a most fortuitous decision—and soon met Steve during my admissions interview. I recall only bits of our conversation at that time (but certainly Steve's probing to see if I had interests in writing interventions), though I recall most clearly his affable demeanor and his assurances that I would be a good fit with the program.

Near the completion of my doctoral studies, I had a methodological critique of phonological awareness intervention literature published—one that I was well suited to do on my own given my professional background, prior research experience in the area, and the fact that I had been working several years with colleagues in the Hearing and Speech Sciences Department at UMCP to develop a phonological awareness intervention program. Unbeknown to me at the time, several prominent scholars had contacted Steve to inquire about his young upstart student (my words) who was zeroing in on flaws in their studies. Steve advocated on my behalf and assured them I was acting as an independent scholar trying to do important, rigorous work, work made possible by his mentorship and that of many other faculty in special education and communication sciences and disorders. He later assured me that all was good.

After graduating and working a few years at the University of Washington, our relationship naturally transitioned to one of colleagues occasionally working on projects together (of course, Steve still was and is informally a mentor), and this is when his jocular and cheerful nature really became most obvious to me. Being referred to by him as "grasshopper" (a reference to a mentor-mentee relationship between characters in the 1970s television series *Kung Fu*), seeing his clever use of *Peanuts* comic strip characters or ideas from the *Hitchhiker's Guide to the Galaxy* to convey information about writing, hearing his tales of misadventures and pranks when he was younger (and some more recent), laughing with him at conferences and meetings, often when he is self-deprecating—one cannot possibly feel glum or dispirited when Steve is present.

Steve is an accomplished, respected, and impactful academic recognized inside and outside the United States for his scholarship; hence, this book in his honor. The rigor and volume of his work (certainly he should be coronated as king of writing research meta-analysis at this point) is well known. What also stands out about Steve's work ethic is his graciousness toward other scholars, me included, and the greater education profession. There are many individuals and groups in K-12 schools, universities, non-profit organizations, and government agencies who can attest that Steve has given much of his time and expertise across the decades, often just informally. He is sought out not only because of his masterful knowledge of the literature and research methodology, his ability to size up the practical implications of actions and associated reactions, and his devotion to the profession and those who it serves, but because he genuinely cares about others' success and wants to

help them attain their highest levels of excellence. Some recent counsel Steve offered to me regarding keeping a strong focus on theoretical frameworks for situating my scholarly endeavors, regardless of the applied nature of the work, has pushed me to be much more mindful of the importance of theory in all kinds of writing research, and I firmly believe this has indeed helped me produce better research and can help the field be more grounded.

Affable. Assuring. Jocular. Cheerful. Accomplished. Respected. Impactful. Gracious. Genuine. These are just some of the traits that Steve possesses that earn him great distinction as a scholar, friend, and person. Knowing Steve, he might read this list of words and contemplate if he could arrange them into a mnemonic acronym and if he couldn't, I'm sure his advice would be to pick synonyms for some of these words to make it easier to create one. His mind is always working to produce better teaching and learning, and he is always willing to offer good advice.

## Tien Ping Hsiang

### *A Fantastic Journey of Studying Writing Instruction from a Cross-Cultural Perspective*

As a researcher who studies the teaching of Chinese writing, I have been reading Steve and Karen's publications since 2005. When I was a master student struggling on choosing a research topic, my supervisor, Dr. Der-Hsiang Huang showed me one of his papers which focused on self-regulated learning. Steve and Karen's articles were referred by Dr. Huang. The articles not only helped me to recognize "Self-Regulated Strategy Development (SRSD)," but also inspired me to study the teaching of Chinese writing from a new perspective.

Later, in my doctoral dissertation, I studied fifth grade students' Chinese writing process and developed a SRSD Chinese writing program for my grade 5 students. I found that although Chinese writing style is not the same as English writing, SRSD and process writing could improve students' Chinese writing motivation and performance.

In order to explore more possibilities in studying writing instruction, I finally plucked up the courage to contact Steve in 2010. In the email, I requested a meeting with Steve during his summer vacation. I couldn't believe that I received his prompt reply. Steve agreed to meet me in his office at Vanderbilt University. I remember I brought one of Steve's book and requested Steve's signature.

The cooperation between Steve and I has continued since 2010. The topics we study range from teachers' Chinese writing instructional practices to teachers' beliefs about teaching writing. Recently, we have focused on teaching Chinese characters which is the core ability of Chinese writing. Our cooperation can keep going because of Steve's diligence and eagerness to learn more.

Steve made a great contribution in each of our academic reports. In order to know the differences between English and Chinese writing, he has reviewed much literature on Chinese writing and teaching Chinese writing. We decided research purposes, methods, and instruments all together. While I spent a lot of time on translating Chinese concepts into English, Steve did his best to introduce what Chinese writing is to the international academic society based on his understanding of Chinese writing. As a stranger to Chinese writing, Steve knows what foreign researchers want to learn from our studies. As an experienced researcher in teaching writing, Steve can always glean much more from the data than I can. Whenever I had an idea about writing research, even though the study was not easily done, Steve always said: Let's do it!

What I have learned from Steve is more than academic research. As a college instructor who teaches teaching Chinese language arts in primary schools, when I wondered why I need to publish English papers, Steve said: Then I can know what you have done in studying Chinese writing. When I asked Steve why he would cooperate with co-authors, he said: Most of the time, I learn and gain from the process of cooperation. When I was not sure how much time I should devote to working, Steve let me know that family always comes first! Steve is just like my supervisor of life, teaching me to live in the world from a broader perspective.

It is my honor to work with Steve! He is an innovative learner, honest researcher, and trustworthy leader. Steve has made great contributions to studying and introducing Chinese writing instruction to the world during the past ten years. It is a fantastic journey of doing writing research with Steve from a cross-cultural perspective. Both of us have learned to appreciate the similarities and the differences. The most important thing Steve has taught me is, we do research because for the love of it. Love to learn and love to add positive power to the world. The seeds Steve has planted in the garden will blossom beautiful flowers!

## Charles "Skip" MacArthur

### *Signs of Steve*

A plate full of lemon slices
A glass of iced tea;
Show that Steve is here.

The ball swishes through the net, not touching the rim;
Famous for his set shot not his jump,
His black Converse high tops follow him off the court,
Peeking out beneath suit or robe.

A smile full of warmth,
Greetings for all, reflect the truth
Of his generosity to colleagues and students;

How often did we gather with students and colleagues,
For dinner at PCRC?
How often did he listen and offer advice and opportunity?

A heavy meta concert,
Thunderous chords and fancy guitar licks,
Meta-analyses rain down;
'Til one can't write a paper without citing at least one.

The ultimate sign – SRSD,
Created with his wife and colleague, Karen,
Woven from a thousand threads of theory, practice, and hard research,
Spread around the world
To serve teachers and students of all ages.

Maybe you'll spot him on one of his daily walks,
On the beach in San Diego,
Or the hills of Arizona,
Or the streets of Paris,
Or anywhere.
Walk with him and talk; you'll learn something.
Fortunate I was to meet him when we were young scholars
To work together and learn with him for so many years.

## Rui A. Alves

### *Seeing the Whole Writing Research Galaxy from Steve's Shoulders*

I can't remember the first time I heard or read about Graham, Steve Graham. But it surely must have been circa 2000, when I started learning something about writing cognition. Graham, for me, was foundational not only in conceptualizing writing but also in helping to establish myself as an international writing scholar. Let me then focus on these two gratitude aspects of my own development as researcher.

As with writing, there is hardly a single influence in developing one's own conceptions. As wisely modeled by Graham, it takes a community to achieve writing and it takes a village of sages to educate a researcher. In my case, I am indebted to giant sages working across both sides of the Atlantic. From the US side, Steve Graham, Dick Hayes, Ginger Berninger, Ron Kellogg, Chuck Bazerman, and Malt Joshi. From the Europe side, Michel Fayol, Sven Strömqvist, Liliana Tolchinsky, Piero Boscolo, and São Luís Castro.

The establishment of writing as an international research topic is a recent endeavor. One can argue that its foundations were laid in the last two decades of the 20th century and were hugely facilitated by globalization, by the explosion of digital communication technologies, most notably the internet, and by Steve. Among

the giants cited above, Steve is arguably the one who contributed the most to settle writing as a lively and respected international topic of research. One needs only to look at the dozens and dozens of his scholarly contributions, to the legion of his collaborators and coauthors spread over the globe, and generally to the global influence of his groundbreaking research work. One can easily check his tremendous influence over the globe by plotting citations to his work over the world map, as depicted for instance in the metrics dashboard of his profile in Web of Science (https://www.webofscience.com/wos/author/record/2023278). Steve is surely the modern giant of writing research and his influence legendary as proven by the current festschrift volume.

It was not up until 2008 that I finally met Steve in an international writing conference held at the University of Poitiers. From Steve and his keynote, I was immediately impressed by his oratory skills, by the value of effective communication to achieve influence, and by the importance of aligning as many sources of evidence as possible to help transform educational practices. In many ways, his series of meta-analyses are the best current effort to inform and improve writing instruction. From that first lecture I attended, I retain Steve's search for alignments in writing research, to discern patterns and to inform further research and writing instruction. In that occasion, Steve was aligning support to evidence-based practices from three sources of inquiry: meta-analyses of experimental studies, meta-analyses of single-subject design research, and meta-synthesis of qualitative studies of exceptional teachers of writing. Crucially, I was strongly moved by his joy when the three sources of evidence did align, supporting an instructional practice! Such alignment sounded almost magical and mystical as when a scientist discovers something about the world or when a child watches the three stars in Orion's belt. I believe that it is this joy of aligning that fuels Steve's incessant scientific queries. Furthermore, this joy nurtures the interest and intelligence of those many who collaborate with him. Ever since that first meeting, he generously engaged me in several of his editorial projects, many of my students learned with him and his research, I asked his support for many scholarly activities, and in a way, he became Orion to me. Indeed, from the shoulders of Orion, we can see the whole galaxy of writing research. Thank you so much for the ride, Steve!

## Xinghua Liu

### *To the Lighthouse*

I keep on seeking purposes for what I am doing and looking for answers, just like a small boat drifting in the ocean looking for a lighthouse to show the way. In 2008, after teaching English as a foreign language at my university for three years, with eagerness to understand my students and to explore what I could bring to the classroom, I embarked on the PhD journey in UK, studying applied linguistics with a

special focus on foreign language education. In 2015, three years after obtaining my PhD, I found myself increasingly bothered by two questions:

First, how I could help my students better learn and become autonomous language learners at the end of their two-year university foreign language program?
Second, what were the purposes of my research and how I could make meaning out of it?

I looked around for a teacher and mentor to show me answers to the two questions.

The opportunity came that year when I was awarded a fellowship from the China Scholarship Council, enabling me to conduct one-year post-doctoral research abroad. I contacted Steve Graham and asked whether I could join his team on writing research. I wanted to work with Steve because I was interested in learning SRSD and running meta-analysis (when I was at ASU, I learned that Steve's wife, Karen Harris was the mother of SRSD and I was also lucky enough to join her SRSD projects to learn how to deliver the program). My one-year stay at ASU turned out to be most memorable… to work with Steve and Karen has proved to be the best decision I have ever made.

Steve is a teacher with whom any junior researchers may dream working. He trusts junior researchers and allows them to take responsibilities. I can still recall the very first meeting Steve and I had in the office. He checked with me what I expected from my stay at ASU. I said I wanted to know more about meta-analysis, and he nodded. Later on, I was involved in two meta-analysis projects, namely "the Impact of Reading Interventions on Writing" and the "Effectiveness of Literacy Programs Balancing Reading and Writing Instruction." Steve has been generous in sharing. He spent time in demonstrating how to look for key information for meta-analysis, how to code variables, and how to run statistical analysis. I am truly indebted for the apprenticeship with Steve.

Steve is kind and fun to work with. Whenever he came to the office room in the morning, with sun glasses on and a flip phone in his hands, he would peep through the glass window of my room, smiling, and asked "Hey, how are you doing, Kevin?" At noon time, he would pop up, knock at my door, smile at me, and say "Kevin, cookie time." He knew those cookies were my favorites. We then walked together from the office building to the snack shop on campus. In the snack shop, Steve would have one to two large cups of iced tea, grasp a handful of lemon slices, and squeeze them into the iced tea. Along the way to and returning from the snack shop, we talked about research projects, paper writing, headlines, and children. The "cookie conversation" was the most inspiring and relaxing moment in a day.

Steve has also always been ready to help. As I was interested in corpus studies, I once asked Steve whether it was possible to obtain some corpus data from a third party testing organization. Without a single moment of hesitation, he smiled at me and said he would like to have a try. In the following days in the office, I heard Steve making phone calls to enquire about the possibility. He also spent time sending emails and filling application forms to request the data. Though eventually we were denied the access to the data, I was thankful for the time and energy Steve had

devoted to my interest. In 2018, when my wife had medical emergency and I asked him about treatment options in the USA, he was quick to gather many useful information and took all the way to reach out to his friends for further assistance. A friend in need is a friend indeed. Steve is a man of utmost kindness and trustworthiness.

Steve is a superhero for those who have ever worked with him, and I have been wondering how he managed such high motivation for work and life. The opportunity to ask about this question came when we drove on the highway. I plucked up the courage and asked "Steve, would you mind sharing why you have been so motivated?". Steve smiled, as always, and gave me five reasons, out of which two helped solve puzzles I had over the past years. One was the love for children. Through his research, Steve said he could help children better learn to write, particularly for those with learning difficulties. Another was his passion for advancing science so that he could better help people in need.

As a Chinese saying goes, art is long, but life is short. Whenever I feel the bottleneck, I would look to Steve, with his true passion to scientific exploration, the genuine love to students, and a loving heart to family and friends. The one year with Steve at ASU is life changing to me. Words cannot express enough my heartfelt thanks to him. To the lighthouse, Steve. From the bottom of my heart, I wish him peace, prosperity, and good health in the years to come.

## Michael Hebert

### *The Emperor and the Bad Penny (and Other Fables)*

Steve is many things. He is a friend, mentor, advisor, boss, collaborator, husband, and father. He is a writer, researcher, theorist, psychologist, and editor. He has also had many nicknames: Graham Cracker, Super Graham, Stevemo. To me, he will always be The Emperor.

In honor of Steve, I wrote three short fables of my time as his doctoral student.

#### The Professor and the Lemonade

The professor was always generous with his doctoral students. He was generous with his time. He was generous with opportunities. He was generous with his laughter.

However, the professor also had high expectations. He modeled hard work and persistence. As long as he had iced tea and lemons, the professor could work for hours on end. Although his students worked hard to impress him, they would sometimes dread the calls into his office. They knew it meant an unrelenting day of reading and coding.

When even the most energetic students would tire and get drowsy, the professor pressed on. The students never complained. They did not want to be disrespectful of the time the professor set aside for them. However, they thought distracting him with a prank might give them reprieve for a day.

Once, they put pictures of the professor and his wife on a video card. To their surprise they heard nothing about it. The professor went on as usual. This led to an escalation in the students' pranks. However, they still heard nothing and still fell asleep while coding.

Finally, the students put a box of Life cereal, a lemon, and a glass on the professor's desk.

*Moral: When life give you lemons (or when your pranks are lemons), make lemonade.*

### Stevemo and the Crayons

Steve's doctoral students sometimes lost pens in his office. In fact, the students noticed that they often lost pens when Steve was around. When Steve stopped by their office, he sometimes asked for a pen to write something. He then distracted them with hard questions before absconding with the pen.

One day, the students finally caught Steve in the act. To their surprise, Steve just laughed and said, "I'm liberating the pens!"

The next time Steve left town, the students devised a prank to get him back. They snuck into his office and replaced all of his pens with crayons. They also put a picture of Elmo on one of his books and retitled it, "Writing Better with Stevemo."

Upon his return, the students went to his office to code studies with Steve. As they hoped, it wasn't long before Steve needed a pen. The students held their breath as Steve reached for his pen cup. However, Stevemo only briefly looked at the crayon he pulled out. Then, he used it to write.

*Moral: You can take Stevemo's pens, but you can't take away his motivation to write.*

### The Emperor and the Bad Penny

One day, the professor declared to his students, "You should call me The Emperor."

The students laughed, but the professor persisted. "Yes, The Emperor has a nice ring to it. As my first act, I decree that my students shall call me The Emperor from this day forward."

This went on for quite some time. After several weeks, the students came up with a plan. They created t-shirts with the professor's face on several different emperors, including emperor penguins. The students even included The Emperor's wife in some of the pictures.

The students then attended a party at the Emperor's house. About 30-minutes through the party the students revealed their t-shirts. It took the Emperor another hour to notice.

Not long after, The Emperor started calling one of the students "Bad Penny."

*Moral: Be careful what you joke about with your students!*

## *One Big Lesson*

All of these fables capture the biggest lesson I learned from Steve: the work is important, and the writing is most important, but it should always be done with kindness and good humor. If we can always remember that, we will multiply our impact on the writing community through all of the relationships we build. That's Steve Graham's legacy.

# Contents

# Contributors

**A. Angelique Aitken** Assistant Professor of Education at Penn State, studied with Steve Graham at Arizona State University where she earned her PhD. Currently, she is developing POWER, an SRSD framework for use in inclusive social studies contexts. Angelique first met Steve at a recruitment event for Peabody College and appreciated his affable nature, interest in writing, and the fact that he wore Chucks high tops. A few days later Angelique saw Steve referenced in the lifestyle magazine, Real Simple (impact factor unknown) on what your handwriting says about your personality and thought, "wow, this guy seems legit, I guess I will go to Vanderbilt."

**Rui A. Alves** is Associate Professor of Psychology at the University of Porto. His main research interests are the cognitive and affective processes in writing, which he studies using experimental methods and logging tools. He and his team have studied the development of text production using an own–developed handwriting logging software (HandSpy) and developed many evidence-based writing interventions addressing cognitive, motivational, and self-regulatory aspects of writing. He serves on several editorial boards and is associate editor of the journal *Reading and Writing*, Springer. Currently, he is the interim coordinator of the European Literacy Network, ELN, and serves on the executive boards of ISAWR and EARLI.

**Ana P. Azzam** is a psychologist graduated from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio). She holds a master's degree in Neurocognition and Language from the Faculty of Psychology and Educational Sciences of the University of Porto (FPCEUP). Currently, she continues her studies as a PhD student at FPCEUP, also collaborating with research projects, and working as an external collaborator at the University. She is passionate about writing, working with data analysis, and studying linguistics related to mental health topics.

**Jenna Basile** is a doctoral scholar in the Special Education program at George Mason University. Her research interests focus on writing instruction and teacher preparation for special education. She serves on various research teams, most

notably as a graduate research assistant on two USDOE funded research grants examining writing instruction. Previously, she worked as a special educator in the Fairfax County Public School district in Northern Virginia and earned a Master of education at George Mason University as well as a Bachelor of Education at the University of Rhode Island.

**Pamela Shanahan Bazis** PhD, is an assistant professor in the Department of Special Education and Communication Disorders in the College of Education and Human Sciences at the University of Nebraska–Lincoln. Her research and scholarship are aimed at lexical literacy skill instruction and intervention. She has served in various instructional and supervisory roles at the K-12 level, including classroom teacher, reading specialist, instructional coach, and district dyslexia coordinator. Pam also provides professional development to elementary teachers on effective reading and writing strategies. Her research experiences include qualitative, single-case, mixed-methods, and randomized control group designs.

**Virginia Berninger** (PhD Psychology, Johns Hopkins University) is a research and clinical psychologist who was on faculty at Harvard Medical School (2 years) Tufts School of Medicine (3 years) and University of Washington (30 years) where she advised PhD and masters students, taught courses, and served as principal investigator for NIH grants on typical literacy development and intervention for low achieving readers and writers (1989–2008) and a multidisciplinary learning disabilities center (genetics, brain, diagnostic, and instructional research) (1995–2006, 2011–2016). Now Professor Emerita, University of Washington, her research resulted in 339 publications and was translated into assessment and instructional tools.

**Pascal Bressoux** is Professor of Educational Sciences at University Grenoble Alpes. He is honorary member of the Institut Universitaire de France. His research focuses on the relationship between teaching practices/devices and students' acquisitions. He specialized in longitudinal studies and multilevel modelling. He has been working as an expert in the Technical Advisory Group and in the Questionnaire Expert group for the OECD PISA, TALIS, and Starting Strong surveys.

**Megan C. Brown** PhD (they/them), is the director of the Global Center of Excellence for Research, Analytics, and Data Science at Starbucks. They have expertise in leading mixed method teams, resolving business challenges and opportunities with many methods, and sharing insights from extremely technical solutions with non-technical audiences. While in academia, they studied how dialect variation affects language and literacy development.

**April Camping** is a postdoctoral research associate at Texas A&M University in the Department of Teaching, Learning, and Culture – Center for Urban School Partnerships. Her research focuses on evidence-based literacy instruction for K-12 students and factors influencing writing development for bi- and multilingual

students. April currently supports grant-based research and professional development for Self-Regulated Strategy Development (SRSD) for writing as well as multiple meta-analyses of writing instruction.

**Stephen Ciullo** PhD, is an associate professor of Special Education at Texas State University. His research includes studies that explore the extent to which classroom writing instruction for students with learning disabilities (LD) aligns with research evidence. Dr. Ciullo serves as PI or Co-PI on four grants funded by the Institute of Education Sciences (IES). These grants include exploration studies, a meta-analysis, and an intervention study of text-based writing for students with LD. Dr. Ciullo teaches courses in the area of learning disabilities, inclusion, and instructional strategies. He earned his doctorate from the University of Texas at Austin and taught Special Education for six years.

**Alyson A. Collins** PhD, is an associate professor of Special Education at Texas State University. Prior to earning her doctoral degree from Vanderbilt University, Dr. Collins taught for nine years in Texas public schools. Her research focuses on exploring writing instruction for students with disabilities, effective interventions for these students, and professional development models that support teacher implementation of evidence-based practices. Dr. Collins serves as PI/Co-PI on four Institute of Education Sciences (IES) grants examining effective writing instruction. She is also Co-Director of an Office of Special Education Programs (OSEP) preparation grant promoting interdisciplinary collaboration between special educators and school psychologists.

**Francisca Costa** holds a master's degree in Clinical and Health Psychology, a bachelor's degree in Psychological Sciences and in Education Sciences. She has experience in research in educational projects and has been collaborating with other research projects and scientific initiatives in the field of psychology, including in the organization of international conferences. Her research interests are related to clinical psychology, psychopathology, emotion regulation, expressive writing, and education.

**Jeff Elmore** is a machine-learning and computational-linguistics engineer and data scientist. He worked for 20 years at MetaMetrics, Inc. developing novel natural language processing and semantic networking methods and tools to support research in literacy, vocabulary development, and language assessment. He is now Chief Technology Officer for Poetics AI where he is focused on conversational computation and story understanding using large language models to support authors of fiction and creative nonfiction texts.

**Megan Erickson** has earned her MEd in Educational Psychology from the University of Utah and an MS in Curriculum and Instruction from Western Governors University. She is the coauthor of Advantage Math, a mathematics curriculum for Grades 2–6, and worked as a professional development facilitator for six years. She is currently a school counselor in Davis School District, Utah.

**Michel Fayol** (PhD Psychology and Educational Sciences) is an emeritus Professor at Clermont-Auvergne University. He was previously in the Scientific Montpellier University (USTL (4 years), then in Dijon Bourgogne University (18 years) where he founded and headed a CNRS laboratory (LEAD), and at the same time partially in Geneva University (5 years). Lastly he moved to Clermont-Auvergne University (head of the LAPSCO CNRS laboratory) (8 years). He was then in charge of the Cognitive science in the Human department sciences of the National Research Agency (ANR) and participated in the European Research Council (as a member of the ERC Starting Grants). He has been studying development and learning in children (3–12) and adults.

**Jill Fitzgerald** is research professor at The University of North Carolina at Chapel Hill. She is a former primary grades teacher and reading specialist. Her most recent research involves text characteristics and vocabulary. She has published over 100 works and has authored, edited, and coedited several books, including the *Handbook of Writing Research*. She is a *Journal of Educational Psychology* associate editor and has served on numerous national and international editorial boards and panels. Jill has received national/international honors including the International Reading Association's Dina Feitelson Award for Research and the Reading Hall of Fame membership.

**Erin FitzPatrick** PhD, is an assistant professor in the Department of Special Education at the University of North Carolina at Charlotte. She serves on the editorial boards of Learning Disabilities Research and Practice, Teacher Education and Special Education, and Remedial and Special Education. She also served on the International Literacy Association's Writing Task Force. Her research interests include writing, self-regulated strategy development, and practice-based professional development.

**Karen R. Harris** is Regents Professor Emeritus and Research Professor at Arizona State University. She is currently co-PI of three national instructional grants. Her research focuses on validating instructional approaches derived from integrating multiple theoretical views and evidence-based practices, as well as effective practice-based PD for these approaches. Dr. Harris developed the Self-Regulated Strategy Development (SRSD) model of strategies instruction, deemed an evidence-based practice. She is a former editor of the *Journal of Educational Psychology* and was lead editor of the *APA Educational Psychology Handbook*. She has authored 14 books and more than 200 peer-reviewed articles and chapters. Dr. Harris' awards include the APA Division 15 (Educational Psychology), Career Award, the AERA Division C Scribner Award for a program of work that has significantly influenced learning and instruction, the AERA Division K Exemplary Publication Award for research in professional development, and the Council for Exceptional Children Career Award for Research.

**John R. Hayes** is emeritus professor of psychology at Carnegie Mellon in Pittsburgh, Pennsylvania. Engaged in writing research since 1979, he has created models to describe the cognitive and affective aspects of writing and explored specific writing issues. He and his colleagues have designed strategies for teaching revision to college freshmen, described how texts convey an impression of the writer's personality to readers, evaluated the impact of linguistic experience in writing, and explored the role of working memory on the writing process. Recently, he has created model of the writing processes of primary school writers.

**Michael Hebert** PhD, is a new associate professor in the School of Education and the Director of the UCI Writing Project. His primary research interests include the impacts of writing on reading, writing assessment, and the development of writing interventions. Dr. Hebert has more than 55 publications, including two influential Carnegie Corporation reports: *Writing to Read* and *Informing Writing*. He is also the co-editor of the third edition of *Best Practices in Writing Instruction*. He has received more than $9 million in grant funding, including a new 2022 Pandemic Recovery grant from the Institute of Education Sciences to test a professional development intervention to enhance reading achievement for rural students in Nebraska. He also received one of the first Early Career and Mentoring Program awards from the National Center for Special Education Research in 2013. He currently serves on the editorial board for the *Journal of Educational Psychology*. Dr. Hebert was previously a classroom teacher and reading specialist in schools in New Hampshire, Massachusetts, California, and the Navajo Nation.

**Tien Ping Hsiang** is an assistant professor and the program coordinator for Primary Education in the Faculty of Education at University of Macau. She was a visiting scholar in the Mary Lou Fulton Teachers College at Arizona State University in the fall of 2016. Dr. Hsiang enjoys writing, and she has won many writing competition awards since she was a little girl. She loves to teach children and has won several awards for teaching and learning assessment since she was a primary school teacher. Her research interests include writing instruction, teaching Chinese as a first/second language, and textbook analysis.

**Teresa Jacques** graduated in 2019 with a master's degree in Clinical and Health Psychology from the University of Porto, Faculty of Psychology and Educational Sciences. She holds a specialization course in Learning Difficulties in Reading and Writing from the same university. She is currently a PhD candidate at the University of Porto, studying the emotion regulation outcomes of expressive writing. Her interests include writing, experimental psychology, cognition, and emotion.

**Xuan Jiang** PhD, is the assistant director of the Center for Excellence in Writing at Florida International University (Miami, Florida, US). Dr. Jiang has led writing tutors in conference presentations as well as academic and creative publications. Her research interests include writing tutors' professional development, well-being and morale at workplace, issues and strategies in academic writing, transnationalism and translingualism in TESOL, writing, writing pedagogy, and writing tutoring.

She has published her studies as peer-reviewed journal articles and book chapters in these fields. Dr. Jiang has collaborated cross disciplines for conference grants, research grants, and transdisciplinary publications.

**James S. Kim** is an expert on literacy intervention and experimental design. He is professor of education at Harvard University where he leads the READS Lab (Research Enhances Adaptations Designed for Scale in Literacy), a research-based collaborative initiative to identify and scale adaptive solutions for improving children's literacy learning and outcomes. His current research priority is to understand how building children's domain knowledge and reading engagement can foster long-term improvements in reading comprehension. The Model of Reading Engagement (MORE), a spiraled and sustained content literacy intervention developed by READS Lab, has been shown to improve first to third-grade students' literacy outcomes.

**Sharlene A. Kiuhara** is an associate professor in the Department of Special Education, College of Education, and served as an Institute of Education Sciences (IES) Post-Doctoral Fellow at Vanderbilt University in Special Education. She was nationally recognized as an Early Career Publication Award recipient by the Division for Research of the Council for Exceptional Children. Her current research focuses on intervention development for students with or at risk for learning disabilities in self-regulation and constructing written arguments to develop mathematical reasoning.

**Kenny Kruse** lives in San Francisco and is the Dean of Instruction at Richmond Charter Academy, a Title I middle school. He holds an MEd in Education, Culture, and Society from the University of Utah and an MFA in Creative Writing from the University of Alabama. He has taught literature and writing with college-in-prison programs in Utah and Alabama and education programs in Myanmar, Thailand, and Cambodia.

**Joel R. Levin** is a distinguished professor emeritus of Educational Psychology at the University of Wisconsin–Madison and the University of Arizona. An award-winning researcher with approximately 425 professional publications, one strand of his research has focused on instructional strategies that are presumed to improve students' information processing and retention. Another research strand is the development of a variety of statistical procedures that are applicable in large-sample, small-sample, and single-case time-series research contexts. Dr. Levin has been a vocal advocate for enhancing the scientific credibility of educational intervention research through the adoption of rigorous methodological standards and statistical analyses.

**Amanda L. Lindner** Ph.D., is an assistant professor of Literacy in the Department of Curriculum and Instruction at Texas A&M University–San Antonio. She is a quantitative reasoning fellow and teaches undergraduate preservice teachers as well as graduate students seeking their reading specialist certification. Her research

focuses on the literacy development of diverse learners and the literacy knowledge and practices of teachers.

**Xinghua Liu** PhD, is an associate professor at School of Foreign Languages, Shanghai Jiao Tong University, China. He serves on the editorial boards of *International Journal of TESOL Studies* and *International Journal of Chinese Language Teaching*. He is currently a member of the steering committee for International Society for the Advancement of Writing Research. His research interests include intercultural rhetoric, corpus and SFL-based discourse analysis, cognitive processes of written production, individual differences in writing classrooms, and writing-based diagnosis and treatment for patients with mental disorders.

**Charles A. MacArthur** PhD, is Professor Emeritus at the University of Delaware. Major research interests include writing development and instruction, self-regulated strategies, and adult literacy. He has directed ten federally funded research projects on writing instruction, most recently a project evaluating a curriculum for college developmental writing – Supporting Strategic Writers. He has published over 125 articles and book chapters and coedited or written several books, including the *Handbook of Writing Research*, *Best Practices in Writing Instruction*, *Adult Education Literacy Instruction: A Review of the Research*, and *Developing Strategic Writers Through Genre Instruction*. In 2020, he was named an AERA Fellow.

**Linda H. Mason** is professor and the endowed director of the Helen A. Kellar Institute for Human disAbility (KIHd) at George Mason University. She earned her PhD at the University of Maryland. Dr. Mason conducts research in writing instruction including combined instruction for reading and writing from text. She is a former Editor-in-Chief of *Learning Disabilities Research & Practice* and a founding Editor of the *Journal for Inclusive Post-secondary Education.* Dr. Mason was awarded a Fulbright Scholarship and the Distinguished Early Career Research Award (DR-CEC) in 2011; in 2021, she received CEC's Research Award and the Division for LD Presidential Service Award.

**Debra McKeown** PhD, is an associate professor and Reta Haynes Fellow in the Department of Teaching, Learning, and Culture at Texas A&M University. Her research interests are focused on effective and efficient professional development models, evidence-based writing instruction, self-regulated strategy development, and use of technology for literacy instruction.

**Clarence Ng** is Professor of Education and Associate Dean for Research in Education and Arts at the Australian Catholic University. His research on reading and writing is informed by both sociocognitive and sociocultural theories of learning and engagement. He has applied the Vygotskian concept of perezhivanie to researching learner identities and language learning. He currently leads an Australian Research Council funded project promoting writing engagement and achievement for students coming from low-income families using the concepts of perezhivanie and funds of knowledge and identity.

**Amber B. Ray** PhD, is an assistant professor in the Department of Special Education at the University of Illinois at Urbana-Champaign. Her research interests include writing and reading interventions and instruction to help students with disabilities and diverse learning needs succeed. Her research focuses on strategy and self-regulation approaches to instruction and methods of professional development for teachers and school leaders on effective writing and reading instruction. Dr. Ray works on developing evidence-based instruction to cultivate the learning of all students, including those with disabilities and has conducted research with elementary through high school students and adults.

**Jackie Eunjung Relyea** is an assistant professor of literacy education at North Carolina State University and faculty associate at READS lab at Harvard Graduate School of Education. Her research interests center on literacy development and instruction of K-6 students from linguistically diverse backgrounds. Her work focuses on evidence-based literacy interventions that emphasize improving multilingual students' academic vocabulary knowledge, informational text comprehension, and content knowledge. She has won several awards, including the UNC-Chapel Hill Outstanding Young Alumnus Award and the Harvard Postdoctoral Award for Professional Development. She has published in several top-tier journals including *Reading Research Quarterly* and *Journal of Educational Psychology.*

**Peter Renshaw** is Emeritus Professor of Education at the University of Queensland and Honorary Professor at the Australian Catholic University. He has drawn upon Vygotskian theories for over three decades to investigate children's learning and recently began deploying perezhivanie to theorize learning as an emotionally engaging and holistic process. He has explored perezhivanie in relation to reading and writing and environmental experiences. He collaborates with scholars in Australia, Finland, and Hong Kong to consider how educators might address the environmental crises of the Anthropocene.

**Amy Gillespie Rouse** PhD, is an associate professor in the Department of Teaching and Learning at Southern Methodist University in Dallas, Texas. Her research interests include developing and measuring the impacts of writing interventions for struggling writers and students with disabilities. She also examines teachers' writing instruction and their use of writing-to-learn practices. Recent research includes developing simulated literacy teaching experiences for future elementary educators using Mursion software and student avatars. Dr. Gillespie Rouse serves as an associate editor for the *Journal of Learning Disabilities* and on the editorial boards for *Exceptionality*, *Journal of Educational Psychology*, and *Teaching Exceptional Children*.

**Bruce Saddler** is an associate professor at the University at Albany, Department of Educational and Counseling Psychology. His current primary research interests center on writing disabilities. He has presented at over 200 local, regional, national, and international research and professional conferences and has published over 40 scholarly articles in national and international journals including *Journal of*

*Educational Psychology*, *Journal of Special Education*, *Learning Disabilities Quarterly*, and *Reading and Writing Quarterly*. He has also authored two books on writing.

**Micheal P. Sandbank** PhD, is an assistant professor in the Division of Occupational Therapy and Occupational Science at the University of North Carolina at Chapel Hill, a former special educator, and an internationally renowned expert on early interventions for children on the autism spectrum. She is the lead researcher on Project AIM, a comprehensive and ongoing meta-analysis of all nonpharmacological interventions designed for young autistic children, which was selected as one of the Top 20 Advances in Autism Research of 2020 by the Interagency Autism Coordinating Committee of the US Department of Health and Human Services. Dr. Sandbank earned her doctoral degree in Special Education at Vanderbilt University.

**Ashley Sandoval** PhD, is a research associate at Southern Methodist University. Her primary educational interests include reading and writing interventions for struggling students and adult literacy. Dr. Sandoval's most recent project involves development of an Android app to support learning for adults with low literacy skills.

**Tanya Santangelo** is an associate professor in the School of Education at Arcadia University. She has also been an associate editor of *Journal of Educational Psychology,* board member for the Council for Exceptional Children's Division for Research, and research fellow with the George W. Bush Institute's Middle School Matters initiative. Prior to earning her doctorate in special education from the University of Maryland, Dr. Santangelo spent nearly 10 years as an educator—working with amazing kiddos and their families. She is the (co)author of numerous publications, targeting both scholarly and practitioner-oriented audiences. Currently, her main areas of interest and scholarship include: developing, validating, and disseminating effective practices for teaching and assessing writing; increasing educators' knowledge and successful implementation of evidence-based practices; and using meta-analytic methods to synthesize and advance educational research.

**Bernard Slusarczyk** Doctor in Educational Sciences, is a college principal and researcher attached to the Acté laboratory (Knowledge Transmission Education Activity) at Clermont-Auvergne University. His work on the study of written verbal production in the school environment (primary and secondary school) is based on longitudinal and cross-sectional studies carried out in an ecological environment. He is studying questions related to the improvement of students' writing performance and the study of educational devices favoring this improvement such as philosophical writings in secondary school.

**Nicole Patton-Terry** PhD, is the Olive and Manuel Bordas Professor of Education in the School of Teacher Education, Director of the Florida Center for Reading Research (FCRR), and Director of the Regional Education Lab—Southeast at Florida State University (FSU). Dr. Terry's research, innovation, and engagement

activities concern young learners who are vulnerable to experiencing poor language and literacy achievement in school, in particular, African American children, children growing up in poverty, and children with disabilities.

**Malynda Tolbert** is an assistant professor in the School of Education at Utah Valley University. She received her PhD in Special Education from the University of Utah. She has experience as a curriculum and instruction specialist, clinical professor, professional development instructor, and K-12 classroom teacher. Her research interests include investigating pre- and in-service development programs for teachers to implement evidence-based classroom practices and research intervention strategies for students with and at risk for disabilities.

**Gary A. Troia** PhD, CCC-SLP is associate professor of special education in the Department of Counseling, Educational Psychology and Special Education at Michigan State University. Dr. Troia is co-editor of the journal *Topics in Language Disorders* and an author of the preschool curriculum *Promoting Awareness of Speech Sounds* (*PASS*), published by Attainment Company. He has authored over 70 research papers, book chapters, and white papers and has given numerous presentations about his work in the areas of phonological processing and awareness, writing assessment and instruction, and teacher professional development in literacy.

**Naomi Weintraub** is a professor and head of the School of Occupational Therapy, the Faculty of Medicine, at the Hebrew University of Jerusalem. She is the founder and head of the Neuro-Developmental Disabilities and Writing Research Laboratory. Her research focuses on understanding the underlying functions of handwriting and keyboarding. She has developed several handwriting and keyboarding standardized tests in Hebrew and Arabic as well as keylogging software. She also studies the functional implications of neurodevelopmental disorders (mainly dysgraphia, dystypia, and Developmental Coordination Disorder) on the participation in daily life activities and quality of life of children and young adults.

**Kausalai (Kay) Wijekumar** PhD, is Houston Endowment Chair, Chancellor's EDGES Fellow, and Director for the Center for Urban School Partnerships at Texas A&M University. She is a passionate advocate for all children and dedicates her research, development, and dissemination efforts to eradicating illiteracy. She has developed web-based intelligent tutoring systems for reading and writing in English and Spanish. She has further developed and integrated teacher professional development systems using the web-based tools. Her recent work integrates podcasts for families to encourage reading and writing at home. She is a musician, avid gardener, and world traveler.

**Murphy K. Young** is a doctoral student at Southern Methodist University and a University PhD fellow. She holds a BS in Interdisciplinary Studies and an MEd. in Curriculum and Instruction, both from Texas A&M University. Prior to joining the PhD program at SMU, she taught middle grades Language Arts and Social Studies for four years.

# Part I
# Prologue

# Honoring Steve Graham: Personal Reflections on an Outstanding Scholar, Colleague, and Partner

**Karen R. Harris and April Camping**

**Abstract** In this chapter, we present our personal reflections regarding the work of, and our time spent collaborating with, Dr. Steve Graham in the area of writing. First, Dr. Karen Harris, Steve's spouse and colleague for over 40 years, share memories and milestones in their work together, as well as other areas. Then, Dr. April-Camping, Steve and Karen's final doctoral student, shares her reflections based on having worked with Steve as a doctoral student and for 3 years afterwards. We are honored to be a part of this festschrift for Steve. The chapters in this book honor the influential body of work he has contributed to the field while also contributing influential new work and insights.

**Keywords** Steve Graham · Meta-analyses · SRSD · Festschrift · Writing

K. R. Harris (✉)
Arizona State University, Tempe, AZ, USA
e-mail: Karen.R.Harris@asu.edu

A. Camping (✉)
SRSD Online - Writing to Learn, East Greenwich, RI, USA
e-mail: april.camping@gmail.com



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_1

As we write our reflections, Steve does not know that this book is in the works, as it is meant to be a surprise. Xinghua (Kevin) Liu, Michael Hebert, and Rui Alves are the editors of *The Hitchhiker's Guide to Writing Research: A Festschrift for Steve Graham*. They asked me to write a personal reflection on Steve's contributions to the field and our work together for over four decades. This book honors the influential body of work Steve has led as well as contributed to over these years, and contributes influential new chapters to the field. This book also honors who he is – an outstanding researcher with the ability to be a leader, a team member, and as they noted, a "teacher, tutor, friend, and colleague." Steve has been many of these things to me over these years, as well as my partner, my spouse, a wonderful father to our daughter, my best friend, and the love of my life. He has also made me laugh every day. He loves to tell stories and to tease. I well remember our daughter as a young girl, coming to me and saying, "Mommy, make him stop!"

I am delighted to have this chance to share my reflections. Writing this piece has created a wonderful trip though many years of memories. I decided to start at the beginning – how I first met Steve, how we came to work together, and where we are now. Several former students who worked with Steve (and typically with me as well) are involved in this book. We have been so fortunate to work with so many outstanding doctoral students who have gone on to do important work in schools, research institutions, and universities. We will not be advising any more doctoral students. Our final doctoral student recently graduated, Dr. April Camping (whom we co-advised). Her reflections conclude this chapter.

## 1 How It All Began

In August of 1978, I arrived at Auburn University eager to start the doctoral program in Emotional and Behavioral Disorders. I was the only one to show up for this brand-new program; thus, the program was closed. What to do? I decided I would need to look for a new doctoral program the next year in this area because in addition to the program being closed, I now had no assistantship.

Serendipity, a theme that has occurred frequently across my life, intervened (Bembenutty, 2022b). A little background will help to set the stage. Before arriving at Auburn University, I taught in an alternative special education program for students ages 13 to 21 in Lincoln, Nebraska from 1975 to 1978. During that time, I also completed my master's degree at the University of Nebraska - Lincoln, in the Department of Educational Psychology. My concentration areas were statistics, learning, and child development. I also completed certification coursework in the area then called mental retardation, currently referred to as intellectual disability or cognitive disability.

I received my undergraduate degrees from the University of Northern Colorado (UNC) in 1974 after completing high school in Illinois in 1971. I had started with a major in Education of the Blind at Illinois State University in 1971. I then made the 'absolutely brilliant decision' to switch majors to Education of the Acoustically

Handicapped, and transferred to UNC. I was worried about the availability of jobs in education of the blind because most were residential and there were not a lot of jobs available. There were, however, barely more jobs available in education of the deaf. This led me to complete two undergraduate degrees at UNC (dual degrees was not an option), one in education of the acoustically handicapped and one in elementary education. I was right to be concerned. The law that guaranteed all children, including those with disabilities, a free and appropriate education did not pass until 1975 (PL 94–142). There were few jobs available in special education in 1974. I completed my student teaching in one of the first Total Communication Programs for children who are deaf in Denver, Colorado. I received permission to do my elementary education student teaching in a kindergarten classroom in Morgantown, West Virginia. I asked to student teach there because during high school I was part of a tutoring group and tutored a young girl in an area of Chicago where many families formerly from Appalachian coal mining areas lived. From 1974 to 1975, I taught fourth grade at the Cassville Grade School in Cassville, West Virginia, a "coal town."

All of this came into play when I arrived at Auburn University in 1978. I had learned so much about education, learning, teaching, disability, inequity, and poverty. I was offered an assistantship in the Foundations of Learning department by Drs. Glenelle and Gerald Halpin because of my background. I typically taught four sections a quarter of the Introduction to Educational Psychology course and worked with both of them until my graduation in 1981. I remained a doctoral student in special education, however. I thought I would leave, but wanted to take courses that first year. Dr. Lemoine Miller, my primary advisor throughout my doctoral program, suggested I take a course on learning disabilities. This was a new area in special education, as was emotional and behavioral disorders. During this course, I realized many of the adolescent students in special education I had taught had learning disabilities as well as emotional and behavioral challenges. This course helped shape the rest of my life. I decided to stay and complete the doctoral program in learning disabilities.

I loved my time at Auburn University. To some degree I was a fish out of water, however, because so many of the students there came from well-off families. I was a midwestern woman living on a doctoral student assistantship. The undergraduate women in the College of Education dressed for class as if for a fancy evening out. Many in my classes were quite straightforward about getting the teaching degree so they could get a job if they weren't married at the end of college, or if they needed to make some money in the early years of marriage. We called it the "Mrs. Degree" (there were also many young people dedicated to becoming teachers). Dr. Miller, our good friend for decades, recommended that I audit and serve as a teaching assistant in Introduction to Special Education, a course I had never had. Serendipity arrived again.

On the first day of the Introduction to Special Education class, I sat in the back of the room. The professor, whom I had not yet met, was late and there was a steady buzz of talk across the room. We were all about to leave when a young man walked in. He had hair down to his shoulders and was wearing beat up blue jeans, a t-shirt, and flip flops. No one stopped talking. He cleared his throat a couple of times at the

front of the room and then said, “I am Dr. Steve Graham and this is Introduction to Special Education.” Suddenly, you could hear a pin drop in that room. While most of the students looked a bit surprised (or shocked), I was happy to see someone whose background seemed to be more like mine.

Steve was a visiting Assistant Professor in the Department of Special Education. In addition to assisting in his class, I joined him in working on a paper about research in spelling. He misspelled the word misspell throughout the entire first draft, something we laughed about then and still do today. Over the first 2 years, I worked with him occasionally on other projects in spelling and handwriting, and attended conferences on learning disabilities where he introduced me to others in the field. Long story short, somewhere in the second year of this doctoral program, we realized we were becoming more than friends and began dating. Dr. Miller had been Steve’s advisor during his master’s degree in learning disabilities at Valdosta State College in Georgia. It was a bit awkward that we were dating and we tried to keep it a secret. When Steve decided to tell Lemoine, his former advisor’s response was, “Tell me something I don’t already know.” I don’t think you can keep a secret in towns that small!

Steve and I had a great deal in common. We both came from working class families (his father had been enlisted in the army, mine was a shoe clerk at a department store when I was born), had moved about every 3 years, were the first in our families to graduate from college, were ‘hippies’ of the baby boomer generation, were strongly committed to social justice and civil rights, and had previously taught children or adolescents living in low income or poverty areas. We both were passionate about improving instruction, especially for students from marginalized, underserved groups. Steve had done his dissertation in reading, but then developed a strong interest in writing, in part because so little research had been done there. He started with work in spelling and handwriting, and to this day is recognized as one of the world’s leading researchers in handwriting. He then moved into studying composing, writing genres, and the writing process.

I had become a “theory aficionado” during my master’s program; I studied affective, behavioral, cognitive, developmental, and social theories (Albert Bandura’s social learning theory and Bronfenbrenner’s ecological systems theory, for example). Studying theories continued throughout my doctoral work, and continues today. My teaching experiences had convinced me that no single theory alone was sufficient for understanding how K-12 students learned and that our students, especially those denied equity, needed and deserved far more effective instruction (for greater detail, see Bembenutty, 2022a, b; Harris, 2018). By the end of my doctoral program, I had begun working to integrate theories with powerful research behind them into an instructional method.

In 1982, Steve and I were married and the Self-regulated Strategy Development (SRSD) instructional model was born (albeit a first iteration). We joined our passions and initially focused on improving students’ composing abilities. Steve waited patiently (for the most part) while I put the pieces of findings from several bodies of research together to form an instructional process. Our first study ensued (Harris, 1982; Harris & Graham, 1985). We selected our first focus in writing,

Steve selected or developed our outcome measures, and I designed and wrote the lesson guides, created the teacher and student materials, and led SRSD instruction with students.

Several well-meaning colleagues told us that working from multiple theories would never work. We were stubborn and energized, however, and continued this work together. In the early years, we published several SRSD studies together focused on instruction in writing (e.g., Graham & Harris, 1993; Harris & Graham, 2009) and one on word problem-solving (Case et al., 1992). We have also collaborated on several books (e.g., Graham & Harris, 2005; Harris & Graham, 1992, 1996; Harris et al., 2008, 2012). Several of my favorite memories involve long walks on the beach during "workations," where we thought out and designed upcoming studies, papers, or books. People sometimes comment that it must be difficult to live and work together, but for us it has been a joy.

## 2 Where We Are Now

We have been faculty members at Purdue University (2 years), University of Maryland (23 years, our daughter was born and raised here), Vanderbilt (8 years), and now Arizona State University (11 years). We have continued to work together, although for over a decade we have each led work in differing areas. Steve has been committed to meta-analyses of the growing bodies of research in writing, the policy implications of the findings, developing the Writer(s)-Within- Community model, writing assessment, and evidence-based practices in writing. He has published an impressive legacy across these areas. I used to tease him about "giving a little boy a hammer" because of the number of meta-analyses he had going. This work, however, has provided the field with a wealth of knowledge about where we are in writing research and where we need to go (e.g., Graham, 2018; Graham et al., 2013, 2015a, b, 2018a, b, 2020, 2021; Graham & Harris, 2018a, b; Graham & Hebert, 2011; Sandmel & Graham, 2011). Many more meta-analyses, and other research studies, are currently in the works.

Over this time period, I have focused on developing web-based support for SRSD instruction, integrating reading and writing instruction, and investigating and validating effective practice-based professional development for SRSD instruction in writing in special education and inclusive classrooms (e.g., Harris et al., 2012a, b, 2021, in press; Festas et al., 2015; McKeown et al., 2019; Wijekumar et al., 2017). Steve has joined in this work, often leading or contributing to selection and scoring of the writing assessment, and more.

We have both reached the point in our careers where discussions of becoming emeritus professors have occurred, and I made the change in 2022, becoming Regents Professor Emeritus. Steve will make decision in a couple more years. We are focusing our time on conducting the work that matters so much to us, completing grants, and writing up the data we have. I don't expect either of us to completely retire for some years to come; I look forward to following the work Steve continues

to contribute to the field, our continued work together, our travels, family time, and adventures to come. We also look forward to continuing our work with Dr. April Camping. She shares her thoughts next.

## 3 Reflections from Our Final Doctoral Student: April Camping

It was fall of 2016 when I first met Dr. Steve Graham and Dr. Karen Harris. I had reached out to them while applying to a doctoral program at ASU, and although I had come across their research during my time as a K-12 teacher, I was naïve to the actual scale/scope of their work. I became Karen's doctoral student; Steve became my co-adviser when we began work on one of the many meta-analyses we would undertake over the next 4 years. They explained to me that I would be their last doctoral student. Karen invited me to join in this chapter for this reason. Throughout my 4 years as their final doctoral student, I learned about their collective impact on the field of literacy education research. I was also witness to and recipient of their unrivaled generosity.

As the field is well aware, Steve has contributed seminal insights about writing instruction through his work on over 20 meta-analyses. This is one strand of his generosity: generosity of knowledge. Steve's willingness to share a career's worth of knowledge with a neophyte scholar in terms of the literature and the meta-analysis endeavor itself is something for which I will always be in debt. His patience in teaching was coupled with an emphasis on my autonomy, and his passion for the work was contagious.

Relatedly, working with Steve and Karen lead to a wealth of opportunities. This second strand, generosity of opportunity, was particularly impactful for me as I worked to establish an academic vita. They offered me experiences at first authorship, seats on important committees, lead presenter ventures at conferences, chances to review manuscripts, and too many opportunities to count for collaborative research.

I finish this brief reflection by noting a final strand: Steve and Karen's generosity in terms of validation. Life happens throughout the course of a doctoral degree. But amidst the many personal and professional challenges I faced, they validated both the legitimacy of these challenges as well as my ability to keep going in spite of them. I successfully defended my dissertation thanks to Karen and Steve's unrelenting support, despite a global pandemic. I have a number of projects still in process with them that I look forward to finishing; some were completed during my time as a Postdoctoral Research Associate at Texas A & M. Now, as Research Director at SRSD Online – Writing to Learn, I look forward to applying Steve's Writer(s)-Within-Community model in research with culturally and linguistically diverse students. Steve wrote (Graham, 2021) "I am at the end of this story but hopefully not at the end of the story." As we all shift to new chapters, I hope to continue the story of impactful research and unbounded generosity.

## References


Bembenutty, H., (Ed.) with Graham, S. (2022a). An interview With Steve Graham: A pioneer of the understanding of how writing develops, is taught effectively, and can be used to support reading and learning. *Contemporary pioneers in human learning and development: Teaching, research, and applications* (pp. 51–70). Information Age Publishing.

Bembenutty, H., (Ed.) with Harris, K. R. (2022b). An interview with Karen R. Harris: The pioneer and developer of the Self-Regulated Strategy Development (SRSD) model of writing instruction. *Contemporary pioneers in teaching and learning* (Vol. 2). Information Age Publishing.

Case, L. P., Harris, K. R., & Graham, S. (1992). Improving the mathematical problem-solving skills of students with learning disabilities: Self-regulated strategy development. *Journal of Special Education, 26*, 1–19.

Festas, I., Oliveira, A., Rebelo, J., Damião, M., Harris, K. R., & Graham, S. (2015). Professional development in self-regulated strategy development: Effects on the writing performance of eight grade Portuguese students. *Contemporary Educational Psychology, 40*, 17–27. https://doi.org/10.1016/j.cedpsych.2014.05.004

Graham, S. (2018). The writer(s)-within-community model of writing. *Educational Psychologist, 53*, 258–279.

Graham, S. (2021). Through the looking glass: Reflections of a writing scholar. Acquired wisdom series, S. Nieto, F. Erickson & P. Winne (Eds.). *Education Review, 28*, 1–21. https://doi.org/10.14507/er.v28.xxx

Graham, S., & Harris, K. R. (1993). Self-regulated strategy development: Helping students with learning problems develop as writers. *Elementary School Journal, 94*, 159–182.

Graham, S., & Harris, K. R. (2005). *Writing better: Effective strategies for teaching students with learning difficulties*. Brookes.

Graham, S., & Harris, K. R. (2018a). Evidence-based writing practices: A meta-analysis of existing meta-analyses. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Design principles for teaching effective writing: Theoretical and empirical grounded principles* (pp. 13–37). Brill Editions.

Graham, S., & Harris, K. R. (2018b). An examination of the design principles underlying a self-regulated strategy development study based on the writer(s)-within-community model. *Journal of Writing Research, 10*, 139–187.

Graham, S., & Hebert, M. (2011). Writing-to-read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*, 710–744.

Graham, S., Harris, K. R., & McKeown, D. (2013). The writing of students with LD and a meta-analysis of SRSD writing intervention studies: Redux. In L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 405–438). Guilford Press.

Graham, S., Harris, K. R., & Santangelo, T. (2015a). Research-based writing practices and the common core: Meta-analysis and meta-synthesis. *Elementary School Journal, 115*, 498–522.

Graham, S., Hebert, M., & Harris, K. R. (2015b). Formative assessment and writing: A meta-analysis. *Elementary School Journal, 115*, 524–547.

Graham, S., Liu, K., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapel, J. (2018a). Balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly, 53*, 279–304.

Graham, S., Liu, K., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018b). Reading for writing: A meta-analysis of the impact of reading and reading instruction on writing. *Review of Educational Research, 88*, 243–284.

Graham, S., Kiuhara, S., & MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research, 90*, 179–226.

Graham, S., Aitken, A., Hebert, M., Santangelo, T., Camping, A., Harris, K. R., Eustice, K., & Ng, C. (2021). Do children with reading difficulties experience writing difficulties? A meta-analysis*, 113*(8), 1481–1506.

Harris, K. R. (1982). Cognitive-behavior modification: Application with exceptional students. *Focus on Exceptional Children, 15*(2), 1–16.

Harris, K. R. (2018). Educational psychology: A future retrospective. *Journal of Educational Psychology, 110*(2), 163–173. https://doi.org/10.1037/edu0000267

Harris, K. R., & Graham, S. (1985). Improving learning disabled students' composition skills: Self-control strategy training. *Learning Disability Quarterly, 8*, 27–36.

Harris, K. R., & Graham, S. (1992). *Helping young writers master the craft: Strategy instruction and self-regulation in the writing process*. Brookline Books.

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.). Brookline Books.

Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *British Journal of Educational Psychology, 6*, 113–135.

Harris, K.R., Camping, A., & McKeown, D. (in press). A review of research on professional development for multicomponent strategy-focused writing instruction: Knowledge gained and challenges remaining. In DeSmedt, F., Bouwer, R., Limpo, T., & Graham, S. (Eds). Conceptualizing, designing, implementing, and evaluating writing interventions. Brill Publishing.

Harris, K. R., Graham, S., Mason, L., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Brookes Publishing.

Harris, K. R., Graham, S., & Urdan, T. (Eds.). (2012). *APA educational psychology handbook: Theories, constructs, and critical issues* (Vol. 1–3). American Psychological Association.

Harris, K. R., Lane, K. L., Driscoll, S., Graham, S., Wilson, K., Sandmel, K., Brindle, M., & Schatschneider, C. (2012a). Tier 1, teacher-implemented self-regulated strategy development for students with and without behavioral challenges. *Elementary School Journal, 113*, 160–191.

Harris, K. R., Lane, K. L., Graham, S., Driscoll, S., Sandmel, K., Brindle, M., & Schatschneider, C. (2012b). Practice-based professional development for self-regulated strategies development in writing: A randomized controlled study. *Journal of Teacher Education, 63*(2), 103–119.

Harris, K. R., Young, S., Camping, A., & Graham, S. (2021). Yes they can: An RCT of Tier 2 SRSD plus transcription and oral language skills for writing to inform using source text at 1st and 2nd grades. Manuscript submitted for publication.

McKeown, D., Brindle, M., Harris, K. R., Sandmel, K., Steinbrecher, T. D., Graham, S., Lane, K. L., & Oakes, W. P. (2019). Teachers' voices: Perceptions of effective professional development and classwide implementation of self-regulated strategy development in writing. *American Educational Research Journal, 56*, 753–791.

Sandmel, K., & Graham, S. (2011). The process writing approach: A meta-analysis. *Journal of Educational Research, 104*, 396–407.

Wijekumar, K., Harris, K. R., Graham, S., & Meyer, B. (2017). A web-based intelligent tutor for supporting elementary classroom teachers in persuasive writing. In S. Crossley & D. McNamara (Eds.), *Handbook of educational Technologies for Literacy* (pp. 294–301). Taylor & Francis.

# Part II
# Writing Theories and Models

# Modeling the Impact of Emotions on Descriptive Writing


**John R. Hayes**

**Abstract** In this chapter, I propose a process model for how writers express their emotions through word choice in written text. I also present a computer simulation of the model to demonstrate that it can actually choose words as writers would.

The model assumes that writers recognize many families of words. Each family member is related but conveys a different shade of meaning. For example, rain might be described as a sprinkle, a drizzle, a shower, a storm, a deluge, or a frog-strangler. In the model, when the translation process selects language to express ideas from the proposer, emotions influence which word is chosen from the relevant word family.

The connection of this model to a volume dedicated to Steve Graham's body of work is this. The model proposed here is teachable. It could be taught as part of SRSD to make writers aware of their options for expressing degrees of emotion or, more generally, degrees of meaning. As the computer simulation of the model shows, difficulty in expressing degrees of meaning could result from limited vocabulary. A teacher could potentially help writers who have difficulty expressing emotions through vocabulary instruction or by instruction in the use of a thesaurus.

**Keywords** Emotion · Computer simulation · Word choice

Steve Graham does a variety of things extremely well. For me, three big areas stand out. The first is the development, with Karen Harris, of the Self-Regulated Strategy Development method (SRSD)—a revolutionary and highly effective approach to teaching writing to young or disadvantaged students (Harris & Graham, 1992; Graham et al. 1998; Graham & Harris, 2018; Harris et al., 2002, 2006; Harris &

The author thanks Karen Schriver and Michael Hebert for their insightful and very helpful feedback on drafts of this chapter.

J. R. Hayes (✉)
Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: jh50@andrew.cmu.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_2

Graham, 2017; 2018). Second is the series of meta-analyses he conducted with a variety of co-authors that show the effectiveness of SRSD, and more generally, offer insights into what works in teaching writing to young or disadvantaged writers (Graham & Harris, 2003; Graham & Perrin, 2007; Graham & Hebert, 2011; Graham & Sandmel, 2011; Graham et al., 2007, 2012; Rogers & Graham, 2008). These analyses are extremely valuable resources for researchers. The items cited here referencing SRSD and the meta-analyses represent a small sample of Graham's and of Harris's publications that I could have referenced. Finally, I admire Steve's contributions to the modeling literature. Steve's writing model (Graham, 2018) provides a detailed and insightful account of many of the factors influencing the writing process, including both the characteristics of the writer and the characteristics of the community in which the writing is done.

Among the features that are unique in Steve's model is a concern for the impact of emotion on writing. Other modelers have been concerned with related variables. For example, David Galbraith (1999) has provided a clever model to account for the impact of writers' dispositions, that is, their attitudes and biases, and Hayes (1996, 2012) has emphasized the importance of motivation in his models. However, neither author has dealt specifically with the impact of emotions on writing. In this chapter, I will try to augment Steve's model by describing specific mechanisms to account for the impact of both emotions and dispositions (or biases) on a particular kind of writing: descriptive writing in which authors inform their readers of their impressions of the agents and actions, sights and sounds involved in events.

## 1 A Framework for Descriptive Writing

To theorize about the impact of emotion on descriptive writing, I have drawn on my 2012 writing model (Hayes, 2012). This model (shown in Fig. 1) provides a framework for describing a variety of kinds of writing, such as essays, journals, reports, etc. The 2012 model postulates a set of processes and resources used by writers as well as the relations among them.

The parts of the model that are most relevant to representing descriptive writing are highlighted in black in Fig. 1. Briefly, this is what the parts of the model do:

1. The *Current Plan* in descriptive writing is a plan to write a description of an event.
2. The *Proposer*, using inputs from the writer's senses, working memory and long-term memory, formulates ideas for expression and passes them to the translator.
3. The *Translator* chooses language to express the ideas from the proposer and passes them to the Transcriber.
4. The *Transcriber* gives the language physical form as text or graphics, and becomes part of the text-so-far.
5. The *Text-So-Far* includes all the text produced so far by the writing process.

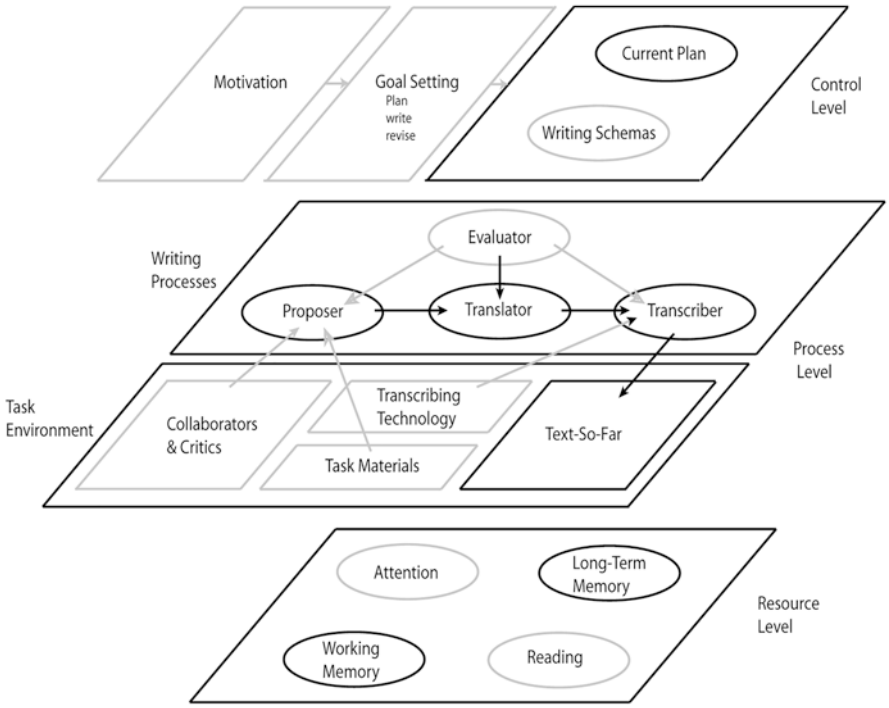



**Fig. 1** Hayes' 2012 model of writing processes

6. *Long-Term Memory* is where memories for the event and communications about the event are stored.
7. *Working memory* holds and manipulates information over short periods of time. It is central to problem solving, decision-making, and many other cognitive processes. *Working Memory*'s special role in buffering text production will be discussed later.

To describe a particular variety of writing, it is typically necessary to augment the 2012 model with additional processes and resources and to specify some processes in greater detail. For example, to describe workplace writing from multiple sources, Leijten et al. (2014) added several features to the 2012 model including inputs from collaborators and critics, inputs from multiple digital sources, graphics as part of the text-so-far, and a process for motivational management.

To model descriptive writing, I have augmented my 2012 model in two major ways. First, recognizing that a descriptive writer is first of all, an observer, I have modeled the processes by which sensory inputs from an event are represented in the writer's long-term memory (see Fig. 2). This addition is important for three reasons. First, it recognizes the role of sensory input in descriptive writing. Second, as we will see, the processes involved in storing sensory information in long-term memory may be influenced by emotion. And third, it is necessary to specify the structure of

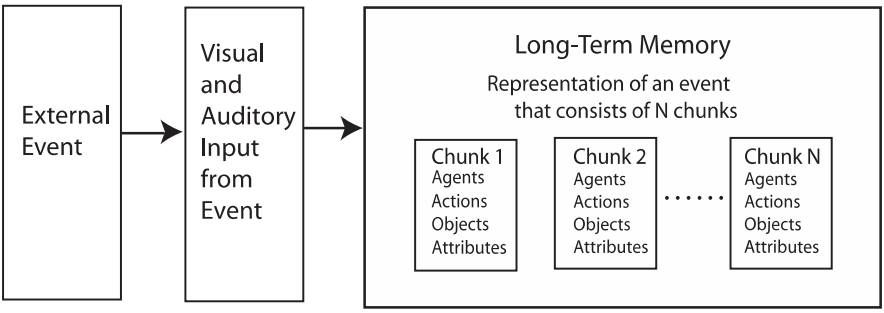



**Fig. 2** How an observed event is stored in long-term memory

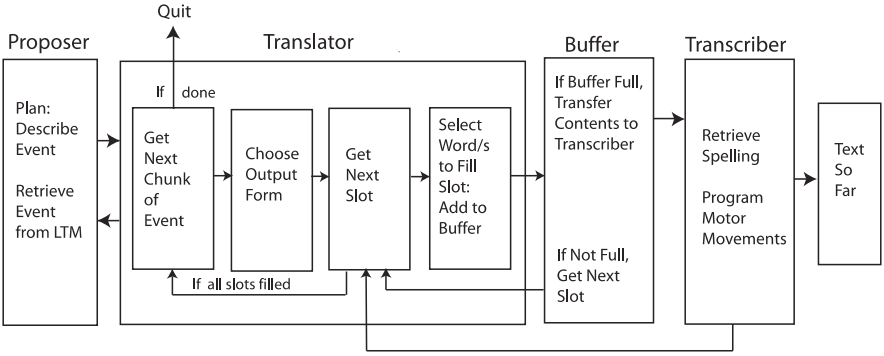



**Fig. 3** A model of descriptive writing

these memories to understand how other writing processes can use them. As Fig. 2 shows, I have assumed that memories for an event are stored in a sequence of temporally or causally ordered chunks. The chunks may be coded either verbally or non-verbally, depending on the nature of the sensory input. The chunks include information about the agents, actions, and objects, together with their attributes that the writer has observed. For example, a chunk might contain non-verbal information that a hawk (agent) had struck (action) a dove (object) with lightening speed (attribute). In addition, the chunks include information about the writer's focus of attention. For example, one writer might focus on the agent of an action, in which case, the agent would be at the top of the list in that chunk within the writer's memory. However, another writer might be focused on the object of the action, in which case, the object would be on the top of the list in that writer's memory,

Second, I have made two elaborations of the Translation process (see Fig. 3). I have added detail about the subprocesses operating within the Translator. These subprocesses allow the Translator to use the parts of the memory structures stored in long-term memory in an appropriate order. For example, they allow the Translator not only to process the chunks from Long-Term Memory in temporal order, but also to correctly interpret information in the chunks about the writer's focus.

I have also added a buffer between the Translator and the Transcriber, the Buffer represents that part of Working Memory that stores the language that we have thought about, but have not yet written down. Many of us have had the experience of thinking of an addition to our text and saying to ourselves, "I better write that down before I forget it." The buffer hold the language we produce before we commit it to paper.

Turning now to the big picture, this is how the model of descriptive writing works. The process starts when the writer observes or hears an event and stores a representation of it in Long-Term Memory. Writing starts sometime later when the writer makes a plan to write a description of the event, perhaps for a letter or a personal journal. When this occurs, the Proposer retrieves information about the event from Long-Term Memory and makes this information available to the Translator. The Translator then accesses the first chunk of the event and chooses an output form for translating the information into language. These output forms are grammatical structures with a series of open slots to be filled. Some of the possible forms and possible slots (represented as subjects, verbs, and objects) are shown in Table 1.

The Translator then chooses the next slot to be filled and selects words to fill it. The output form and the words chosen to fill the slots determine how the chunk will be translated into language. When the form is completed, it is then stored in the Buffer. Since the Buffer is part of Working Memory, its capacity depends on the current load on working memory.

If the Buffer is full, its contents will be transferred to the Transcriber, which transforms the information into text or graphics, and is then added to the Text-So-Far. When this process is complete, control is passed back to the Translator to finish the chunk or, if finished, to get the next chunk. If the Buffer is not yet full, control is also passed back to the Translator to get the next slot. When all of the slots in the first chunk have been filled, the Translator gets the next chunk. The process continues until all of the chunks have been translated. At this point, any content remaining in the Buffer is transferred to the Transcriber for outputting and the process stops. The language that has been produced by this process is recorded in the Text-So-Far and also becomes part of the representation of the event that has been stored in Long-Term Memory.

**Table 1** Some grammatical forms generated by the Translator

| Form | Slots required to complete form |
|---|---|
| S-V-O | subject (slot)-verb (slot)-object (slot) |
| O-V-"by"-S | object (slot)-verb (slot)-"by"-subject (slot) |
| O-"was"-V | object (slot)-"was"-verb (slot) |

## 2 What Descriptive Processes Are Influenced by Emotion?

There are three occasions en route to creating a description that seem most likely to be influenced by biases and emotions:

1. The occasion when an event is initially observed.
2. The occasion when the Translator translates an element into language, and in particular, when the Translator chooses the form of the output and the words to fill the slots in the output form.
3. The occasion when the writer's description of the event is stored in memory.

### *2.1 The Initial Observation*

Substantial evidence indicates that biases and emotions can influence the initial encoding of a visual experience. For example, Payne (2001) briefly primed white American undergraduates with a picture either of a black face or a white face. The participants then identified whether a second picture represented either a gun or a tool. Payne found that his participants were quicker to identify a gun and more likely to mistake a tool for a gun when primed with the black face. Similarly, Unkelback et al. (2008) showed New Zealand undergraduates pictures of people wearing either Islamic headgear (e.g., turban or scarf) or no headgear, who were, in addition, either carrying a gun or a tool. Participants were instructed to press a button marked 'shoot' if the participant was carrying a gun, but not to do so otherwise. The authors found that participants were more likely to press 'shoot' for those wearing Islamic headgear. These authors also found that when they induced anger in participants, the participants were more likely to 'shoot' and presumably to perceive guns whether the people had Islamic headgear or not. Many studies have found results similar to Payne's and Unkelback's (Correl et al., 2002; Duncan, 1976; Payne et al., 2002; Zeelenberg et al., 2006). These studies strongly suggest that biases and emotions can influence a person's initial representation of an event.

### *2.2 During the Translation Process*

Choosing the output form for describing an event (see Fig. 3, the second box in the Translator) and choosing words to fill the slots (see Fig. 3, the final box in the Translator) could both be influenced by emotions. For example, Forgas (1999a, b) studied how moods influence the way people write requests. In his studies, Forgas predicted that people in a bad mood would write more polite requests because they would have less confidence that their requests would be positively received. As predicted, participants reliably wrote more polite requests when they were induced to have a bad mood rather than a good one.

Presumably, requests that differ in politeness can differ both in output form and in word choice. We can imagine that being in a bad mood might often have the opposite effect. Suppose, for example, a manager is waiting for an overdue report from an assistant. In a good mood, he might say, "By the way, did you finish that report?" But in a bad mood, he might say, "Finish that report ASAP!" Of course, writing a request differs from writing a description. None-the-less, Forgas' results make it seem plausible that output forms and word choices could also be influenced by emotion when writing a description.

## *2.3 Storing the Description in Long-Term Memory*

Finally, a person's language used when describing an event can combine with the initial memory representation of the event to create a new representation of the event in long-term memory. If the person describes the event using extreme language, either in writing or conversation, we would expect the combined memory would represent a more extreme representation than the one originally stored. Further, we might expect that if the person recounts their description more than once, their language might become still more extreme. Therefore, if the person's first description employs extreme language, we would expect that their second description, based on the new combined-memory representation, would be more extreme than the first. I will explore this hypothesis further below.

There is also substantial evidence that language describing an event, which is experienced after the event, can influence people's memory for the event. For example, Loftus and Palmer (1974) asked participants who had watched films of a car crash either, "How fast were the cars going when they *hit* each other?" or "How fast were the cars going when they *crashed* each other?" When the question included the word *crashed*, participants reported significantly higher speeds than when it included the word *hit*.

The participants in the Loftus and Palmer study were in a different situation than a person writing a description. In the Loftus and Palmer study, they presented critical language in the form of a question. In writing a description, the critical language is created by writers themselves. Whether this makes an important difference isn't clear. Never the less, the Loftus and Palmer study and other related studies (Loftus, 1975; Loftus et al., 1978) do indicate that language experienced after an event can, in some circumstances, influence memory for the event.

My objective here is not to prove that emotions can influence writing. The evidence reviewed above suggests that they do. Rather, my task is to propose plausible mechanisms by which that interaction can occur. Next, I will propose such mechanisms for descriptive writing.

## 3 How Emotions Can Influence Translation During Description

People have a wealth of linguistic knowledge that allows them to express a wide variety of shades of meaning when describing observed events. For example, they know that "downpour" describes a more intense rain than does "drizzle." Generally, people have a range of lexical options that vary in intensity for describing events. One can describe a laugh as a "titter," a "snicker," a "chortle," a "giggle," a "guffaw," a "belly-laugh," a "side-splitter," or a variety of other alternatives. I will call such collections of words that describe the same action or thing with different intensities a "word family." Part of language learning involves acquiring the ability to match the intensity of a descriptor to the intensity of the thing described; that is, learning to describe things as one's language community prefers to characterize them. For example, it is easy to agree which of Figs. 4 and 5 most people would choose to represent a titter and which a side-splitter. If writers use language in a way that does not match community norms, their language could be described as biased, imprecise, or unidiomatic.

Central to my account of how emotions can influence description is the idea that writers can use their knowledge to select descriptors that are neutral (or more precise) by community standards. They can also choose words that are more or less intense than neutral to express their biases and emotions. In particular, I propose that under the influence of emotion, a writer may prefer descriptors that indicate that an action is more or less intense or that an object is more or less valuable than a neutral description would indicate. For example, in a car accident, the angry victim might say that his "very expensive car was *destroyed*"; the guilty driver might say the "old wreck was *dented*," and a neutral observer might say, "the car was *damaged*."

To illustrate concretely how emotions and biases might influence the translation process, I created a computer simulation program that embodies the ideas discussed above. The program writes two successive descriptions of an event in a way that is influenced by the emotions of the simulated writer.

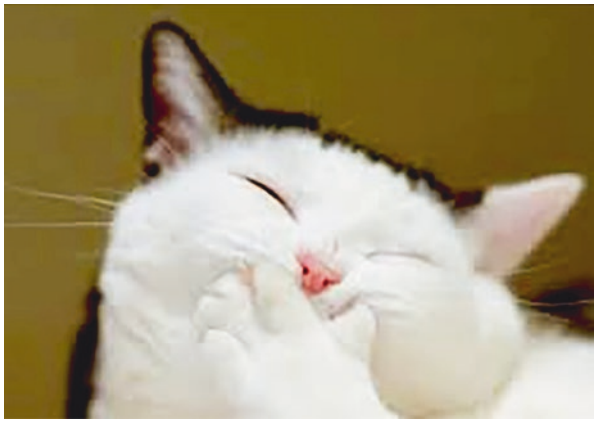


**Fig. 4** A titter

**Fig. 5** A side-splitter

## 4 The Event to Be Described

To provide something for the simulation model to describe, imagine this simple event:

> A shopper, Sam, walks through a flea market. He accidentally bumps the table of a merchant, George. Some of the objects on George's table fall to the ground and are damaged.

The simulation program provides written descriptions from each of three points of view: (1) Sam, the shopper, (2) George, merchant, and (3) Edna, a passer-by, who has a neutral point of view. Each of the characters writes a first and then a second description of the event, each of which is simulated by the program.

## 5 A Simulation Model for Composing Descriptions

I created a simulation model, programmed it in the Python language, and gave it three types of information:

- The three characters' emotional states
- The three characters' knowledge of the event
- A vocabulary for translating the writers' memories into language

1. ***The three characters' emotional states***

   George, the merchant, is assumed to be angry in degrees from 1 (mild annoyance) to 5 (enraged). Sam, the shopper, is assumed to feel guilty in degrees from 1 (mildly sorry) to 5 (deeply ashamed), and Edna is assumed to be neutral. Before running the program, a human operator informs the program of the identity of the narrator and the intensity of their emotion.

2. ***The three characters' knowledge of the event***

   As indicated in Fig. 2, each observer of the event forms a long-term memory representation of the event, which includes the agents, actions, and objects involved in the event together with their associated attributes. Some of the attributes—such as intensity, quantity, and value—can vary in magnitude from low to high. For example, the intensity of an auto accident can vary from *a fender-bender* (low) to a *wreck* (high). We assume that the initial representation is non-verbal since it is based on non-linguistic sensory input rather than on hearing a verbal description.

In Table 2, I have represented Edna's non-biased memory of the event. The table shows that Edna knows that Sam bumped George's table with an intensity of 0.3 on a scale of 0 to 1. (She may not know Sam's or George's names, but she can distinguish the men visually.) She knows that George's things, which she judges to have a value of 0.35, fell to the floor with an intensity of 0.7. Finally, she knows a quantity (magnitude 0.7) of George's things were damaged with an intensity of 0.6.

Since emotions can influence the narrator's initial representations, Sam's and George's memory for the event may well be different from Edna's. For example, George, the merchant, might represent the intensity of the bump and the damage as greater than Edna did. Further, the writer's focus on aspects of the event, such as agents or actions may also differ. I represented the writer's focus by placing focused parts of each element at the top of each column in Table 2. Thus, Edna's focus for element 1 is on Sam. George's focus, in contrast, might be on his table. I assume that the writer's focus can influence the form used in translating the memory into language. Therefore, because her focus is on Sam, Edna might write, "He bumped the table." George, who is focused on his table, might write instead, "My table was crashed into by some idiot."

As I have indicated in Fig. 3, the Translator also influences the form of the output. Thus, although the writer's focus in long-term memory might suggest one form,

**Table 2** Edna's memory for the event

| Chunk 1 | | Chunk 2 | | Chunk 3 | |
|---|---|---|---|---|---|
| Agent: Sam | | Object: things | | Object: things | |
| Action: Bumped | | Owner: George | | Owner: George | |
| Intensity: | 0.30 | Value: | 0.35 | Value: | 0.35 |
| Object: Table | | Action: Fall | | Quantity: | 0.70 |
| Owner: George | | Intensity: | 0.70 | Action: Damage | |
| | | From: Table | | Intensity: | 0.60 |
| | | To: Floor | | | |

the Translator might change that form. For example, the writer's focus might suggest he write, "I made serious mistakes," but because of fear of consequences, that form might be changed in the Translator to the passive voice, leading him to write, "Mistakes were made."

It is likely that all three writers would have different initial memories for the event. However, since the objective of the simulation is to illustrate the effect of emotion on lexical choice, all three writers have been assigned the same initial long-term memory representation of the event so that the effect of emotion on lexical choice can be directly compared across writers.

3. ***A vocabulary for translating the writers' memories into language***

   The vocabulary used by the simulation includes word families, each with words expressing a wide variety of intensities. I assigned each word a scale value from 0 (for least intense) to 1 (for most intense).[1] In Table 3, the second column for each family shows the intensities I assigned to each of the words.

## 6 What the Program Does

The simulation program carries out most of the activities represented in Fig. 3. It takes the information about the event stored in long-term memory and processes it chunk by chunk. It chooses a grammatical structure and words to express each of the chunks. The length of the text that the program produces, therefore, depends on the length and complexity of the representation of the event in long-term memory. Since the buffer is not implemented in this simulation, the program passes the language it creates directly to the transcriber, which prints out the text. It might be interesting to implement the buffer in a future simulation program because it could

**Table 3** Word families used by the simulation model (intensity shown after each word)

| Word families | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Bump | | Things | | Fall | | Quantity | | Damage | |
| Touched | 0.10 | Junk | 0.10 | Descended | 0.10 | Very few | 0.10 | Soiled | 0.10 |
| Nudged | 0.20 | Trifles | 0.15 | Glided | 0.20 | A couple | 0.12 | Scratched | 0.15 |
| Bumped | 0.30 | Trinkets | 0.35 | Dropped | 0.40 | Three | 0.15 | Chipped | 0.20 |
| Hit | 0.40 | Things | 0.50 | Fell | 0.50 | A few | 0.20 | Cracked | 0.30 |
| Struck | 0.45 | Collectables | 0.65 | Tumbled | 0.60 | Several | 0.40 | Damaged | 0.40 |
| Tripped over | 0.60 | Antiques | 0.75 | Plummeted | 0.70 | Many | 0.70 | Broken | 0.60 |
| Upset | 0.70 | Antiquities | 0.80 | Hurtled | 0.80 | Most | 0.80 | Smashed | 0.75 |
| Crashed into | 0.80 | Heirlooms | 0.90 | Crashed | 0.85 | Nearly all | 0.90 | Shattered | 0.85 |
| Smashed into | 0.90 | Treasures | 0.95 | Smashed | 0.90 | All | 0.95 | Destroyed | 0.90 |

[1] Different individuals would likely assign slightly different values to the words. What is important here is that individuals can assign intensity values to the words and that those values will generally agree with others in the linguistic community.

suggest an explanation for language bursts (Chenoweth & Hayes, 2001), the tendency for writers to produce texts in bursts of language separated by pauses.

When the human operator gives the simulation program the identity of the writer and the degree of the writer's emotional involvement, the program produces a written description based on the writer's representation of the event in long-term memory and on the intensity of the writer's emotion. If the writer is Edna, the program writes a neutral description of the event. If the writer is George, the program writes a description that emphasizes the seriousness of the accident. How much he emphasizes the seriousness will depend on how angry George is. If the writer is Sam, the program will minimize the seriousness of the accident, the guiltier Sam feels.

To understand how the program expresses emotion, take as an example how the program chooses words to express Sam's impact on George's table. First, the program chooses the appropriate word family to describe an impact. In Table 3, that is the "bump" family. Next, the program notes the intensity of the impact as recorded in the writer's memory. As shown in Table 2, the intensity was 0.30. (Remember, all writers are assumed to have the same memory.) Then, the program calculates a "preferred" intensity by adding or subtracting a quantity based on the strength of the narrator's emotion. With this calculation, Edna's preferred (and neutral) intensity would be greater than Sam's and less than George's. Having calculated the preferred intensity, the program then chooses the word that most closely matches the writer's preferred intensity.

The program's output (see Table 4) shows that George, who is angry, says that Sam "crashed" into the table; Edna, who is neutral, says he "bumped" into the table; and Sam, who is feeling guilty, minimizes the seriousness of the event and says that he "nudged" the table.

Some of the word choices in Table 4 seem unnatural, for example, g*lided* and *plummeted*. This occurred because I chose the words for use by the simulation program by consulting word lists from an on-line thesaurus. As a result, the program's output does not reflect the word frequencies of normal human speech. (If I had selected words on the basis of their English word frequency, the program's output would seem more like that of a native English speaker.)

**Table 4** Output of the simulation program for three levels of emotional bias

| Writer | First text | Second text |
|---|---|---|
| *Sam, who feels guilty,* wrote | I nudged the table. The trifles glided to the floor.<br>A few were chipped. | I touched the table.<br>The junk glided to the floor.<br>Three were scratched. |
| Edna, *who is neutral*, wrote | He bumped the table.<br>The trinkets plummeted to the floor.<br>Many were broken. | He bumped the table.<br>The trinkets plummeted to the floor.<br>Many were broken. |
| George, *who is angry*, wrote | He crashed into the table.<br>The antiques hurtled to the floor.<br>Most were smashed. | He crashed into the table.<br>The antiques hurtled to the floor.<br>Nearly all were destroyed. |

Once the first description is complete, the language chosen to describe the event is combined with the information about the event that is already in long-term memory to produce a new representation of the event. The program then produces a second description for each writer based on the modified representation currently in long-term memory. This allows us to compare a first and second text (see Table 4). Because of the change in the representation in long-term memory, the second text may differ in some details from the first. For example, as shown in Table 4, in Sam's first text, he says that he "nudged" the table, but in his second, he says that he merely "touched" it.

If this effect of descriptive writing on memory for an event can be verified for human writers, it is reasonable to expect that it might also be observed with any linguistic recounting of the event, such as describing it to a friend in conversation.

## 7 Discussion

In this chapter, I responded in a small way to Steve Graham's call for attention to the impact of emotion in writing. I have proposed two models to account for the impact of emotion on description. The first, a box-and-arrow model, is more loosely structured, but broader in scope. The second, a computer-simulation model, is more tightly structured, but narrower in scope. The broader model postulates three phases in the process of description:

Phase 1. An initial observation phase in which a representation of the event is stored in long-term memory.
Phase 2. A language production phase in which the writer constructs a written description of the event.
Phase 3. A memory modification phase in which the initial memory for the event memory is combined with the language produced in phase 2.

Each of these phases provides an opportunity for emotion to influence the descriptive process.

The computer-simulation model is consistent with the box-and-arrow model but aims for greater rigor. It provides a concrete working mechanism to account for how lexical choice is influenced by emotion and how the lexical choices can influence the writer's memory for an event. The central idea for the simulation was based on the human ability to distinguish among shades of meaning.

In designing the two models, I focused on the impact of emotions on descriptive writing. However, the ideas contained in the models may apply to a variety of affective states, such as attitudes, biases, dispositions, moods, and beliefs. They may also apply to other forms of writing, such as narration or argumentation.

It is important to recognize that the ideas I presented here are speculations that may be false. Even if the computer simulation does work, that does not mean that the simulated processes are the ones that humans use. It only shows that the proposed processes can produce the anticipated results. The models must be tested to be sure

they align with human performance. Perhaps people's lexical choices are not influenced by emotions. Perhaps the writer's memory for an event is not affected by what the writers writes. In either case, we would **not** expect writers' second description to differ from their first, contrary to the predictions of the models. Perhaps the effect on memory does not depend on writing. Or perhaps just thinking about the event could modify the memory for it. Writing might not be essential at all. Clearly, the models can and should be tested.

## *7.1 Why Bother with a Computer Simulation?*

Computer simulations can provide some benefits that box-and-arrow type models cannot. Simulations can give concrete embodiment to theoretical ideas. That is, simulations can demonstrate that procedures designated rather vaguely by an arrow or two in a box-and-arrow model can be converted into instructions for and made to work by a very literal-minded computer that won't put up with vagueness. This enforces a degree of rigor on the theorist, requiring that procedures and the flow of information be very clearly specified.

For example, in constructing a simulation model of children's knowledge telling (Hayes, 2011), I was forced to think very explicitly about the structure of the memory required to support the processes a child might use when writing and how that memory structure was to be accessed. The simulation presented here required me to provide a very specific mechanism for selecting words on the basis of the writer's preferred intensities. Of course, we can think rigorously without creating computer simulations, but creating them does help force us to depict processes and their relationships more explicitly.

The process of constructing a computer simulation can have other valuable side effects. For example, constructing the model described here made me pay attention to some issues that I might have thought about, but hadn't. It made me realize that a person with a very limited vocabulary would have difficulty expressing their emotions precisely in language. They may observe shades of meaning without having the range of lexical options to put them into language. And if they do attempt to express their emotions in written, spoken, or internal language, the words they choose may create an inaccurate memory of their own emotional experience. Nuance in expression and memory, available to those with rich vocabularies, may be denied to those without them.

Creating simulation programs is not every researcher's cup of tea. But for those who choose to do it, it can be a valuable source of information and can encourage precision in thinking about writing processes that constructing box-and-arrow models do not. In this chapter, I hope to have made a modest contribution to Steve's call to consider emotions in writing and to raise questions about how to model writing processes more generally.

## References


Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication, 18*, 80–98.

Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*(6), 1314–1329.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology, 34*(4), 590–598.

Forgas, J. P. (1999a). Feeling and speaking: Mood effects on verbal communication strategies. *Personality and Social Psychology Bulletin, 25*(7), 850–863.

Forgas, J. P. (1999b). On feeling good and being rude: Affective influences on language use and request formulations. *Journal of Personality and Social Psychology, 76*(6), 928–939.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Studies in writing: Vol. 4. Knowing what to write: Conceptual processes in text production* (pp. 139–160). Amsterdam University Press.

Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educational Psychologist, 53*(4), 258–279.

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323–344). Guilford Press.

Graham, S., & Harris, K. R. (2018). An examination of the design principles underlying a self-regulated strategy development study. *Journal of Writing Research, 10*(2), 139–187.

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744.

Graham, S., & Perrin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476.

Graham, S., & Sandmel, K. (2011). The process writing approach: A meta-analysis. *The Journal of Educational Research, 104*(6), 396–407.

Graham, S., Harris, K. R., & Troia, G. A. (1998). Writing and self-regulation: Cases from the self-regulated strategy development model. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 20–41). Guilford Publications.

Graham, S., Kiuhara, S., McKeown, D., & Harris, K. R. (2007). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 99*(3), 445–476.

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879–896.

Harris, K. R., & Graham, S. (1992). Self-regulated strategy development: A part of the writing process. In M. Pressley, K. R. Harris, & J. T. Guthrie (Eds.), *Promoting academic competence and literacy in school* (pp. 277–309). Academic.

Harris, K. R., & Graham, S. (2017). Self-regulated strategy development: Theoretical bases, critical instructional elements, and future research. In R. Fidalgo & T. Olive (Series Eds.) & R. Fidalgo, K. R. Harris, & M. Braaksma (Vol. Eds.), *Studies in writing series: Vol. 34. Design principles for teaching effective writing* (pp. 119–151). Brill.

Harris, K. R., & Graham, S. (2018). Self-regulated strategy development in writing: Premises, evolution, and the future. *BJEP Monograph Series II, No. 6 – Teaching and learning writing* (pp. 113–135).

Harris, K. R., Graham, S., Mason, L. H., & Saddler. (2002). Developing self-regulated writers. *Theory Into Practice, 41*(2), 110–115.

Harris, K. R., Graham, S., & Mason, L. H. (2006). Improving the writing, knowledge, and motivation of struggling young writers: Effects of self-regulated strategy development with and without peer support. *American Educational Research Journal, 43*(2), 295–340.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Randall (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Erlbaum.

Hayes, J. R. (2011). Kinds of knowledge-telling: Modeling early writing development. *Journal of Writing Research, 3*(2), 73–92.

Hayes, J. R. (2012). My past and present as writing researcher and thoughts about the future of writing research. In V. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 3–26). Psychology Press.

Leijten, M., van Waes, L., Schriver, K., & Hayes, J. R. (2014). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research, 5*(3), 299–337.

Loftus, E. F. (1975, Oct.). Leading questions and the eye witness report. *Cognitive Psychology, 7*(4), 560–572. https://doi.org/10.1016/0010-0285(75)90023-7

Loftus, E. F., & Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning and Verbal Behavior, 13*(5), 585–589.

Loftus, E. F., Miller, D. G., & Burns, H. J. (1978). Semantic integration of verbal information into a visual memory. *Journal of Experimental Psychology: Human Learning and Memory, 4*(1), 19–31.

Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology, 8*(2), 181–192.

Payne, B. K., Lambert, A. J., & Jacoby, L. L. (2002). Best laid plans: Effects of goals on accessibility bias and cognitive control in race-based misperceptions of weapons. *Journal of Experimental Social Psychology, 38*, 384–396.

Rogers, L. A., & Graham, S. (2008). A meta-analysis of single-subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879–906.

Unkelbach, C., Forgas, J. P., & Denson, T. F. (2008). The turban effect: The influence of Muslim headgear and induced affect on aggressive responses in the shooter bias paradigm. *Journal of Experimental Social Psychology, 44*(5), 1409–1413.

Zeelenberg, R., Wagenmakers, E., & Rotteveel, M. (2006). The impact of emotion on perception: Bias or enhanced processing? *Psychological Science, 17*(4), 287–290.

# Promoting Writing and Writing Engagement Through Self-Regulation, Motivation, Effective Instruction and *Perezhivanie*


**Clarence Ng and Peter Renshaw**

**Abstract** Graham's extensive research on writing has drawn attention to four engagement enablers in writing: self-regulation, motivation, reading-writing connections and effective writing instruction. Following a brief review of Graham's research on these enablers, we describe Graham's theoretical work on community and writing in his writer(s)-within writing model. In this theoretical discussion, Graham has taken a sociocultural turn, drawing attention to complex influences on writing and writing engagement derived from personal and social realms. Building on Graham's theoretical model, we re-conceptualise the personal and social realms in terms of Vygotsky's notion of perezhivanie, highlighting: (1) the dynamic unity of personal and situated influences on children's writing; (2) the interdependence of emotion and intellect in writing; and (3) students' sense-making and self-making during the writing process. We elaborate this theoretical perspective using two cases from our research projects. The first case discusses students' reflective writing following their experiences during a forest excursion designed as a narrative place-based pedagogy that engages children in imaginative role-play and inquiry activities. The second case was a design-based investigation examining the use of emotionally charged literacy activities to create a dramatic situation to re-engage a group of disaffected students in reading and writing activities. In each case, there is attention to the sense-making, self-making and emotionality arising from experience and its transformation across time/space.

**Keywords** Writing engagement · Writing motivation · Writing communities · Vygotsky · Perezhivanie · Identity

C. Ng (✉)
Australian Catholic University, Brisbane, QLD, Australia
e-mail: Clarence.Ng@acu.edu.au

P. Renshaw
University of Queensland, St Lucia, QLD, Australia



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_3

## 1 Introduction

Writing is important for learning, working and conducting everyday life in technology-driven societies where writing is required when we tweet, text and compose emails. Writing benefits learning because it 'provides students with a tool for visibly and permanently recording, connecting, analyzing, personalizing, and manipulating key ideas in text' (Graham & Hebert, 2011, p. 712). Despite its importance, not every student develops writing proficiency in school. Students who cannot write well will 'enact a toll', as poor writing abilities and skills may limit academic, career and personal development (Graham, 2006). There is considerable concern that many students fail to acquire adequate proficiency with writing by the time they complete their compulsory education. Large-scale tests in Australia, the UK and USA have shown that many students, especially those coming from various disadvantaged backgrounds, have fallen behind in their writing achievement and failed to develop writing skills. An important task for policy makers, educators and teachers alike is to develop different ways to promote writing, especially among students who find it challenging to learn to write due to various constraints derived from socioeconomic status, cultural backgrounds and physical conditions.

In this chapter, we adopt an engagement perspective to explore the issue of improving writing. We consider writing engagement an inherent part of the writing process, as engagement is required at every stage of writing wherein students plan, revise and edit their writing. We begin this chapter with a discussion on writing engagement, clarifying this concept and its importance to writing. We then provide a snapshot review of Graham's work with a focus on empirical evidence explicating how significant enablers facilitate and promote writing engagement. This is followed by a brief discussion of Graham's theoretical model of writers-within-community (WWC) that has the potentials to expand current research and promote a dynamic understanding of writing engagement. Building on this effort, we discuss the Vygotskian concept of perezhivanie and clarify how this concept may be used to promote writing and writing engagement. We discuss two cases derived from our work, elaborating how perezhivanie promotes engagement in writing.

## 2 Writing Engagement

Writing engagement can be understood as personal attributes and social constructions (Ng & Graham, 2018). As personal attributes, engagement in writing can be examined based on students' writing motivation, abilities to self-regulate writing and other significant personal characteristics critical for supporting writing. This individualistic conceptualisation aligns with Guthrie and colleagues (2004)'s description of reading engagement in terms of behavioral (time on task), emotional (enjoyment), and cognitive responses (use of cognitive strategies). More recently, Ng and Graham (2018, p. 617) has provided an alternative conceptualisation and

defined engagement in reading and writing as "emerging out of a social where individuals reciprocally interact with each other", aligning with a Vygotskian perspective on learning and development. This sociocultural understanding highlights the importance of situating engagement in the context of different literacy learning and activities. Both personal and sociocultural perspectives are important for understanding writing engagement, though, most research on writing engagement has taken an individualistic perspective focusing on personal attributes that enable engagement in writing. Graham's work has also followed an individualistic perspective. However, his recent theoretical work (Graham, 2018) has taken a social turn, drawing attention to writing and writing engagement as interactive processes situated in writing communities.

Why does engagement matter in writing? First, writing cannot be completed without sustained engagement. During the writing process, students need to plan, revise and edit their writing. All these writing actions require students to invest substantial effort and time. Second, writing is a complex task; it requires a host of complex knowledge and skills such as knowledge about written language, knowledge about the writing process and writing mechanics. Writing is both personal and social, as it serves not just personal goals such as interest in a topic but also social demands including fulfilling parents' expectation for writing achievement. This suggests that students need to know why they engage in writing and for what purposes they serve when they write. Third, engagement in writing is reciprocally related with motivation. Engaged writers are motivated writers. It follows that those who do not engage in writing will be deprived of significant motivation that supports writing, turning them into reluctant writers. Finally, writing engagement is an instructional issue that many teachers share, as disengaged students tend to expend limited efforts in writing. It is therefore important to develop innovative practices to promote writing engagement, especially for students who are reluctant to write or whose interest in writing has declined as they move up the grades.

## 3 Enabling Writing and Writing Engagement

Promoting writing and improving writing quality and achievement are perennial foci of Graham's research over the past four decades. Graham's work has often included assessment of writing engagement from an individualistic perspective using indicators such as time spent on writing, length of writing, and attitudes to writing. His work on writing intervention has showed that students' regulatory efforts is a particularly important form of writing engagement central to the writing process. Below, we discuss four important contributions to writing engagement based on Graham's extensive research on writing in the past four decades. These contributions include: (1) Self-regulation; (2) connections between reading and writing; (3) writing motivation; and (4) effective writing instruction.

## 3.1 Self-Regulation

An important cognitive resource for promoting writing and writing engagement is self-regulation which can be defined as an integrated set of cognitive, metacognitive, motivation and behavioral strategies that students use during writing. Self-regulation has been included as an inherent part of many writing models (e.g., Hayes & Flower, 1986). Given the complexity of writing, students' abilities to self-regulate the writing process is critical. To write effectively, students are expected to understand the writing process, know what is required of them as a writer, and develop relevant skills and strategies to manage the writing process and sustain engagement during writing. In a landmark review, Graham and Harris (2000) discussed empirical evidence supporting four important predictions related to self-regulation in writing: (1) Skilled writers are more self-regulated than less skilled writers; (2) Developing writers become more self-regulated with age and schooling; (3) Individual differences in self-regulation predict individual differences in writing; and (4) Teaching self-regulatory strategies improves the writing performance of developing and struggling writers.

Specific to writing engagement, self-regulation is important because it involves building blocks that enable writing and manage adjustment that needs to be made during the process in order that students continue to engage confidently (Graham & Harris, 2000). Graham's work has demonstrated clearly that teaching self-regulated strategies to struggling writers and students with writing difficulties benefit their writing engagement and writing quality. In addition, Graham's work has specifically highlighted the importance of providing continual support to self-regulated writing in order that positive effects gained from any self-regulatory interventions can be maintained and generalized (Harris & Graham, 1992).

Thus far, sources of motivation for self-regulating writing have been given limited attention in extant research on writing. The answer to this question is important for writing engagement because students who know what has made the writing process meaningful and motivating are more likely to devote time and effort to self-regulate their writing. In the same vein, it is important to understand why some students fail to self-regulate or use self-regulatory strategies ineffectively or are reluctant to deploy them in writing.

## 3.2 Writing Motivation

Graham and colleagues have affirmed the important role of motivation in promoting writing, engagement and writing achievement (Graham et al., 2017, 2019). A host of enabling motivators include writing self-efficacy (Graham et al., 1993), attitudes to writing (Graham et al., 2019) and goal setting (Page-Voth, & Graham, 1999) has been examined. Aside from these cognitive constructs, peer support has been shown to be an important social mechanism that supports writing and improves writing quality (De Smedt et al., 2020).

Graham's current work on writing motivation has focused on motives for writing, which are motivational incentives explaining why students write. The concept of writing motives draws attention to writing motivation that has not been generally examined in dominant frameworks such as achievement goals and self-determination models. For example, writing motives include not just intrinsic and extrinsic motivation but also regulatory motivation derived from affective responses to writing, such as feelings of being involved in writing and regulating emotions. This is important because students who write mostly for performance reasons in school may hold other motives for writing in out-of-school contexts.

Thus far, studies by Graham and colleagues (e.g., Camping et al., 2020; Graham et al., 2020; Limpo et al., 2020; Ng et al., 2021) have verified seven categories of writing motives, including curiosity (wanting to write for the interest in a writing topic), involvement (wanting to write in order to gain positive experiences), competition (wanting to write for outperforming others), grades (wanting to write for achieving better grades), emotional regulation (wanting to write in order to overcome negative emotions), reliefs from boredom (wanting to write to fill in time) and social recognition (wanting to write to seek praise and approval from others). American and Portuguese studies (Camping et al., 2020; Rocha et al., 2019) found that curiosity was an important motives associated positively with writing achievement and writing quality. The study (Ng et al., 2021) using Chinese samples drawn from Hong Kong and Shanghai showed that social recognition was an important writing motive, linking positively with writing self-efficacy. Importantly, these studies showed that not every writing motive is motivating to the same degree and that writing motives as cultural constructions may have different roles among different cultural groups.

## *3.3 Reading and Writing Connections*

Engaging successfully in writing requires not just motivation and regulatory skills, it also requires other important knowledge such as domain knowledge on a writing topic. In this context, Graham's research highlights the importance of reading and writing connections. This is because reading and writing share common knowledge and cognitive processes (Fitzgerald & Shanahan, 2000), which, according to Graham et al. (2018a, b), include domain knowledge, meta-knowledge about written language, procedural knowledge, and pragmatic knowledge of text attributes. Improving students' understanding of a topic through reading will improve writing, as both tasks rely on identical domain knowledge. Teaching functions and purposes of text, such as knowledge about text structures, helps students write as they can use such meta-knowledge to guide their writing. Teaching reading comprehension strategies enables students to develop strategies such as goal setting, analyzing, and summarizing, which are essential to writing. Finally, teaching knowledge about features of text, words, syntax, and usage helps both reading and writing, as students

use such knowledge in decoding/encoding and comprehending/constructing sentences. In short, improving reading will improve writing, and vice versa.

Based on a large number of intervention studies, Graham and colleagues' meta-analyses (2018a, b), alongside previous studies (e.g., Bus & van IJzendoorn, 1999; Olinghouse & Wilson, 2013) provided empirical support to reading-writing connections. In their meta-analytic study, Graham and colleagues showed that teaching reading improved writing, including writing quality, words written, and spelling. More importantly, the findings showed that the enabling effect of reading instruction on writing was maintained over time.

What does reading-writing connection mean for writing engagement? To write about a specific topic, students are required to have developed sufficient knowledge and understanding in the relevant domain. Otherwise, it will be hard for students to engage in the writing process or students will be reluctant to embark on a writing project in an area that they barely understand. Similarly, students may not be confident to write in a specific genre or construct text structures that they seldom come across in their reading. If students lack basic literacy skills, such as word knowledge, they will find it hard to engage freely in writing, as problems such as a lack of vocabulary and inability to construct sentences will interrupt their engagement. Improving reading and reading instruction should enable writing and writing engagement. Good readers make good writers; good readers are well-prepared to engage in writing (2018a, b).

## *3.4 Effective Writing Instruction*

Effective writing instruction enables writing engagement and promotes writing performance. Graham's research work in the past four decades has provided an empirical foundation supporting this important relationship (Graham, 2019). For example, in a meta-analytical study on adolescent students, Graham and Perin (2007a) found that writing interventions designed to focus on strategy instruction, summarization, peer assistance, goal setting, word processing, sentence combination, inquiry, prewriting activities, process writing approach, study of model writing and grammar instruction significantly improved writing quality. More significantly, Graham's collaborative work with Karen Harris on Self-regulated Strategic Development (SRSD) has verified and disseminated an evidence-based instructional approach for promoting writing, writing engagement and performance (e.g. Graham et al., 2005). Notably, SRSD has been widely tested with over 100 studies conducted by researchers in different countries.

Given the importance of writing instruction, it is critical to know how writing is being taught in classroom. Addressing this issue, Graham has conducted a series of national surveys involving samples of American teachers in both primary and high school settings (Brindle et al., 2016; Cutler & Graham, 2008; Gilbert & Graham, 2010; Graham et al., 2014; Ray et al., 2016). While these survey studies found that teachers used evidence-based instructional practices such as strategy instruction in

teaching writing, these practices were not used frequently and extensively. Insufficient time was given to writing and teaching of writing skills. There was also a lack of attention to students' writing motivation, connection between home and school practices and use of technologies in writing. Similar concerns regarding limited time on writing and infrequent instruction on writing skills were also reported in writing instruction surveys Graham and his colleagues conducted in other countries (Bañales et al., 2020; Graham et al., 2020; Hsiang et al., 2020).

From an engagement perspective, these survey findings suggest that students are ill-prepared to engage confidently in the writing process, as strategy instruction is infrequently delivered. More importantly, students are given limited opportunity to engage in writing because time spent on writing is generally infrequent and insufficient. Furthermore, issues related to out-of-school writing and writing using technologies indicate that teachers across different nations have failed to utilize additional support derived from home and new technologies to support students' writing engagement and achievement. In a review of writing intervention research, Graham and Perin (2007b, p. 330) stated that there is a need to "go beyond the classroom to study writing outside of school including how new technologies such as e-mail, text messaging, and innovations not yet realized impact writing development". Although more than a decade has passed since their review, Graham's call for attention remains valid, especially given that writing in social media has already established as a dominant form of writing for many adolescents and adults in online platforms.

## 4 The Writer(s)-Within-Community Model

Graham and Perin (2007b) discussed and highlighted the importance of capitalizing on research derived from cognitive and sociocultural understanding of writing. A decade later, Graham (2018) has achieved this theoretical connection in his elaboration of the writer(s)-within-community (WWC) model that draws on both cognitive and sociocultural studies of writing. In this model, complex influences derived from personal and contextual realms jointly and interactively affect students' writing, writing engagement and writing outcomes. Writing in this model is supported by two interrelated components – writing community and its individual member's own cognitive capabilities. Writing communities consist of a group of members (e.g., the teacher and students in a grade 6 class) who are governed by common norms and practices and produce writing to achieve shared purposes. While writing communities are flexible structures, each community has a shared physical space, social context and collective histories that enable or constrain writing activities and the production of writing. The second key component of the WWC model is cognitive capacities and resources of individual members in a writing community. Individual members apply control mechanisms such as attention and executive control to regulate their writing actions, as they compose a text to achieve a shared purpose. Cognitive capabilities related to writing mechanics, ideation, regulation, knowledge

and motivation are important cognitive resources that afford or constrain the writing process and its resulting writing products.

The WWC model offers a situated view on writing and draws attention to its complexity due to dynamic influences derived from both the person and context. Although Graham did not explain how writing engagement occurs in the WWC model, it can be assumed that writing engagement is both social and personal. This is because students' engagement in writing is constrained simultaneously by their own cognitive capacities and the characteristics of the writing community where writing is situated. In this sense, engagement in writing depends not just on the presence of developed and refined cognitive enablers such as self-regulation and motivation, but also the characteristics of a writing community, especially the shared purpose of writing. For example, intrinsically motivated writers, compared to extrinsically motivated counterparts, may be less likely to engage in writing in a performance-focused classroom where writing is completed predominantly for assessment purposes. In short, the WWC model provides a balanced treatment of personal and contextual influences on writing and offers a theoretical lens enabling a detailed analysis of complex influences derived from these arenas (see Graham, 2021). Thus far, empirical research guided by the WWC model is at the beginning stage (e.g., Ng et al., 2021).

## 5 Perezhivanie and Writing

Graham and Perin (2007b, p. 330) made the comment that 'we cannot rest on what is already known' when contemplating future research on writing intervention. In alignment with this spirit of advancing research on writing, we devote the remainder of this chapter to the Vygotskian concept of perezhivanie and explore how it can be used to promote writing and writing engagement. From a perezhivanie perspective, writing can be taken as a means to reflect on emotionally charged events where past experiences are relived and future experiences imagined anew. Similar to Graham's WWC model, perezhivanie draws attention to the interaction between the person and the context, but unlike the WWC model, Vygotsky does not separate the person from the context. He takes emotional lived experience – perezhivanie – as an indivisible unit of analysis.

What is perezhivanie? Perezhivanie is widely understood in recent scholarship as emotional lived experiences (e.g. Blunden, 2016; Ng & Renshaw, 2019; Renshaw & Tooth, 2016; Roth & Jornet, 2017; Veresov & Fleer, 2016). In 'The Psychology of Art', Vygotsky (1971) initially linked perezhivanie to dramatic and artistic experiences, both on stage as actors relive and recreate emotional moments, and in real life as audience members reflect on their everyday lives through the lens of the dramatic experience and emotional engagement. At the end of his life, Vygotsky elaborated perezhivanie in 'The Problem of Environment' to propose the 'indivisible unity' between the person and environment to understand child development (Vygotsky, 1994, p. 342). In other words, it is not the environment that is critical for

understanding child development, per se. It is the child's subjective understanding and emotional reactions that are critical in shaping their development.

Vygotsky (1994, p. 339) states that, "the emotional experience (perezhivanie) arising from any situation or from any aspect of his [sic] environment determines what kind of influence this situation or this environment will have on the child. Therefore, it is not any of the factors in themselves (if taken without reference to the child) which determines how they will influence the future course of his development, but the same factors refracted through the prism of the child's emotional experience (perezhivanie)." Using a metaphor of prism, Vygotsky highlights the critical role of refraction in understanding the relationship between perezhivanie and development. The process of refraction draws attention to significant factors and conditions presented in both internal and external realms that an individual might deploy to make sense of an experience. Perezhivanie therefore represents how an individual 'becomes aware of, interprets, [and] emotionally relates to a certain event' (Vygotsky 1994, p. 341). Roth and Jornet (2017) described perezhivanie as a double description (Roth & Jornet, 2017), involving both a specific experience and how it is interpreted (Ng & Renshaw, 2019). Blunden (2016) explains that perezhivanie is always associated with personally significant events or moments that are heightened with emotional responses. In other words, perezhivanie is not just any experience but those that are refracted, emotionally-charged and personally significant.

The key consideration is how the concept of perezhivanie informs writing and promotes writing engagement. Vygotsky maintains that 'art is the social technique of emotion, a tool of society which brings the most intimate and personal aspects of our being into the circle of social life'. (Vygotsky, 1971, p. 249). Writing fits this conceptualisation as an art form whereby past experiences are encapsulated, retold and shared in a textual form through creative orchestration of words, texts and images. From a perezhivanie perspective, writing is both *experiential* and *reflective*. Perezhivanie itself provides the content for writing, as it involves an intimate knowledge of one's experiences that are considered personally significant. Through writing, students describe and share these significant personal moments and experiences. Writing also provides a chance for students to reflect and make further sense of these important moments in their life and enables them to construct meaning (i.e. refract) through words, images and signs that are made available in different modes of writing. This experiential process of writing is engaging, as it involves an authentic experience that is personally significant.

From a perezhivanie perspective, writing is a refractive *sense-making* process. Through writing, lived experiences are not just recounted, but reshaped, represented, and communicated to an audience. This is achieved through retelling that involves purposeful selection of words, sequencing of ideas, and communicating of personal meanings and significance. Therefore, writing is not simply a description of one's experiences, but more importantly, offers a chance to relive and reinterpret these experiences in light of other important considerations associated with a writing task, such as addressing a specific audience. Writing as a sense-making process

promotes continuous engagement. It is highly likely that students are committed to retell and share their experiences that they consider important.

Writing is not just sense-making but also *self-making* (cf. Ng & Renshaw, 2019). Through writing, students develop a better understanding of who they are, based on their reflection and recount of past significant experiences. Writing enables students to become authors where they can take-up a range of "I" positions as the subjects of their own stories. Through their authorship they can explore past, present and possible future selves, elaborating how specific events and experiences have contributed to an identity they embrace or dislike. As an author, they have the opportunity to reflect on the personal and environmental influences that have contributed to their ontological journey. In summary, writing offers a chance to reflect on the self, its past constructions and future possibilities and enables a self-making process, as students re-present their past and build connections with the present and future. This self-making process through writing is motivating, as it turns writing into an ontological project for developing students' identities in areas that they consider personally significant.

Below, we illustrate how perezhivanie promotes writing and writing engagement using two examples drawn from our studies. These cases illustrate how writing and writing engagement could be supported, as students relived and retold their experiences.

## 6 Case 1: Written Reflections Following the Forest Excursion

The context for the first case study is an environmental education program for primary school children that is based on a narrative pedagogy called *storythread* (see Tooth Wager & Proellocks, 1988; Tooth & Renshaw, 2009; Renshaw & Tooth, 2018). S*torythread* programs involve three key activities: first, taking children to a specific place where natural systems, flora and fauna, and ecological processes can be experienced; second, engaging children in dramatic stories based on the history of the place; and third, introducing children to attentiveness strategies such as the Indigenous practice of *dadirri*, in order to deepen the children's sensuous and reflective connection to place (see Ungunmerr-Baumann, 1988, 2002). *Dadirri*, as practiced in the *storythread* program, involves sitting quietly by oneself for some time (15–20 min), listening and being receptive to the surrounding environment, and afterwards sharing the experience with others.

The written extracts analysed below were written by children after a *storythread* program that occurred at Karawatha, a 1000 hectare forest that includes endangered fauna and flora. The central character in the narrative pedagogy is Bernice Volz, a local environmentalist, who advocated with other local citizens in the 1990s to save the forest from development as a housing estate. The children conclude by writing a letter to Bernice to describe their experiences and reflect on what they have learned from Karawatha. The analysis below has been organised to show the experiential

nature of writing, writing as sense-making and as self-making from a perezhivanie perspective.

(i) Writing that brings personal aspects of our being into the circle of social life.

The extract below written by a Year 6 girl is emotional, personal and poetic. Her *perezhivanie* involved both learning about Bernice and Karawatha ("*I now know why you wanted to save this amazing place*") and reflecting on herself and her life.

*Extract 1*

*I now know why you wanted to save this amazing place. In class I never really knew what you meant …This place is like singing my favourite songs. I love singing more than anything! Singing is me. You're very right. There is magic that falls with the leaves and blows with the wind. The such uniqueness of each and every tree. The way 5 leaves are never the same. I found inside this sacred place that all my worries disappeared with the sights and sounds of the city. When we sat still everything that I saw as ordinary I found something so amazing … Something inside me made me want to stay there forever. Watching the insects and colourful birds swoop and glide through the clean air. I've had a hard life but not as hard as others so in that moment I was reminded of the beautiful things and people that life has given me.*

As she describes various experiences in the forest (uniqueness of trees and leaves, and birds gliding through the air) she also reveals her interior self ("*something inside me*"). We now know that she worries about things ("*all my worries disappeared"*) and that singing is central to her identity ("*singing is me*"). The key insight occurred toward the end of the extract when she is moved to reflect on her life in a general way ("*I've had a hard life but not as hard as others*") but her reflection is reshaped by the experience of Karawatha ("*so in that moment I was reminded of the beautiful things and people that life has given me*"). These are very personal aspects of her life that she has chosen to bring into the circle of social life through writing. There is an interlacing of her actual experiences in Karawatha with personal reflections on her interior life. It is this unit – her perezhivanie – that is both enabled and revealed through her writing. The communicative act of writing brought her interior world into the circle of social life that, she reminds herself, is actually "beautiful".

(ii) Writing as *sense-making* involves reflection on lived emotional experience (refraction) that reshapes its significance for the writer.

*Extract 2*

*I love how you spent a lot of your time dedicated to this place and never gave up on your dreams. Karawatha is an amazing forest and I'm so grateful to you for helping this wonderland survive the threats humans pose. You have inspired so many, me included! This is one place I will never forget & hold dear in my heart. Karawatha has made me realise that there is always beauty around us. Whether it's your surroundings or the people who you*

*hold close. The environment is a part of everyone, even if they know it or not. I absolutely fell in love with Karawatha when I first saw it. All the different zones in the one forest, it's stunning! And you would never believe that this giant heartland is sandwiched in between two major cities! I love to write & I would like to be an author when I grow up, and I know that Karawatha will be featured in a lot of my stories, describing its beauty and gracefulness and how one woman helped it through its journey and helped it develop. You really inspire me and even though I might not save a 1000 hectare forest in my lifetime, I know that determination and never giving up is a major key into achieving your dreams. You helped me see that. Thank you.*

This letter to Bernice by a Year 6 girl reveals her sense-making, as she reflects on Bernice and her experiences in the Karawatha forest. Her experiences were emotional, as revealed in the way she describes the forest as "amazing", as a "wonderland", as "stunning" and her response as, "I absolutely fell in love with Karawatha." But it is her reflections on these experiences that indicate her deeper sense-making and the personal significance of the experience. She connects the beauty of Karawatha to a reflection on people close to her ("…*there is always beauty around us*…") and to her goal of becoming an author ("*I would like to be an author when I grow up*"). She concludes by explicitly noting a significant insight about how to achieve her dream ("*I know that determination and never giving up is a major key to achieving your dreams").* Her perezhivanie is not bounded by the actual events and experiences in Karawatha. Rather, her perezhivanie extends beyond these concrete experiences to her subsequent sense-making in the act of writing. In her letter, she interlaces experiences in Karawatha with her goal of being an author and the personal qualities of determination and persistence she will need to achieve her dreams. Again, it is this unit of experience and reflection – her perezhivanie – that is enabled and revealed in writing.

(iii) writing is also a *self-making* process where students can explore past, present and possible future selves.

The extracts below were chosen for their explicit references to self-making.

*Extract 3*

*I never did want to be an environmental advocate, but since I walked through Karawatha it's changed me. I now think more about trees and how they dance in the wind. Birds and other animals how they're so free. I also think more about not dropping litter.*

In Extract 3 the writer, a Year 6 boy, is honest in stating that, unlike Bernice, he had never wanted to be an advocate for the environment but that the experience in Karawatha had changed him. While the text is more direct and less emotive than those analysed above, there are close observations of trees ("*dancing"*) and birds ("*being free"*) arising from his experience. We learn about his "changed self" and his intention to not drop litter.

*Extract 4*
*It is very hard to say or describe what I feel like I am a new girl and a new person. Karawatha is one of the most beautiful places I have ever been and seen and as I said it is hard to describe but one thing pops (into) my head. Thank you thank you for saving Karawatha. My favourite zone was the sandstone ridge. I will always remember this as my first day as an environmental advocate and my first day doing my new favourite hobby, being with nature.*

In extract 4 the writer, a Year 6 girl, describes how she feels as "a new girl and new person" due to her experiences in Karawatha. The sense of a changed self is heightened by her deployment of the future to reflect back on her time at Karawatha (*"I will always remember this as…"*). In this section of the text she imagines her future self reflecting on how she became an environmental advocate ("*my first day as an environmental advocate*").

As Blunden (2016) has proposed, perezhivanie can be theorised as those emotional lived experiences in a person's biography where significant choices were made. In the cases above, we do not have other evidence to know if these lived experiences actually became important steps in an unfolding biography. What we can say is that the emotional experiences in Karawatha, refracted through the process of writing, engaged children in an explicit process of self-making. They sensed new possibilities for themselves and different practices arising from their perezhivaniya.

## 7 Case 2: Emotionally Charged Writing

In this case, a class of Year 4 students from a primary school situated in a low SES suburb in Queensland participated in a partnership project designed to promote literacy learning with a focus on emotional well-being. These students were underachievers in literacy learning. The theme was selected by the teacher who was concerned about her students' abilities to express their feelings comfortably as most of them came from migrant families where English was used as a second language. The program was designed with a series of learning activities that aimed to promote students' learning of emotion-related vocabulary, confidence in sharing their personal feelings and engagement in writing about their feelings and personal experiences. These learning activities included: working on semantic maps of emotion-related vocabulary, sharing of and writing about an emotive moment based on a selfie, and writing about the feelings of a character in a novel. These activities took about two terms (6 months) to complete and were incorporated in the English lessons.

Findings derived from this study showed that students' engagement in writing has improved substantially. This was revealed in the length of their writings, the number of sentences they composed, and the number of emotive words they used

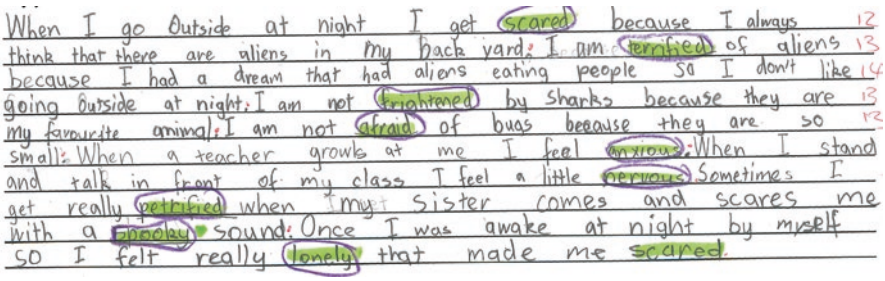

When I go Outside at night I get scared because I always 12
think that there are aliens in my back yard. I am terrified of aliens 13
because I had a dream that had aliens eating people so I don't like 14
going Outside at night. I am not frightened by sharks because they are 15
my favourite animal. I am not afraid of bugs because they are so 16
small. When a teacher growls at me I feel anxious. When I stand
and talk in front of my class I feel a little nervous. Sometimes I
get really petrified when my sister comes and scares me
with a freaky sound. Once I was awake at night by myself
so I felt really lonely that made me scared.

**Fig. 1** Paul's fear

(see Appendix 1 for four examples), which also suggests that students' engagement in writing has improved over the program. An analysis of students' post-program writings showed that students' improved engagement was associated with the opportunities to reliving and reinterpreting their emotional experiences (sense-making) in writing. Such perezhivanie-focused writing requires students to choose carefully relevant events or experiences and to present them in a coherent way. Additional to authentic experiences, this may include the use of imagined experiences and personal perceptions to sharpen the focus. Paul's writing on fear (Fig. 1) illustrates this form of writing engagement. He put together real moments of fear he had experienced, such as walking in the backyard after dark and waking up alone up at night. Between these fearful experiences, he listed amusing descriptions revealing that he was not afraid of sharks and bugs but would be succumbed to 'a teacher who growls'. This adds a sense of dramatic feeling to his writing. When he was asked why he structured his writing in this way, Paul said he wanted others to feel his feelings that 'teachers sometimes are more scary than a shark'.

Perezhivanie-focused writing promotes engagement with the self. It gives students the opportunity to reflect on their self-understanding (self-making). In the second example (Fig. 2), Peta described an emotionally charged moment where her father took her laptop away from her. Peta explained that her father would sometimes take her laptop away if she used it for too long, which she would seldom complain. Nevertheless, in this particular perezhivanie, she was furious because she was in the moment of enjoying her favourite song. In recalling this instance, she imagined herself as 'a rage monster' attacking a village 'full of her dads'. Peta explained that her writing was designed to make others feel her rage and understand how important it was for her to use the laptop to access YouTube videos for fun.

Joana's writing (Fig. 3) is another example showing how perezhivanie focused writing improves self-understanding. Joana relived mixed emotions she experienced during her birthday and explained why it turned out to be the best birthday she had ever had. Joana was upset because no one seemed to remember her birthday. This of course was due to her own expectation and the way that birthdays were being celebrated among her family and her church friends. She cried when her expectation

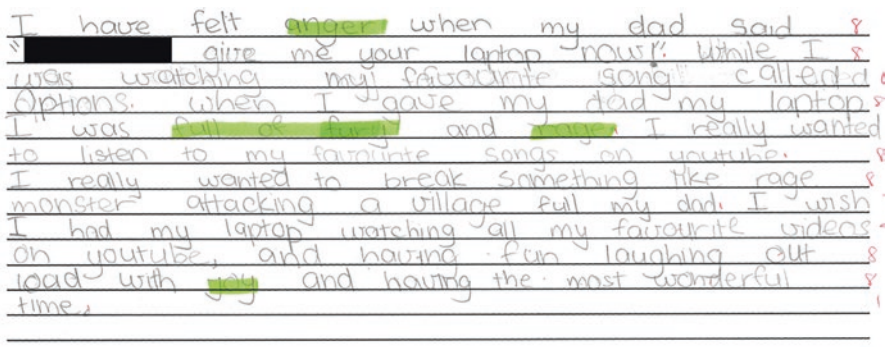

I have felt anger when my dad said "[redacted] give me your laptop now!". While I was watching my favourite song I called options. when I gave my dad my laptop I was full of fury and anger. I really wanted to listen to my favourite songs on youtube. I really wanted to break something like rage monster attacking a village ful my dad. I wish I had my laptop watching all my favourite videos on youtube, and having fun laughing out load with joy and having the most wonderful time.

**Fig. 2** Peta's rage

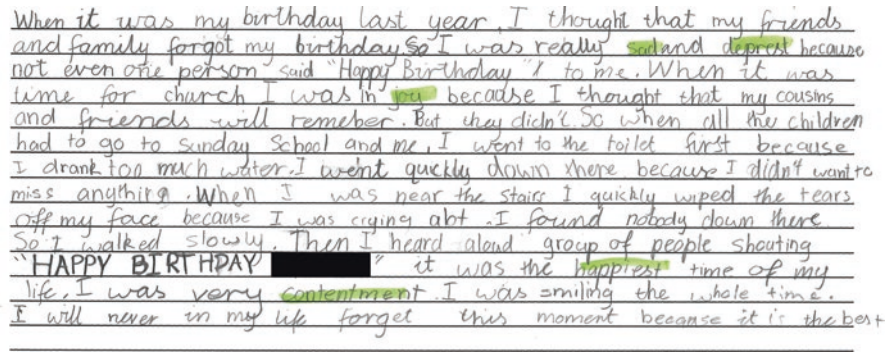

When it was my birthday last year, I thought that my friends and family forgot my birthday so I was really sad and depress because not even one person said "Happy Birthday" to me. When it was time for church I was in joy because I thought that my cousins and friends will remeber. But they didn't. So when all the children had to go to sunday School and me, I went to the toilet first because I drank too much water. I went quickly down there because I didn't want to miss anything. When I was near the stairs I quickly wiped the tears off my face because I was crying abt. I found nobody down there. So I walked slowly. Then I heard aloud group of people shouting "HAPPY BIRTHDAY [redacted]" it was the happiest time of my life, I was very contentment I was smiling the whole time. I will never in my life forget this moment because it is the best

**Fig. 3** Joana's happiest moment

was not being fulfilled, which ironically, turned a simple shout of 'happy birthday' from her church friends into a memorable moment of joy. In this writing, Joana arranged the sequence of events in such a way that others would be able to understand the significance of a simple shout of celebration to her.

Similarly, Peter carefully sequenced the details to magnify his feeling of sadness associated with moving to a new school. In this writing (Fig. 4), Peter relived his sadness in an orderly manner. First it was the unhappiness associated with farewelling the teachers. This was then followed by the saddest moment of leaving his friends. His deliberate compilation of four emotional adjectives amplified his feeling of sadness. At the end, making new friends in the new school seemed to have softened this sadness. However, the way he ended his writing with the word 'kinda' communicated a deep sense of loss, suggesting that he still emotionally attached to his old friends.

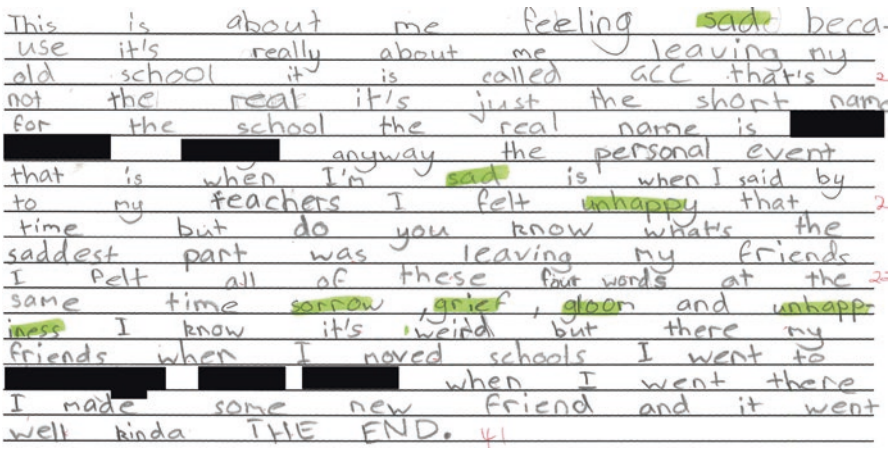



**Fig. 4** Peter's sadness

## 8 Writing *About* Perezhivanie and Writing *as* Perezhivanie

From a perezhivanie perspective, writing is experiential, simultaneously personal and social, and highly reflective. Our cases illustrate that perezhivanie promotes writing and writing engagement in two distinctive ways: writing about perezhivanie and writing as perezhivanie.

*Writing about perezhivanie* highlights the experiential nature of writing that draws its content from one's significant experiences, as illustrated by the extracts from both cases. In contrast to typical school writing tasks that can be personally irrelevant or restricted to prescribed content, perezhivanie-focused writing provides students with emotional and self-related sources to draw upon. Students are motivated by a desire to tell, share and reflect on their experiences related to concrete events or significant moments of heightened emotional engagement. This form of writing is emotionally-charged, reflective and forward looking (cf., Gutiérrez, 2008). It enables students to explore their authorial voices, express personal meanings and reflect on their past and future identities through writing (Ivanic, 1998). While instruction on writing mechanics remains important, the extracts in Case 2 indicate that perezhivanie-focused writing for struggling students motivates their sustained engagement, as they are eager to tell their personal stories and to share them with their peers.

Perezhivanie-focused writing provides a chance to relive past experiences, formulate new interpretations (sense-making), and reorganise the self (self-making) (Renshaw & Tooth, 2016). We called this, *writing as perezhivanie*. As students engage in the process of writing about their emotional lived experiences, they need to craft a written account by choosing appropriate words and emotionally charged descriptions, by selecting important details and highlighting dramatic events, and

by reflecting on the meaning of the events in relation to the audience and in the context of their own lives. Students' texts in both Case 1 and Case 2 illustrate how *writing as perezhivanie* offers an opportunity to revisit and 'live through' one's past experiences, enabling the exploration of constitutional elements, both personal and social, inherent to a perezhivanie. Case 1 showed how students relived memorable events from the excursion, bringing different aspects of their life – past, present and future – into their recounts. Specifically, the requirement to write to Bernice, the environmental advocate, meant that the students needed to tailor their writing to address her personally as they revisited their significant experiences, and more importantly, they needed to re-interpret and elaborate their experiences in relation to their own lives. In this way, writing to Bernice itself became an emotional and memorable event in the lives of the children. *Writing as perezhivanie* is, in fact, a meta-experience (Smagorinsky, 2011). It promotes and sustains engagement in a textual space where emotionally charged experiences are curated, re-enacted and refracted. This form of writing engagement is ongoing, reflective and highly personal.

## 9 Conclusion

Graham's writing research in the past four decades has made a significant contribution to advancing our understanding of writing, teaching of writing and writing achievement. From an engagement perspective, Graham's work has demonstrated the importance of self-regulation, motivation, knowledge and effective instruction to writing and writing engagement. Additional efforts are required to further our understanding on the complexity of writing and writing engagement in light of challenges posed by new technologies that Graham and Perin (2007b) pointed out more than a decade ago. In this regard, Graham's current theoretical work (2018a, b) provides a way to conceptualise online writing communities in social media and different internet platforms and to consider the influences derived from both personal and contextual realms on online writing and engagement. Graham's WWC model also provides a theoretical foundation for the development of writing engagement models based not just on cognitive enablers but also significant influences derived from social relationships, interaction, norms and practices.

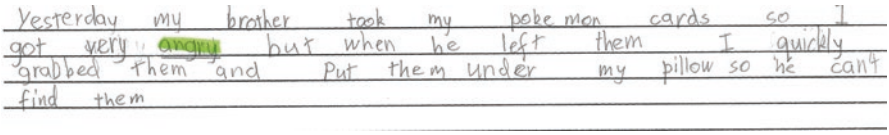
In this chapter, we added to Graham's writing research using the Vygotskian concept of perezhivanie, explicating how this concept can be used to research writing engagement, as illustrated in our cases of research. The cases that were analysed above indicate that students' writing about their personally significant experiences is cognitively demanding and emotionally charged. Students are required to relive their past experiences, consider important constitutional conditions that affect how they feel, how they make sense of what has happened, and how to retell these experiences within the constraints of the specific writing task. As we have shown in our analyses, students' writings drawn from these cases bring highly personal aspects of

students' interior life into the circle of social life, engage them in a refractive sense-making process, and enable them to consider their changing identities and sense of self across time. *Writing about perezhivanie* and *writing as perezhivanie* are processes through which we can come to know and to create ourselves and our social worlds. This is at the core of the educational enterprise. Interestingly, Graham (2021), in a reflective account of his lifelong career on writing research, has engaged in such a form of perezhivanie-focused writing, reliving emotionally charged moments in his life (writing about perezhivanie) and reflecting on how these significant moments have led to his life-long commitment to writing research (writing as perezhivanie).

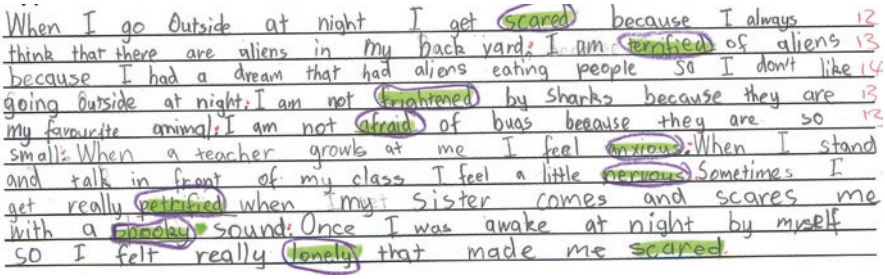
## Appendix 1: Examples of Pre- and Post-program Writings

### *Paul's Emotionally Charged Writing*

#### Pre-program Writing

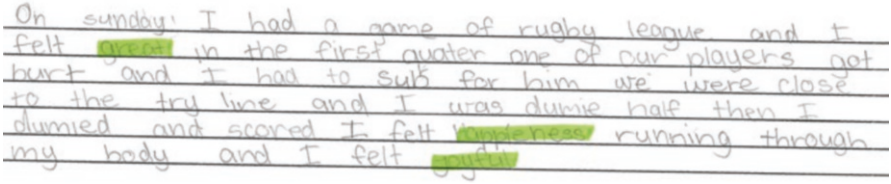

Yesterday my brother took my poke mon cards so I got very angry but when he left them I quickly grabbed them and Put them under my pillow so he can't find them

#### Post-program Writing

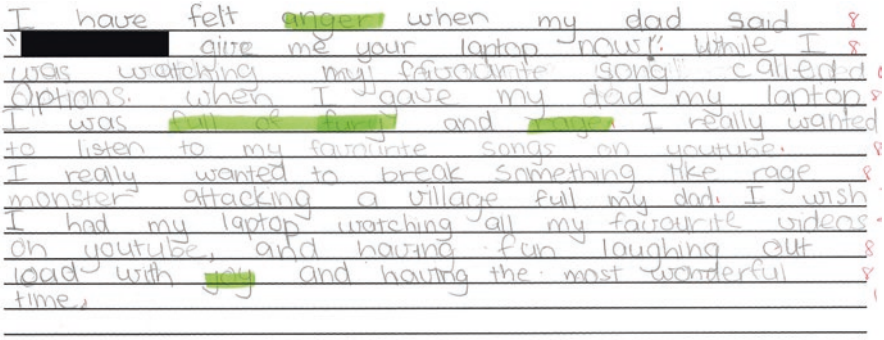

When I go Outside at night I get scared because I always think that there are aliens in my back yard. I am terrified of aliens because I had a dream that had aliens eating people so I don't like going Outside at night. I am not frightened by sharks because they are my favourite animal. I am not afraid of bugs because they are so small. When a teacher growls at me I feel anxious. When I stand and talk in front of my class I feel a little nervous. Sometimes I get really petrified when my sister comes and scares me with a spooky sound. Once I was awake at night by myself so I felt really lonely that made me scared.

## *Peta's Emotionally Charged Writing*

### Pre-program Writing

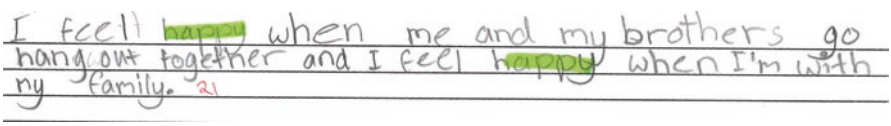

On sunday I had a game of rugby league and I felt great in the first quater one of our players got hurt and I had to sub for him we were close to the try line and I was dumie half then I dumied and scored I felt [redacted] running through my body and I felt joyful

### Post-program Writing

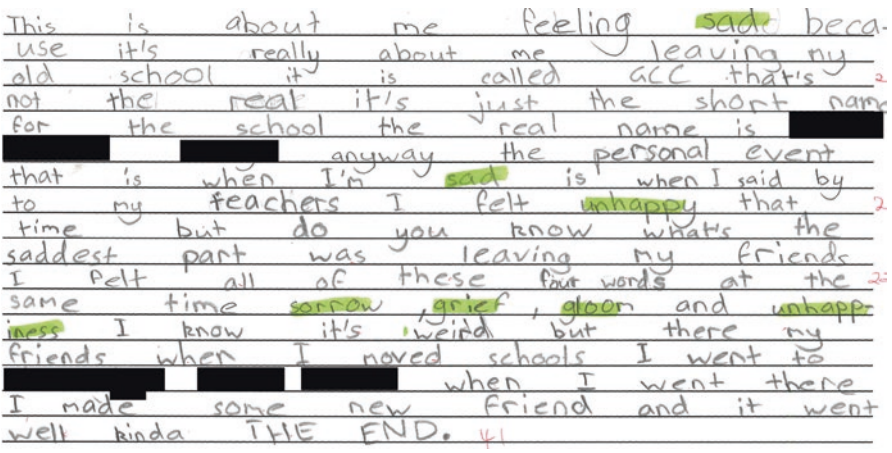

I have felt anger when my dad said "[redacted] give me your laptop now!" While I was watching my favourite song" called Options. when I gave my dad my laptop I was full of [redacted] and [redacted]. I really wanted to listen to my favourite songs on youtube. I really wanted to break something like rage monster attacking a village full my dad. I wish I had my laptop watching all my favourite videos on youtube, and having fun laughing out load with joy and having the most wonderful time.

## *Peter's Emotionally Charged Writing*

### Pre-program Writing

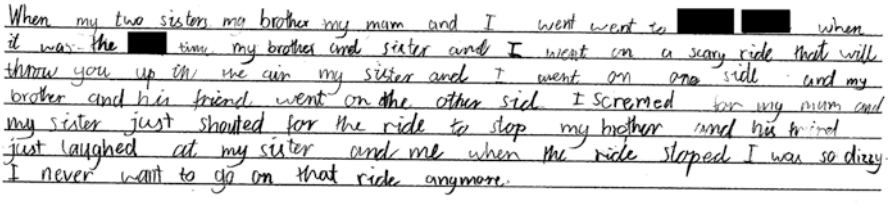

I feell happy when me and my brothers go hang out together and I feel happy when I'm with my family.

**Post-program Writing**

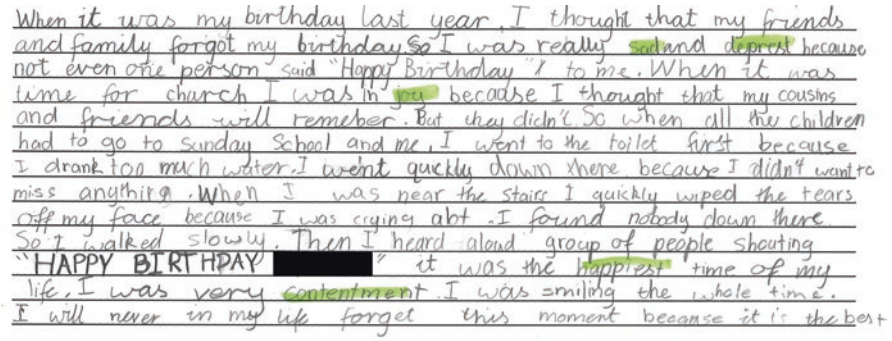

This is about me feeling sad because it's really about me leaving my old school it is called acc that's not the real it's just the short name for the school the real name is [redacted] anyway the personal event that is when I'm sad is when I said by to my teachers I felt unhappy that time but do you know what's the saddest part was leaving my friends I felt all of these four words at the same time sorrow, grief, gloom and unhappiness I know it's weird but there my friends when I moved schools I went to [redacted] when I went there I made some new friend and it went well kinda THE END.

## *Joana's Emotionally Charged Writings*

**Pre-program Writing**


When my two sisters my brother my mam and I went went to [redacted] when it was the [redacted] time my brothers and sister and I went on a scary ride that will throw you up in the air my sister and I went on one side and my brother and his friend went on the other sid I scremed for my mum and my sister just shouted for the ride to stop my brother and his friend just laughed at my sister and me when the ride stoped I was so dizzy I never want to go on that ride anymore.

**Post-program Writing**


When it was my birthday last year, I thought that my friends and family forgot my birthday so I was really sad and depress because not even one person said "Happy Birthday" to me. When it was time for church I was in joy because I thought that my cousins and friends will remeber. But they didn't. So when all the children had to go to sunday School and me, I went to the toilet first because I drank too much water. I went quickly down there because I didn't want to miss anything. When I was near the stairs I quickly wiped the tears off my face because I was crying abt. I found nobody down there. So I walked slowly. Then I heard aloud group of people shouting "HAPPY BIRTHDAY [redacted]" it was the happiest time of my life, I was very contentment. I was smiling the whole time. I will never in my life forget this moment because it is the best

## References


Bañales, G., Ahumada, S., Graham, S., Puente, A., Guajardo, M., & Muñoz, I. (2020). Teaching writing in grades 4–6 in urban schools in Chile: A national survey. *Reading and Writing, 33*, 2661–2696.

Blunden, A. (2016). Translating *Perezhivanie* into English. *Mind Culture and Activity, 23*(4), 274–283.

Brindle, M., Graham, S., Harris, K. R., & Hebert, M. (2016). Third and fourth grade teacher's classroom practices in writing: A national survey. *Reading and Writing, 29*(5), 929–954.

Bus, A., & van IJzendoorn, M. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91*, 403–414.

Camping, A., Graham, S., Ng, C., Atken, A., Wilson, J. M., & Wdowin, J. (2020). Writing motivational incentives of middle school emergent bilingual students. *Reading & Writing, 33*(9), 2361–2390.

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology, 100*(4), 907.

De Smedt, F., Graham, S., & Van Keer, H. (2020). "It takes two": The added value of structured peer-assisted writing in explicit writing instruction. *Contemporary Educational Psychology, 60*, 101835.

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*, 39–50.

Gilbert, J., & Graham, S. (2010). Teaching writing to elementary students in grades 4–6: A national survey. *The Elementary School Journal, 110*(4), 494–518.

Graham, S. (2006). Writing. In P. Alexander & P. Winne (Eds.), *Handbook of Educational Psychology* (pp. 457–478). Erlbaum.

Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educational Psychologist, 53*(4), 258–279.

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education, 43*(1), 277–303.

Graham, S. (2021). Through the looking glass: Reflections of a writing scholar. *Education Review, 28*, 1–21. https://doi.org/10.14507/er.v28.3209

Graham, S., & R. Harris, K. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational psychologist, 35*(1), 3–12. https://doi.org/10.1207/S15326985EP3501_2

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744.

Graham, S., & Perin, D. (2007a). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445.

Graham, S., & Perin, D. (2007b). What we know, what we still need to know: Teaching adolescents to write. *Scientific Studies of Reading, 11*(4), 313–335.

Graham, S., Schwartz, S. S., & MacArthur, C. A. (1993). Knowledge of writing and the composing process, attitude toward writing, and self-efficacy for students with and without learning disabilities. *Journal of Learning Disabilities, 26*(4), 237–249.

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*(2), 207–241.

Graham, S., Capizzi, A., Harris, K. R., Hebert, M., & Morphy, P. (2014). Teaching writing to middle school students: A national survey. *Reading and Writing, 27*(6), 1015–1042.

Graham, S., Harris, K. R., Kiuhara, S. A., & Fishman, E. J. (2017). The relationship among strategic writing behavior, writing motivation, and writing performance with young, developing writers. *The Elementary School Journal, 118*(1), 82–104.

Graham, S., Liu, K., Atken, A., Ng, C., Bartlett, B., & Harris, K. (2018a). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly, 53*(3), 279–304.

Graham, S., Liu, K., Bartlett, B., Ng, C., Harris, K., Aitken, A., Barkel, A., & Talukda, J. (2018b). Reading and writing: A meta-analysis of the impact of reading and reading instruction on writing. *Review of Educational Research, 88*(2), 243–284.

Graham, S., Harris, K. R., Fishman, E., Houston, J., Wijekumar, K., Lei, P. W., & Ray, A. B. (2019). Writing skills, knowledge, motivation, and strategic behavior predict students' persuasive writing performance in the context of robust writing instruction. *The Elementary School Journal, 119*(3), 487–510.

Graham, S., Skar, G. B., & Falk, D. Y. (2020). Teaching writing in the primary grades in Norway: A national survey. *Reading and Writing, 34*, 529–563.

Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., & Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology, 96*(3), 403–423. https://doi.org/10.1037/0022-0663.96.3.403

Gutiérrez, K. D. (2008). Developing a sociocritical literacy in the third space. *Reading Research Quarterly, 43*(2), 148–164.

Harris, K. R., & Graham, S. (1992). Self-regulated strategy development: A part of the writing process. In M. Pressley, K. R. Harris, & J. Guthrie (Eds.), *Promoting academic competence and literacy in school* (pp. 278–312). Academic.

Hayes, J., & Flower, L. (1986). Writing research and the writer. *American Psychologist, 41*, 1106–1113.

Hsiang, T. P., Graham, S., & Yang, Y. M. (2020). Teachers' practices and beliefs about teaching writing: A comprehensive survey of grades 1 to 3 teachers. *Reading and Writing, 33*, 2511–2548.

Ivanic, R. (1998). *Writing and identity: The discoursal construction of identity in academic writing*. John Benjamins.

Limpo, T., Filipe, M., Magalhães, S., Cordeiro, C., Veloso, A., Castro, S. L., & Graham, S. (2020). Development and validation of instruments to measure Portuguese third graders' reasons to write and self-efficacy. *Reading and Writing, 33*, 2173–2204. https://doi.org/10.1007/s11145-020-10039-z1-32

Ng, C., & Graham, S. (2018). Improving literacy engagement: Enablers, challenges and catering for students from disadvantaged backgrounds. *Journal of Research in Reading, 41*(4), 615–624.

Ng, C., & Renshaw, P. (2019). An indigenous Australian student's perezhivanie in reading and the evolvement of reader identities over three years. *Learning, Culture and Social Interaction, 22*, 100310.

Ng, C., Graham, S., Lau, K.-L., Liu, X., & Tang, K.-Y. (2021). Writing motives and writing self-efficacy of Chinese students in Shanghai and Hong Kong: Measurement invariance and multigroup structural equation analyses. *International Journal of Educational Research*. https://doi.org/10.1016/j.ijer.2021.101751

Olinghouse, N., & Wilson, J. (2013). The relationship between vocabulary and writing quality in three genres. *Reading & Writing, 26*, 45–66.

Page-Voth, V., & Graham, S. (1999). Effects of goal setting and strategy use on the writing performance and self-efficacy of students with writing and learning problems. *Journal of Educational Psychology, 91*(2), 230.

Ray, A. B., Graham, S., Houston, J. D., & Harris, K. R. (2016). Teachers use of writing to support students' learning in middle school: A national survey in the United States. *Reading and Writing, 29*(5), 1039–1068.

Renshaw, P., & Tooth, R. (2016). Perezhivanie mediated through narrative place-responsive pedagogy. In A. Surian (Ed.), *Open spaces for interactions and learning diversities* (pp. 13–23). Sense Publishers.

Renshaw, P., & Tooth, R. (2018). *Diverse pedagogies of place: Educating students in and for local and global environments*. Routledge.

Rocha, R. S., Filipe, M., Magalhães, S., Graham, S., & Limpo, T. (2019). Reasons to write in grade 6 and their association with writing quality. *Frontiers in Psychology, 10*, 2157. https://doi.org/10.3389/fpsyg.2019.02157

Roth, W.-M., & Jornet, A. (2017). *Understanding educational psychology a late Vygotskian, Spinozist approach*. Springer International Publishing.

Smagorinsky, P. (2011). Vygotsky's stage theory: The psychology of art and the actor under the direction of perezhivanie. *Mind, Culture, and Activity, 18*(4), 319–341.

Tooth, R., & Renshaw, P. (2009). Reflections on pedagogy and place: A journey into learning for sustainability through environmental narrative and deep attentive reflection. *Australian Journal of Environmental Education, 25*, 95–104.

Tooth, R., Wager, L., & Proellocks, T. (1988). Story, setting and drama – A new look at environmental education. *Australian Journal of Environmental Education, 4*, 31–34.

Ungunmerr-Baumann, M. R. (1988). Dadirri. *Compass Theology Review, 22*, 9–11.

Ungunmerr-Baumann, M. R. (2002). *Dadirri: A reflection by Miriam-Rose Ungunmerr Baumann.* Retrieved from http://nextwave.org.au/wp-content/uploads/Dadirri-Inner-Deep-Listening-M-RUngunmerr-Bauman-Refl.pdf

Veresov, N., & Fleer, M. (2016). Perezhivanie as a theoretical concept for researching young children's development. *Mind, Culture, and Activity, 23*(4), 325–335.

Vygotsky, L. S. (1971). *The psychology of art*. M.I.T. Press.

Vygotsky, L. S. (1994). The problem of environment. In R. Van der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 338–354). Blackwell Press.

# More Motivating than Cherry Pie? The Writer(s) Within Community Model of Writing Through a Motivation Theory Lens


**A. Angelique Aitken**

**Abstract** Until now, the Writer(s) Within Community (WWC) Model of Writing has been characterized as a socio-cognitive model. Dr. Steve Graham presented writing as a socially situated activity within a community of writers that requires an individual's cognitive processes to create a written product. However, the WWC is also a motivational writing model. Although Graham referred to some motivational aspects of the WWC model, his references to motivation were predominately superficial (e.g., motivation, intrinsic motivation) rather than grounded in specific motivation constructs or theories. This chapter expands upon Graham's current theory to view it through a motivational lens of Bandura's social cognitive motivational theory. I describe minor modifications to the current WWC's conceptual framework with writing at the center of the bullseye and added motivational constructs in the writer(s) and writing community components. After delving into how specific motivational constructs interact with the writer(s) and the writing community, as well as the motivational effects the writer(s) and community have upon one another, future directions for analyzing motivational aspects of the WWC are discussed. Tangible examples are provided in the context of Karen and other founding members of a non-profit when they welcome their first intern, Stevie, and other members to their growing writing community.

**Keywords** Writing · Writing theory · Motivation theory · Socio-cognitive theory

The research was supported, in part, by a grant from the U.S. Department of Education, Institute for Education Sciences (R324B160033). The opinions expressed by the authors are not necessarily reflective of the position of or endorsed by the U.S. Department of Education.

A. A. Aitken (✉)
The Pennsylvania State University, State College, PA, USA
e-mail: angelique.aitken@psu.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_4

In over 40 years of scholarship, Dr. Steve Graham has been one of the most influential writing authorities on writing development, effective teaching strategies, and how writing supports learning. From over 20 meta-analysis reviews to dozens of writing intervention studies, to the numerous students and colleagues he has mentored, Graham's influence reaches far. Yet, despite this long legacy, I believe his most influential contribution has been a more recent development, his Writer(s) Within Community (WWC) Model of Writing, which Graham describes as a sociocognitive view of writing.

Previously, writing theorists and viewed writing as either a socially-constructed action or a cognitive process occurring within the writer's mind (Graham, 2018a, b). Further, most writing researchers held either a social or cognitive view without much attention to the alternate point of view (Newell et al., 2011). With the camps divided, neither integrated the benefits of the other's perspective in their research, teaching, or understanding of writing. Instead, each group stayed stuck in a unique version of my-side bias.

With the WWC, Graham integrated both perspectives for a richer and deeper understanding of writing, stating "writing involves an interaction between the context in which it occurs and the mental and physical actions writers and their collaborators are able to enlist and engage" (Graham, 2018a, p. 271). On one hand, the act of composing requires intensive cognitive skills and physical actions occurring, primarily, within the writer (e.g., working memory, executive control, knowledge, beliefs, ideation, translation, emotions, personality traits); however, the assumption that writing occurs in a cognitive vacuum is imprudent. Rather, writing is also inherently a social activity occurring within the context of a writing community (Barton, 1991; Bazerman, 2016; Hull & Schultz, 2001). Like the writer, the community involves multiple layers of influences (e.g., writing tools, goals, community writers and collaborators, collective community history and purposes, and influences from outside the core community) that inform and influence the community's written artifacts. While the cognitive and social aspects may seem to be distinct units and have been traditionally presented that way, several tenets underlying the WWC indicate that two components continuously interact with each other: writer(s) and the writing community.

Despite carefully describing each component and how they interact, Graham did not fully explore the motivational underpinnings of the WWC model because his focus was bridging sociocultural and cognitive writing theories. Implicitly he provided insights into the motivational aspects of the WWC model in that writing is motivated by a variety of social and individual attributes which affect the writing process and resulting products. However, Graham's references to motivation were predominately superficial (e.g., motivation, intrinsic motivation) rather than grounded in specific motivation constructs or theories, with a few exceptions such as where the aspects of the WWC are themselves a motivational construct (e.g., beliefs). In this chapter, I expand upon Graham's work to demonstrate how motivation is deeply incorporated into the WWC model. After a brief overview of the WWC model, I will discuss how components of the model are supported by relevant motivational theories and research using a fictional writing community and it's expanding number of writers who join the community over time.

# 1 Components of the Writer(s) Within Community Model of Writing

Before demonstrating that the WWC is also a motivational view of writing, one needs to understand the social and cognitive items Graham incorporates into his theory. Throughout the rest of the chapter, I will provide contextual examples involving a writing community, the Lark's Advocacy Group, a non-profit that advocates on issues important to people with disabilities. The founding members of this community include Karen, Mike P., Floyd, Skip, Jill, Lamoine, and Arthur. Early on, they brought on an intern, a Converse-wearing guy named Stevie and over time the writing community membership grows.

## *1.1 Writing Community*

The writing community is a multi-layered entity comprised of multiple people who share similar goals, precepts, and assumptions to fulfill their shared writing purposes (Graham, 2018a, b). A writing community can take many different forms, from a classroom to a nonprofit that advocates for people with disabilities to fans updating a Wiki page (e.g., Wookieepedia, a *Star Wars* fan page). In each of these communities, several factors influence how and why the community members produce written products. The seven basic components of a writing community include:

(a) **Goals**—Intent and objective of a particular writing project or written product
(b) **Tools**—Electronic, online forums, print, pen and paper
(c) **Written Product**—Tangible written documents that can include diagrams, pictures, videos, and other artifacts as part of the written product
(d) **Actions**—Defining writing tasks and goals, designating responsibilities for specific actions, managing discussions, researching content, creating the writing environment, composing, editing, and regulating the physical, motivational, social, and emotional aspects of writing
(e) **Community Members**—A group that includes writers, collaborators, *and the intended readers* and brings into consideration the number of members, their roles and responsibilities, as well as the level of exclusivity
(f) **Physical and Social Environments**—The physical and social contexts where writing occurs
(g) **Collective History**—How previous collective experiences shape a community's practices through defining and redefining values, norms, and views expressed in the community's written products
(h) **Purposes of the Writing Community**—Goals, values, and norms related to writing as well as the intended audiences
(i) **External Influences—**Institutional, political, historical, social, cultural, and other socially-derived communities

### *1.2 Writer(s)*

Just as the written artifacts are influenced by the community, they are simultaneously influenced by the unique attributes of the writer(s). Writing is a cognitively demanding task that requires performing multiple co-occurring skills such as self-regulation, executive function, attention, motor functions, memory, communication, and language skills (Hayes, 2012). However, the writer's cognitive architecture and ability to manage these tasks can make writing more or less difficult (Mayer, 2012; Paas & Sweller, 2014). Graham (2018a, b) describes each aspect of the writer through four components:

(a) **Long-Term Memory Resources**—Knowledge about content and writing, writing beliefs, identity, and values
(b) **Control Mechanisms**—Attention, working memory, executive control (intentions, plans, monitoring, reacting)
(c) **Production Processes**—Conceptualization, ideation, translation, transcription, reconceptualization
(d) **Modulators**—Emotions, personality traits, physical state

## 2 Motivation & Writer(s) Within Community Model of Writing

Although the social (community) and cognitive (writers) aspects of this model are presented separately to accommodate the linear nature of this chapter, the four core underlying tenets reflect the core aspect of the WWC model: that the writing community and writer(s) continually interact to shape writing and the written products. The confluence of the writer(s) and the writing community illustrates and resembles multiple layers of motivation, as illustrated in Fig. 1. A written product (or one that is in progress) is at the center of the WWC motivation model. We see a writer in the process of composing a written artifact. Surrounding the writer are the aspects of her writing community ranging outward from proximate to external influences on the written product. The writer and writing community as one interactive unit, together, motivate writing and influence the writing process.

Figure 2 is a more detailed illustration of the writer with her cognitive and motivational processes interacting as she writes. The major cognitive processes are illustrated in a similar manner and layout to Graham's (2018a) representation with a few modifications. First, the computer represents the act of writing, at the center of the model in both figures. Second, a woman is depicted to represent the writer(s). The theory posits that multiple writers could be co-writing (e.g., Google Docs), but for simplicity, one is depicted here. Surrounding her head are the three primary types of cognitive processes, presented in a similar layout to Graham; however, not every aspect of long-term memory is depicted here in its entirety due to space constraints

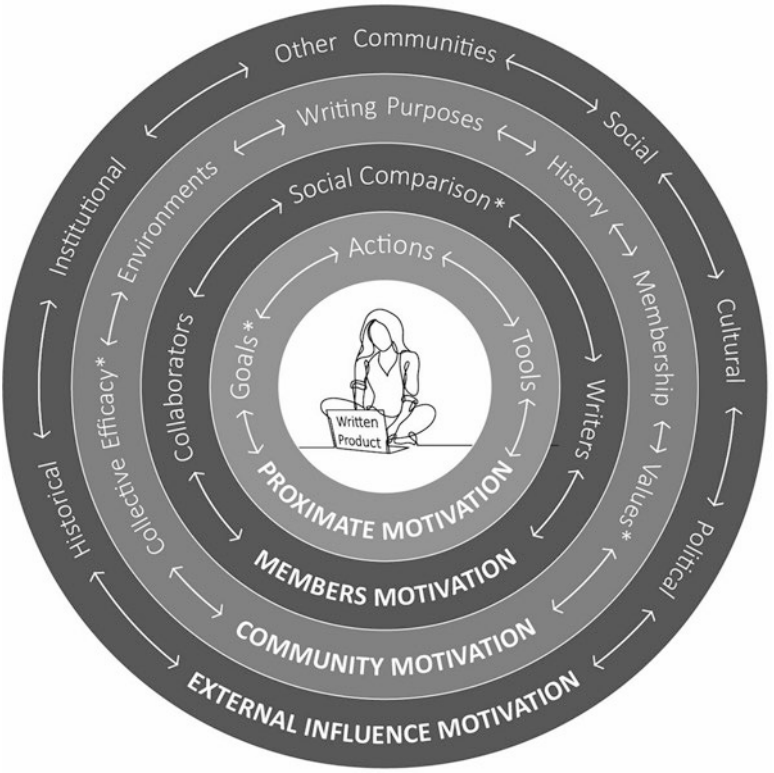



**Fig. 1** Writer(s) within community motivation conceptual framework

(e.g., "knowledge" in the figure includes content knowledge, reading and writing knowledge etc.,). Third, each item that Graham grouped as "modulators" is portrayed across the writer's torso because I envision them more related to the "heart" (i.e., personality, emotions) and body (physical state) than as cognitive functions. Finally, items noted by an asterisk are the motivational constructs addressed in this paper.

Similarly, this depiction of the writing community aligns with Graham's, with a few modifications. First, the center of the community is writing (written product) and the writer(s). While the circular format was used to align with Graham's model, the shading is intended to represent a target with the written product as the bullseye or focused goal. The target represents each layer interacting with the other layers and the writer(s), while the bidirectional arrows indicate same-level interaction. The proximate motivation level includes items that were at the center of the community in the original the WWC model. One exception is "written product" because this model has a computer at the center, which can represent, not only current writing but previously written documents from the writer(s) and the writing community accessed by the computer. Further, the community's written documents could be considered tools in that they guide the community (e.g., community guidelines and

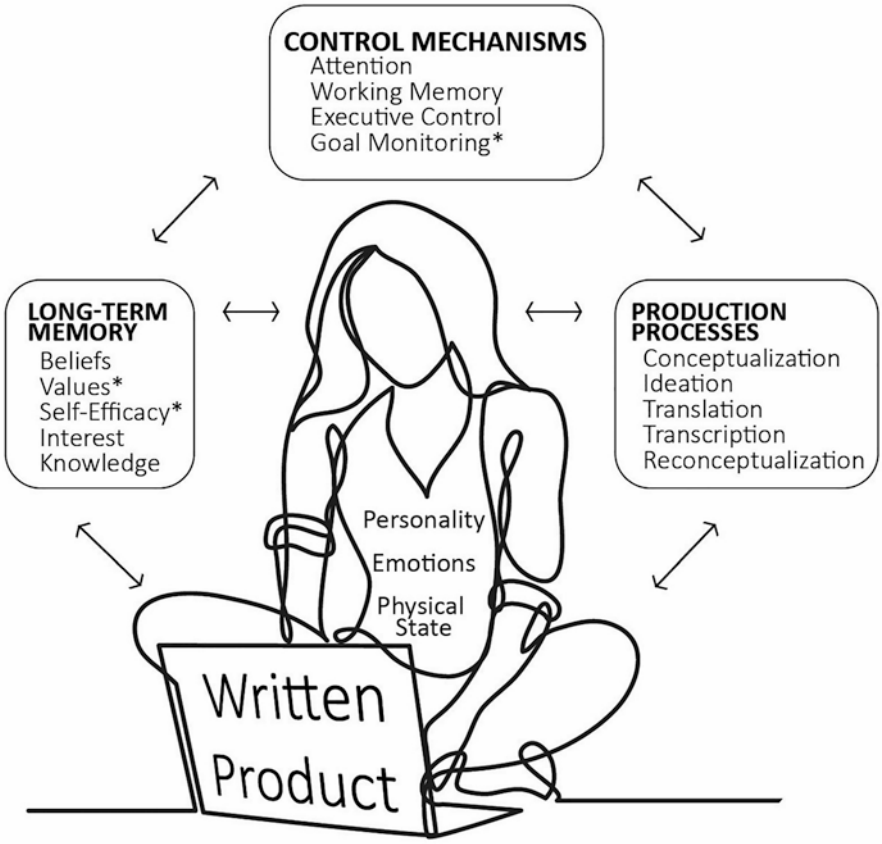



**Fig. 2** Writer(s) motivational and cognitive processes

bylaws) or as writing models. As with the writer (Fig. 2), asterisks represent motivational constructs addressed in this paper. Finally, each layer is named, in part, to account for other potential motivational aspects not identified in this paper.

In summary, the figures illustrate that cognitive processes and the aspects of the writing community influence writing motivation and, ultimately, the written product, represented by the computer. Each of the cognitive processes affects and is affected by the motivational aspect of writing, just as the aspects of the writing community similarly interact with motivation. While the WWC fits with aspects of multiple motivational theories, this representation of the WWC most clearly aligns with Bandura's social cognitive theory (1977b, 2001).

## 3 Social Cognitive Theory

Given that the WWC is a socio-cognitive model of writing, the motivational model most aligned with the WWC is social cognitive theory (Schunk & Usher, 2012). Components of social cognitive theory which are found in the WWC include triadic reciprocity, enactive and vicarious learning, as well as the following motivational constructs: (a) social comparison, (b) goals and goal setting, (c) values, (d) outcome expectations, and (e) self-efficacy (Schunk & Usher, 2012).

At the core of social cognitive theory, Bandura describes a triadic reciprocality where behavior, personal, and social/environmental aspects interact with one another to explain human motivation and behaviors. For example, environmental aspects (e.g., friends, pop culture, political influences) reciprocally influence personal aspects (e.g., cognition, identity, skills, prior knowledge) that motivate a particular behavior. Both personal and environmental aspects influence how, why, and how well an individual (or a group of individuals) engages in a certain behavior, such as writing (Bandura, 1977b, 2001). Similarly, the behavior (writing) or the result of the behavior (written artifact) impacts personal and social/environmental aspects. To illustrate this relationship, many descriptions of social cognitive theory (e.g., Schunk & DiBenedetto, 2020; Schunk & Usher, 2012) include some form of Fig. 3 to illustrate the triadic reciprocality interactions. Figures 3 and 4 illustrate the parallel between social cognitive theory (Fig. 3) and the WWC (Fig. 4). By replacing "individual" with "writer(s)," "environment/social" with "writing community," and "behavior" with "writing," a modified version (Fig. 4) illustrates the parallel alignment between the WWC and the social cognitive motivational theory.

The underlying tenets of the WWC describe how the writing community and writer(s) influence one another (Graham, 2018a). In Tenet 2, Graham states that "writing is simultaneously shaped by the capacity of the writing community and the capacities possessed by members of the community" (2018a, p. 272). This

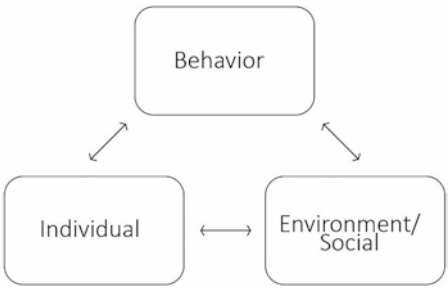



**Fig. 3** Triadic reciprocality, social cognitive theory

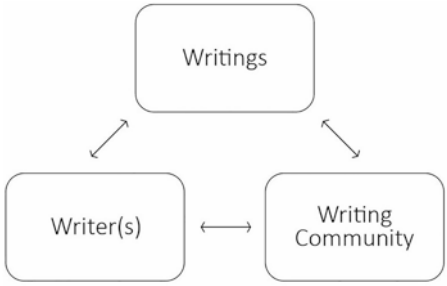



**Fig. 4** Triadic reciprocality, writer(s) within community theory

tenet focuses upon the abundance or lack of capabilities of the different aspects of the writing community and the writer(s). He briefly mentions that this can include an evolving motivational context where he compares a long-standing writing community to a fledgling community. To extend this further and incorporate Bandura's idea of triadic reciprocality (1977b, 2001), the writing community can motivate the writer(s) and vice versa, which influences the written product.

Consider the writing community that takes the form of a disability advocacy non-profit with multiple writing purposes such as engaging with donors, informing the public of upcoming ballot measures, instigating letter-writing campaigns, and lobbying elected leaders with whitepapers to pass laws and fund essential programs. Karen, a lead member of this non-profit, may choose to ask intern Stevie to handwrite thank you notes (i.e., tools are pen and paper) to their donors with the goal of expressing gratitude by way of the distinctiveness of a handwritten note. On one hand, this task may be demotivating because transcription is hard for Stevie; it takes him a long time. Plus, let's face it, Stevie doesn't have the best handwriting. His thank-you notes may be written sloppily (behavior) and offend donors (member as a reader), which lowers their motivation to donate and remain as part of the community. However, for Amber, another member of the writing community who enjoys calligraphy, this writing task of the community could be intrinsically motivating. Further, Karen praised her beautifully written notes which increased her motivation to create this type of written product.

In another scenario, the writing community, collectively, may be more motivated to take on writing tasks to oppose potential legislation "requiring all third-grade students read at grade level or else they shall be retained" because of the expertise of specific community members. Ashley and April, previous special education teachers and lawyers, with their specialized special education knowledge and writing practices, can help expand the reach of this non-profit writing community. Similarly, Ashley and April are excited to write about an issue they believe is particularly important and that they are extremely qualified to tackle.

## 4 Enactive and Vicarious Learning

Another core aspect of the social cognitive theory framework that aligns with the WWC is enactive and vicarious learning. The writing community can provide numerous sources of information and motivation for the members (Schunk & Usher, 2012). Enactive learning occurs from actions taken by the individual, whereas vicarious learning occurs by learning without explicit actions by the learners (Bandura, 1986) or writers. The writing intervention literature supports both types of learning such as strategy-based writing intervention where students engage in the writing process (e.g., Lane et al., 2011; Ray et al., 2019) and observational learning where participants observe a student model while writing (e.g., Braaksma et al., 2004). In the context of the WWC, both types of learning can influence writing motivation. In the legislation example, April may use enactive learning to write to a new audience (congressional members instead of a court), and if persuades congressional members, this can increase her motivation.

With vicarious learning, Zoi who has similar writing abilities and training could learn through observing April's writing technique may feel similarly motivated and confident to engage in future legislative writing tasks (Schunk, 1987). Another form of vicarious learning could include using previously written documents from other community members as a model to help novice writers improve their writing. Having writing models from experienced writers could increase writing motivation for newer members, such as Bruce and Sandy, because it reduces task difficulty, and they are more likely to be successful when using a previously successful model.

Using a written artifact is not the only type of modeling that occurs in a writing community; observing what more experienced writer(s) do can help guide a novice. For example, Karen may coach Stevie (the intern) by telling him how she prepares her writing space, identifies her writing goals, and interacts with collaborators to learn how he can become a more effective writer. As Stevie becomes a stronger writer, his writing self-efficacy and motivation should improve. Whether learning enactively or vicariously, the writing community strengthens as the members grow in their writing skills.

## 5 Motivational Constructs

The social cognitive theory includes multiple motivational constructs that authentically align with aspects of the WWC, including (a) social comparison, (b) goals and goal setting, (c) values, (d) outcome expectations, and (e) self-efficacy (Schunk & Usher, 2012).

## 5.1 Social Comparison

Given the interaction between the writer (person) and writing community (environment), developing as a writer occurs through social comparison, where an individual compares themselves to others (Wheeler & Suls, 2005). Social comparison of a similarly competent and situated individual who succeeds at a task can increase the observer's motivation that they too will be successful (Schunk, 1987). In the WWC framework, a writer compares their writing performance (e.g., compositions) and other writing skills (e.g., research, editing, self-regulation, collaborations, number of publications on a CV) to other writers in the writing community. Take a cohort of doctoral students; if Gustav has a manuscript accepted to a journal, other cohort members, Vera and Bob, experience increased self-efficacy, motivation, and likeliness to follow in their cohort-mate's footsteps. However, it is important to note that while social comparison can motivate adults, results for children are more variable (e.g., Schunk, 1983; Guay et al., 1999).

## 5.2 Values

One area the WWC explores is the effect of values on the writer(s), the community, and the interaction between the two. Values, as a motivational construct, refer to the perceived importance of a belief, activity, or learning activity. Individuals will be motivated to act in a way that supports their values (Wigfield & Eccles, 2002; Wigfield et al., 2004). Graham (2018a) went beyond discussing the impact that values and beliefs have on both the individual writers and the community; he also integrated the motivating effect that the interaction between the values of the writer(s) and community value(s) have on one another. Being a part of a writing community can shape the values of the individuals. For example, when Kevin joined our imaginary non-profit, he did not have a particular affinity for disability rights or writing. Although he had always been a "good writer," what he valued most when he accepted the job was the ability to pay his rent. Over the years, however, through his involvement with the writing community, his values widened to include those of the community and other community members. Similarly, adding April and Ashley to the writing community increased the importance placed on special education-specific issues. Both examples illustrate that knowledge has a significant influence on values (Graham, 2018a, b). As Kevin's knowledge about disability rights advocacy increased, his personal values motivated him to write on behalf of the non-profit. Similarly, when Ashley and April joined the community, their knowledge expanded the knowledge of all the community members, so they could understand and value an area of new but related issues.

Values can impact specific aspects of the writer(s) and the writing community (Graham, 2018a). For one, values can include the beliefs of the importance of the

act of writing as well as the message of a written artifact. For example, Michael, the non-profit's silver-tongued "front man," shares the values of the non-profit, but writing is not his thing. Instead, he would rather schmooze in person to achieve the community's goals. However, if he must write, he will write, begrudgingly. On the other hand, Amy, a freelance writer, loves writing, whether it be disability advocacy, issues related to homelessness, or writing about NASCAR. She values writing over content. Values associated with content or issues (Michael) versus the act of writing (Amy) likely motivate Michael and Amy to participate in the writing community in different ways. Similarly, participation in and commitment to a writing community can be informed by members' values, such as alignment of overarching values and how a writer feels about their role in the community. If Gary does not value his role as an issues manager because he wants more responsibility, he may leave the community. Let's say over the next year, Mike P., Arthur, Lamoine, Floyd, Paul, Colin, and Julia all retire or leave the non-profit. The individual values of those who take their places, interacting with those who remain, can change the community's values, goals, and purposes for writing. With Karen, and the others including Stevie (what a loyal guy!) remaining, the collective community's motivation may shift to focus more upon special education-specific issues rather than disability advocacy more broadly.

## *5.3 Goals and Goal Setting*

The WWC addresses writing goals, intentions, and purposes throughout both the *writer(s)* and the *writing community* components. However, Graham only touches on the motivational aspects of goals and goal setting. Instead, he focuses on how goals function to direct the individual or the community and goals as part of the cognitive processes. He does discuss how the overarching purposes of a writing community interact with goals, motivation, and other aspects of the WWC. Given the numerous times Graham includes the concept of goals in his theoretical papers (2018a, b), it is important to consider how this traditional motivation construct applies to the WWC.

According to the social cognitive theory, goal setting provides direction and motivation to achieve a desired outcome. Progress toward a goal motivates an individual to continue to take steps towards achieving the goal (Locke & Latham, 2002). Similarly, goal achievement further motivates future goal-setting and goal-directed behavior (Bandura 1986; Schunk, 1990). Environmental feedback on goal progress also influences an individual's self-efficacy, which can impact motivation and performance (Locke & Latham, 2019; Schunk & Swartz, 1993) Both personal traits (e.g., prior performance, self-efficacy, actual ability, values) and environmental/social attributes (e.g., community norms and goals, feedback from others, actions by others) can influence goal commitment and, therefore, the required actions needed to attain the goal (Locke & Latham, 1990, 2019). Furthermore, previous

research has shown that goal-setting for writing tasks improves writing performance and motivation (e.g., Page-Voth & Graham, 1999; Graham et al., 1995; Schunk, 2003)

Based on these basic tenets of social cognitive theory, goals are also a motivating factor in the WWC. The writing community is guided by a set of writing purposes (goals), which are valued by the community and its members. More specific writing goals, which may be set by the writer(s), are then inextricably linked by the community. The community structure provides direction and motivation for goal-directed behavior. For instance, writer(s) in the disability advocacy non-profit would likely be less committed and motivated to set writing goals associated with writing *Star Wars* fan fiction because that topic is outside the community's writing purpose, norms, and goals.

Furthermore, community members' feedback and actions can influence the writer's writing motivation. If leader Karen provides feedback on intern Stevie's writing goals in a way that is knowledgeable, specific, and authentically supportive, this can enhance his writing motivation. Secondly, community member's actions, such as multiple members' offering to collaborate on Alyson's proposed writing goal, can similarly elevate her motivation. Conversely, individuals can affect the purposes of the writing community. As new members join the community, their individual writing goals can shift, expand, and influence the community's writing purposes (e.g., April and Ashley's expertise expanding the non-profit's writing purpose to include special education-related issues).

One other aspect of goals worth mentioning in relation to the WWC is the effort required to achieve a goal. Although many persist more with more difficult goals, there is a tipping point where tasks that require too much effort have a demotivating effect (Schunk & Usher, 2012). While overly lofty community goals could demoralize the writing community, in relation to the WWC, I see it exemplified most prominently in the cognitive aspect of Graham's model. As he illustrates, writing is a complex cognitive process that requires a writer to juggle control mechanisms, long-term memory resources, production processes, and unique modulators (i.e., emotions, personality traits, physical state). With all the content and communication skills (e.g., oral language, specialized writing knowledge, knowledge of the writing community) required to compose, the cognitive architecture has to process, integrate, and manage a lot of information in order to write. If the goal-related tasks are too difficult, the control mechanisms will run less efficiently.

For instance, a less proficient writer's working memory would try to hold too much information that more proficient writers can access in their long-term memory (MacArthur & Graham, 2016; Torrance & Galbraith, 2006). This can create cognitive overload resulting in difficulties with focus, which in turn inhibits the production process (e.g., ideation is more difficult), lowers motivation, and results in less effective writing (Bruning & Horn, 2010; Flower & Hayes, 1980; Graham, 2018a). The difficulty associated with writing may have such a demotivating effect that some members may avoid writing, passively participate in their writing community, or leave the community altogether.

## *5.4 Self-Efficacy and Collective Efficacy*

Self-efficacy is the level of perceived ability to learn or take action in order to achieve a goal. It holds a central role in social cognitive theory as it is critical for motivation and behavioral change (Bandura, 1977a; Schunk & Usher, 2012). Self-efficacy evolves through four fundamental means: (a) enactive and performative experiences, (b) vicarious experiences, (c) social persuasion, and (d) psychological processes (Schunk & Usher, 2012). With writing specifically, research has shown that writing self-efficacy correlates with writing performance and outcomes (e.g., Bruning & Kauffman, 2016; De Smedt et al., 2016; Graham et al., 2017; Pajares & Valiente, 2006).

Graham addresses self-efficacy, but not as thoroughly as if he interpreted the theory through a motivational lens. Graham acknowledges self-efficacy as a belief in the writer's long-term memory (2018a, b) and that it impacts one's emotions (as a modulator; Graham, 2018a). But when we add the layer of socio-cognitive motivational theory, the influence of self-efficacy is much more pronounced. Each of the four sources of self-efficacy development is influenced by the writer(s), the writing community, or both.

Enactive and performative experiences provide the most reliable type of feedback to influence self-efficacy because it is the actions of the individual that inform their ability. Therefore, when Shar, a writer at the non-profit, has her editorial on the pending reading legislation picked up by a national news source, the cognitive interpretation of this success heightens her writing self-efficacy (Schunk & Usher, 2012). Shar's writing success not only can influence her self-efficacy but also that of other similarly situated writers through vicarious experience. As with the doctoral student example, Matt, a newer member of the writing community, may increase his self-efficacy based on Shar's model and other successful models in the community. Similarly, input from others in the form of social persuasion can influence a writer's self-efficacy. For instance, Shar encouraging Matt through positive, authentic feedback and persuasive messages about his can writing can boost Matt's writing self-efficacy. However, negative messages could lower Matt's self-efficacy, and if Matt does not feel like he can relate to Shar, social persuasion could have no impact on his self-efficacy.

The fourth source of self-efficacy, psychological processes, involves the interpretation of emotional or psychological aspects that can influence self-efficacy, such as feeling nervous or joyful about an upcoming experience. In the writing context, this could occur if Tanya experiences writer's block or writing anxiety, which could lower efficacious beliefs. Alternatively, when Rui writes 2000 words with ease because they feel excited and are in flow (Csikszentmihalyi et al., 2014) these psychological processes can enhance self-efficacy.

Self-efficacy is not influenced by the outcome of a situation but by how the individual interprets the situation. Therefore, not only do these psychological and emotional processes originate within the individual writer, but inputs from the writing community can influence how the individual interprets inputs to affect their

self-efficacy. Suppose the non-profit writing community has a 30 plus-year history of success in fending off cuts to the Americans with Disabilities Act. Depending on the writer(s), this could evoke anxiety or confidence. Ralph feels anxious their writing will "break the streak," while Kay is boosted by the long-standing reputation of the community. So, the individual's interpretation of the same community input could impact Ralph and Kay's writing self-efficacy differently. Similarly, vicarious experiences can be interpreted negatively or positively, affecting self-efficacy accordingly. For example, Fien, a visiting scholar from Belgium, may feel less confident after witnessing Hilde (also from the EU) fail at a writing assignment. Verbal recognition from a respected colleague or mentor could impact self-efficacy, such as if Karen verbally encourages Debra telling her she is ready to tackle a more challenging writing project. Vicarious experiences can be interpreted negatively or positively, which affects self-efficacy accordingly. So while self-efficacy is related to the cognitive interpretation of success, input from the community can influence how, why, and if the writer experiences an increase, decrease, or no effect on writing self-efficacy.

## Collective Efficacy

While members of the writing community each have their individual self-efficacy, the writing community has a collective efficacy. Collective efficacy refers to the efficacious beliefs members have for the community's ability as a whole rather than the sum of each member's level of self-efficacy. The development of collective efficacy mirrors that of an individual's self-efficacy, where enactive experiences, vicarious experiences, social persuasion, and psychological processes develop collective efficacy. Further, collective efficacy grows over time as community members learn more about one another, how to effectively work together toward the community's goals, and create appropriate social and physical environments conducive to goal achievement (Goddard et al., 2004). The fact that our non-profit now has over 30 years of success with most of the core founding members still a part of the writing community has a strong collective efficacy because members recognize the non-profit's track record in writing and advocating on behalf of the disability community.

Even as new members, like Ashley and April, join the writing community, they recognize the strong collective efficacy and are motivated to contribute to it. The non-profit's strong collective efficacy impacts the individual's motivation as well as the collective community. Our community members may feel a high sense of job satisfaction, less stress, and more motivation to participate in the community because of the collective efficacy (e.g., Caprara et al., 2003; Petitta & Borgogni, 2011). As a writing community, collective efficacy may motivate group action. For instance, the writing community may feel more emboldened to take on new, but related, writing tasks such as special education issues spearheaded by Ashley.

## 6 Other Motivational Theories and Writer(s) Within Community

This chapter scratches the surface on how the Writer(s) Within Community Writing Model represents not only a socio-cognitive but a motivational theory of writing as well. I argued that the WWC aligns with social cognitive theory (Bandura, 2001; Schunk & Usher, 2012), paper viewed the WWC as a socio-cognitive theory. However, future research could further might view the WWC through other motivational lenses. The WWC and this paper touch on a number of motivational constructs that could be further explained and connected to the WWC. First, much more could be explored about how goals and goal orientations, beliefs, interests, and values relate to both the writer(s) and community. Second, a closer examination of the WWC's four basic tenets and the motivational underpinnings related to (a) how writer(s) and the community are simultaneously shaped (Tenet 1), (b) the capacities or limitation of capacities of writer(s) and community (Tenet 2), (c) variability within the community and individual differences (Tenet 3), and (d) changes within the community and among its members (Tenet 4). Third, a closer examination of other prominent motivational theories and how they align with the WWC could shed additional light on other aspects of the WWC, further grounding it as a motivational writing theory. Briefly, I suggest three such theories.

### *6.1 Self-Determination Theory*

Self-determination theory holds that autonomy, competence, and connectedness impact intrinsic motivation and are required for personal fulfillment (Deci & Ryan, 2000). Future research could compare autonomy and competence to the writer's cognitive needs and connectedness to a writing community as the social need. Furthermore, self-determination theory has multiple mini-theories (Ryan & Deci, 2017; Vansteenkiste et al., 2010) that could be examined to further connect the WWC to self-determination theory. For example, cognitive evaluation theory (Vansteenkiste et al., 2010) posits that autonomy and perceived competence work together to support intrinsic motivation. These two requirements could be analogized to cognitive aspects of the WWC, and perceived competence could have been analyzed as I discussed self-efficacy.

### *6.2 Self-Regulation Theory*

While later socio-cognitive motivation writings describe self-regulation as an aspect of social cognitive theory (e.g., Bandura, 1991), self-regulation has a separate theoretical domain (e.g., Zimmerman et al., 1994) and has been specifically applied to

writing (e.g., Zimmerman & Risemberg, 1997). Therefore, self-regulation deserves a more in-depth analysis of how it connects to the WWC than can be afforded here. Self-regulation is the combination of processes that writers employ to maintain focus, take goal-directed actions, self-monitor their progress, and adjust their actions so that they may reach their goals (Zimmerman & Risemberg, 1997). Admittedly, Graham (2018a) does not explicitly state self-regulation, but he cites seminal sources on self-regulation when he describes the cognitive control mechanisms an individual implements in order to write (i.e., intentions, plans, monitor, react) are part of self-regulation. This connection could be more firmly made. Further, future research could describe the effects the community could have on self-regulation while writing (e.g., a rambunctious, off-task environment versus a quiet, focused one).

### *6.3 Maslow's Hierarchy of Needs*

Just as Maslow identified certain basic needs that must be met before an individual will feel motivated to take action to meet other needs, certain aspects of the WWC must be met before there is motivation to write. For instance, if Natalie skipped breakfast, she will be less motivated to write a whitepaper on the importance of disposable straws for some members of the disability community until she has filled her belly. Also, writers who have difficulty with handwriting or spelling will have difficulty with more challenging tasks (Graham et al., 2011; Scardamalia & Bereiter, 1986). Similarly, if the basic tools (e.g., computer, pens) were unavailable, the social environment unwelcoming, the purposes for writing unclear, and/or disinterested mentors led a writing community at Arizona State University, Angelique may never have felt motivated to write.

## 7 Conclusion

Stevie, who started out as an intern at the beginning of this paper, has moved up in the writing community to VP of Research and Writing, second in command only to Karen. He still wears his Converse All-Stars to even the most elegant fundraising events, and he can't help himself from stealing the intern's pens, but through his participation in this writing community, he grew as a writer and became more committed to advocating for Lark's Non-Profit: An Advocacy Group for Persons with Disabilities.

Over the years, he experienced changes in his values, goals, writing self-efficacy, and social comparisons, not only for himself but with the collective writing community. He now mentors new members of the writing community along with Karen. They have provided enactive and vicarious learning opportunities for dozens of novice writers who eventually leave the non-profit to create writing communities of

their own. And we thank you for all you have done and will continue to do for all your writer(s) and writing communities. May you never have to handwrite a thank you card again. Instead, may you find and write about 42.

**Acknowledgements** Steve, thank-you for welcoming me into your writing community, mentoring me, and inspiring me to make a difference in the world through writing.

You deserve the last slice of cherry pie.

## References

Bandura, A. (1977a). Self- efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215. https://doi.org/10.1016/0146-6402(78)90002-4

Bandura, A. (1977b). *Social learning theory*. Prentice Hall.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248–287. https://doi.org/10.1016/0749-5978(91)90022-L

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Barton, D. (1991). The social nature of writing. In D. Barton & R. Ivanič (Eds.), *Writing in the community* (pp. 1–13). Sage.

Bazerman, C. (2016). What do sociocultural studies of writing tell us about learning to write? In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 11–23).

Braaksma, M. A., Rijlaarsdam, G., Van den Bergh, H., & van Hout-Wolters, B. H. M. (2004). Observational learning and its effects on the orchestration of writing processes. *Cognition and Instruction, 22*(1), 1–36. https://doi.org/10.1207/s1532690Xci2201_1

Bruning, R., & Horn, C. (2010). Developing motivation to write. *Educational Psychologist, 35*(1), 25–37. https://doi.org/10.1207/S15326985EP3501_4

Bruning, R. H., & Kauffman, D. F. (2016). Self-efficacy beliefs and motivation in writing development. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 160–173).

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Steca, P. (2003). Efficacy beliefs as determinants of teachers' job satisfaction. *Journal of Educational Psychology, 95*(4), 821–832. https://doi.org/10.1037/0022-0663.95.4.821

Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J. (2014). Flow. In *Flow and the foundations of positive psychology* (pp. 227–236). Springer. https://doi.org/10.1007/978-94-017-9088-8_15

De Smedt, F., Van Keer, H., & Merchie, E. (2016). Student, teacher and class-level correlates of Flemish late elementary school children's writing performance. *Reading and Writing, 29*(5), 833–868. https://doi.org/10.1007/s11145-015-9590-z

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuit: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*, 277–268. https://doi.org/10.1207/S15327965PLI1104_01

Flower, L., & Hayes, J. R. (1980). The cognition of discovery: Defining a rhetorical problem. *College Composition and Communication, 31*(1), 21–32. https://doi.org/10.2307/356630

Goddard, R. D., Hoy, W. K., & Hoy, A. W. (2004). Collective efficacy beliefs: Theoretical developments, empirical evidence, and future directions. *Educational Researcher, 33*(3), 3–13. https://doi.org/10.3102/0013189X033003003

Graham, S. (2018a). A revised writer (s)-within-community model of writing. *Educational Psychologist, 53*(4), 258–279. https://doi.org/10.1080/00461520.2018.1481406

Graham, S. (2018b). A writer(s) within community model of writing. In C. Bazerman, A. N. Applebee, V. Berninger, D. Brandt, S. Graham, J. V. Jeffery, P. K. Matsuda, S. Murphy, D. W. Rowe, M. Schleppegrell, & K. C. Wilcox (Eds.), *The lifespan development of writing*. National Council of Teachers of English.

Graham, S., MacArthur, C., & Schwartz, S. (1995). Effects of goal setting and procedural facilitation on the revising behavior and writing performance of students with writing and learning problems. *Journal of Educational Psychology, 87*(2), 230–240. https://doi.org/10.1037/0022-0663.87.2.230

Graham, S., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children, 44*(4), 1–12. https://doi.org/10.17161/foec.v44i4.6687

Graham, S., Harris, K. R., Kiuhara, S. A., & Fishman, E. J. (2017). The relationship among strategic writing behavior, writing motivation, and writing performance with young, developing writers. *The Elementary School Journal, 118*(1), 82–104. https://doi.org/10.1086/693009

Guay, F., Boivin, M., & Hodges, E. V. (1999). Social comparison processes and academic achievement: The dependence of the development of self-evaluations on friends' performance. *Journal of Educational Psychology, 91*(3), 564. https://doi.org/10.1037/0022-0663.91.3.564

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29*(3), 369–388. https://doi.org/10.1177/0741088312451260

Hull, G., & Schultz, K. (2001). Literacy and learning out of school: A review of theory and research. *Review of Educational Research, 71*(4), 575–611. https://doi.org/10.3102/00346543071004575

Lane, K. L., Harris, K., Graham, S., Driscoll, S., Sandmel, K., Morphy, P., House, H., & Schatschneider, C. (2011). Self-regulated strategy development at tier 2 for second-grade students with writing and behavioral difficulties: A randomized controlled trial. *Journal of Research on Educational Effectiveness, 4*(4), 322–353. https://doi.org/10.1080/19345747.2011.558987

Locke, E. A., & Latham, G. P. (1990). Work motivation and satisfaction: Light at the end of the tunnel. *Psychological Science, 1*(4), 240–246. https://doi.org/10.1111/j.1467-9280.1990.tb00207.x

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist, 57*(9), 705–717. https://doi.org/10.1037//0003-066X.57.9.705

Locke, E. A., & Latham, G. P. (2019). The development of goal setting theory: A half century retrospective. *Motivation Science, 5*(2), 93–105. https://doi.org/10.1037/mot0000127

MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 160–173).

Mayer, R. (2012). Information processing. In K. Harris & S. Graham (Eds.), *APA educational psychology handbook* (Vol. 1, pp. 85–100). APA.

Newell, G. E., Beach, R., Smith, J., & VanDerHeide, J. (2011). Teaching and learning argumentative reading and writing: A review of research. *Reading Research Quarterly, 46*(3), 273–304. https://doi.org/10.1598/RRQ.46.3.4

Paas, F., & Sweller, J. (2014). Implications of cognitive load theory for multimedia learning. In R. Mayer & R. Mayer (Eds.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 27–42). Cambridge University Press.

Page-Voth, V., & Graham, S. (1999). Effects of goal setting and strategy use on the writing performance and self-efficacy of students with writing and learning problems. *Journal of Educational Psychology, 91*(2), 230–240. https://doi.org/10.1037/0022-0663.91.2.230

Pajares, F., & Valiente, G. (2006). The social nature of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 158–170). Sage.

Petitta, L., & Borgogni, L. (2011). Differential correlates of group and organizational collective efficacy. *European Psychologist, 16*, 187–197. https://doi.org/10.1027/1016-9040/a000035

Ray, A. B., Graham, S., & Liu, X. (2019). Effects of SRSD college entrance essay exam instruction for high school students with disabilities or at-risk for writing difficulties. *Reading and Writing, 32*(6), 1507–1529. https://doi.org/10.1007/s11145-018-9900-3

Ryan, R. M., & Deci, E. L. (2017). Self-determination theory: Basic psychological needs in motivation, development, and wellness. .

Scardamalia, M., & Bereiter, C. (1986). Written composition. In M. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 778–803). Macmillan.

Schunk, D. H. (1983). Developing children's self-efficacy and skills: The roles of social comparative information and goal setting. *Contemporary Educational Psychology, 8*(1), 76–86. https://doi.org/10.1016/0361-476X(83)90036-X

Schunk, D. H. (1987). Peer models and children's behavioral change. *Review of Educational Research, 57*(2), 149–174. https://doi.org/10.3102/00346543057002149

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*(1), 71–86. https://doi.org/10.1207/s15326985ep2501_6

Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading & Writing Quarterly, 19*(2), 159–172. https://doi.org/10.1080/10573560308219

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology, 60*, 1–10. https://doi.org/10.1016/j.cedpsych.2019.101832

Schunk, D. H., & Swartz, C. W. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology, 18*(3), 337–354. https://doi.org/10.1006/ceps.1993.1024

Schunk, D. H., & Usher, E. L. (2012). Social cognitive theory and motivation. In R. Ryan (Ed.), *Oxford handbook of human motivation* (pp. 249–279). Oxford University Press.

Torrance, M., & Galbraith, D. (2006). The processing demands of writing. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 67–80). Guilford Press.

Vansteenkiste, M., Niemiec, C. P., & Soenens, B. (2010). The development of the five mini-theories of self-determination theory: An historical overview, emerging trends, and future directions. In *The decade ahead: Theoretical perspectives on motivation and achievement*. Emerald Group Publishing Limited.

Wheeler, L., & Suls, J. (2005). Social comparison and self-evaluations of competence. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 566–578). Guilford Press.

Wigfield, A., & Eccles, J. S. (2002). The development of competence beliefs, expectancies for success, and achievement values from childhood through adolescence. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 91–12). Academic. https://doi.org/10.1016/B978-012750053-9/50006-1

Wigfield, A., Tonks, S., & Eccles, J. S. (2004). Expectancy value theory in cross-cultural perspective. In D. M. McInernery & S. Van Etten (Eds.), *Big theories revisited* (pp. 165–198). Information Age Publishing.

Zimmerman, B. J., & Risemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22*, 73–101. https://doi.org/10.1006/ceps.1997.0919

Zimmerman, B. J., Greenberg, D., & Weinstein, C. E. (1994). Self-regulating academic time: A strategy approach. In B. H. Schunk & B. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational implications* (pp. 181–191). Erlbaum.

# Part III
# Writing Instructions at School Levels

# Academic Vocabulary in First-Grade Children's Compositions: An Exploration

**Jill Fitzgerald, Jackie Eunjung Relyea, Jeff Elmore, and James S. Kim**

**Abstract** The purpose of the study was to explore the extent of first-grade children's academic vocabulary usage in writing. We additionally investigated the extent to which children who used more academic words tended to write compositions that reflected more phonologically and orthographically unique words and/or words with greater semantic challenge. The sample consisted of 191 first-grade control group children involved in a prior study. In the spring of first grade, after hearing a passage about a science topic, students were prompted to write an argumentative composition. Academic words were counted in compositions. Four predictor variables represented characteristics of words in the compositions—(a) phonological uniqueness (how similar a word sounds in comparison to other words); (b) orthographic uniqueness (how similar a word's spelling is compared to other words); (c) semantic challenge (the age at which a word is commonly known) and (d) semantic concreteness (how tangible or imageable a word's meaning is). The main conclusions were: Slightly more than half of the children included one or two academic words in their compositions, and greater inclusion of academic words in a composition was weakly associated with composition words that were more phonologically unique and more semantically challenging.

**Keywords** Vocabulary · Academic vocabulary · Early grades · Writing · Composition

J. Fitzgerald (✉)
The University of North Carolina at Chapel Hill, Chapel Hill, NC, USA
e-mail: jfitzger@email.unc.edu

J. E. Relyea
North Carolina State University, Raleigh, NC, USA
e-mail: jrelyea@ncsu.edu

J. Elmore
MetaMetrics, Durham, NC, USA
e-mail: jelmore@gmail.com

J. S. Kim
Harvard University, Cambridge, MA, USA
e-mail: james_kim@gse.harvard.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_5

## 1 Introduction and Rationale

The overarching purpose of the present study was to explore the extent of first-grade children's academic vocabulary usage in writing. We also investigated the extent to which children who used more academic words tended to write compositions that reflected more phonologically and orthographically unique words and/or words with greater semantic challenge.

Academic vocabulary is defined as words that appear more frequently in academic disciplinary (content area) discussion and texts than in informal communications (cf. Nagy & Townsend, 2012). Academic vocabulary consists of two types of words—domain-specific and general academic words. Domain-specific academic vocabulary is defined as a lexicon of relatively low-frequency words that are unique to a specific disciplinary area (e.g., Nagy & Townsend, 2012). General academic vocabulary is defined as words: (a) that meet the definition of academic—that is, words that appear more frequently in academic disciplinary discussion and texts than in informal discussion and text; and (b) that, unlike domain-specific words, appear across a variety of academic disciplines (e.g., Baumann & Graves, 2010; Coxhead, 2000). Examples of domain-specific academic words that appeared in the present study compositions are: *air*, *earth*, *extinct*, *habitat*, *oxygen*, *species* (science words); and *shelter* (a social studies word). An example of a general academic word from present study compositions is *conclusion*.

### *1.1 Why Does Academic Vocabulary Matter and Does It Really Matter for First-Grade Children?*

Children's acquisition of academic vocabulary knowledge is important for several reasons. Academic vocabulary knowledge is crucial for academic success and school achievement (e.g., Biemiller, 2009; Nagy & Townsend, 2012). Also, there is a strong positive association between students' vocabulary knowledge in general and reading comprehension (e.g., Verhoeven et al., 2011). That association has recently also been documented for academic vocabulary in particular in a few studies (e.g., Connor et al., 2015; Meneses et al., 2018).

Historically, at least in United States schooling, there has been a strong tendency to delay learning of challenging disciplinary topics and introduction of academic vocabulary until later years. However, over the course of the past two decades, some researchers have argued that academic vocabulary acquisition in preschool and the early grades is crucial to instantiating a basic conceptual system that supports current and later disciplinary reading and writing and content area learning (e.g., Neuman & Wright, 2014). Organizations such as the National Association for the Education of Young Children (Engel et al., 2016) and the Common Core State Standards (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010) emphasize academic vocabulary acquisition for all elementary-grade children, including kindergarten children.

## 1.2 Why Explore Academic Vocabulary in Compositions?

School, college, and many career settings demand that students develop and exercise successful written communication abilities. That success requires students' knowledge of academic language register, including specialized academic vocabulary (e.g., Truckenmiller & Petscher, 2020). In spite of the prominent role that composing plays in schooling, to date, very little is known about students' actual use of academic language, and academic vocabulary in particular, in students' composing and compositions. The few prior studies of academic vocabulary (or academic language generally) in students' compositions have tended to involve upper elementary and/or middle grades students (e.g., Galloway & Uccelli, 2019; Truckenmiller & Petscher, 2020). For example, one set of researchers found that fourth- and eighth-grade students who had higher levels of academic language, including academic vocabulary, tended to be more likely to pass a state writing assessment (Truckenmiller & Petscher, 2020).

## 1.3 Theoretically, What Does Word Use in Compositions Imply About a Student's Mental Word Store?

According to the lexical quality hypothesis (Perfetti, 2007), the likelihood that an individual will attempt to write a word, or an academic word in particular, derives from the quality of the individual's mental representation of that word. When a child has a strong mental representation of a word's phonology, orthographic form, and meaning, the child will more likely understand the word when reading and, by extension, will be more likely to attempt to use the word when writing (Dobbs & Kearns, 2016; Perfetti & Stafura, 2014).

Hypothetically, words that are both phonologically and orthographically similar to other words are likely to be easier for students to use when composing. Prior research supports that contention in that words that are orthographically similar to many other words are recognized by readers faster than are more distinctive words (e.g., Andrews, 1997). It is possible that a similar phenomenon occurs when writing. That is, phonological and orthographic patterns should be processed more efficiently because they may be better instantiated mentally and therefore may provide a "well-characterized attractor basin" (Yarkoni et al., 2008, p. 977). On the other hand, words that are unique phonologically and orthographically are likely to be more challenging to write because their sound and spelling patterns would be less well represented mentally.

Similarly, stronger and embellished semantic mental representation allows efficient, automatic, and reliable meaning retrieval during reading (Perfetti & Stafura, 2014). Again by extension, a well-connected mental network of academic words in a particular domain could theoretically facilitate students' word usage in compositions (Fitzgerald et al., 2020).

**Implications of the Lexical Quality Hypothesis for Academic Words**

Academic words are known to be morphologically complex and abstract in meaning (Schleppegrell, 2004). For example, *extinct* consists of two syllables. The syllable, *ex,* appears to function as an affix, but the word does not have a base word that would be immediately recognizable. Moreover, the meaning of *extinct* is not likely to be imageable. Academic words are also likely to be more phonologically and orthographically unique in comparison to other words. Again using *extinct* as an example, the combination of letters, *inct*, making the sounds */in-k-t/* would not be common orthographic or phonological combinations in English. Notably, academic words, especially domain-specific words, are also by definition words that appear infrequently in oral and written text. An implication is that academic words' tendency toward uniqueness, complexity, and rarity may present challenge to students' development of high quality mental lexical representation, at least potentially more challenge than other words.

## *1.4 But What About the First-Grade Situation?*

How much academic vocabulary should we expect to see in *first-grade children's* compositions? First-grade children's lexicon, and especially their academic vocabulary lexicon, is nascent. Although the development of a high-quality lexicon has roots even before schooling, first-grade children have not had significant opportunity to develop a deep and wide lexicon through enduring exposure to academic vocabulary in oral or written text. In that sense, one might expect that first-grade children's academic word repertoire would be budding rather than burgeoning. Similarly, first-grade children would not have had long-term opportunity to acquire orthographic and phonological word patterns (that is, the patterns referenced above in Sect. 1.3). As a result, their mental store of such patterns would be embryonic.

Still, the recent researcher and organizational policy emphasis on facilitating young children's academic vocabulary and disciplinary knowledge development (noted above in Sect. 1.1) could be impacting current classroom instruction. Young children may be receiving more academic vocabulary instructional and more academic vocabulary exposure in texts than in the past. Such exposure could strengthen children's academic vocabulary stock. Indeed, one publisher's most recent first-grade core reading programs contained more academic words than in earlier decades (Fitzgerald et al., 2021). From 1962 to 2013, on average the decade by decade rise in academic words was 54%. The estimated number of academic words in the 1962 core reading program was 1.51, and in 2013 it was 13.56.

Also, why might we expect an association between *first-grade children's* use of academic vocabulary with phonological, orthographic, and semantic word characteristics? One possible reason is related to how composing is introduced to young children in the early school years. For decades now, researchers and practitioners have encouraged young children to use "invented spelling" during composing. Invented spelling frees young learners from formal correctness and supports their

effort to focus on meaning. In that context, young children might be likely to use academic words in compositions when they have reasonably strong understanding of the words' meanings, even if they are unable to spell the words correctly. That is, their focus on instantiating meaning in text could prioritize semantic characteristics of words over phonological/orthographic characteristics. In such a context, a low-quality mental store of phonological/orthographic features of academic words may be inconsequential for first-grade children's word choice during composing.

### *1.5 What Is Needed Now?*

Better understanding of the extent to which first-grade children actually do use academic words could help to move the field forward for policy and practical reasons. If there is even minimal evidence of young children's use of academic vocabulary, that evidence could support the possibilities for academic vocabulary learning emboldened in some current policies. Relatedly, practitioners might be further encouraged to introduce disciplinary topics and academic vocabulary in a sustained manner.

Such findings might also imply qualifications to the lexical quality hypothesis. For example, perhaps during the nascent phases of lexical development, under certain conditions (e.g., instructional emphasis or task requirements) young children might give preference to high quality semantic mental representations over phonologic/orthographic representations.

To date, virtually nothing is known about the roots of children's mental academic word quality instantiation or about their use of academic vocabulary when composing. The present study represents an initial exploratory first step toward such understanding.

## 2 Research Questions and Design

The overarching purpose of the study was to explore first-grade children's academic vocabulary inclusion in compositions in response to an academically-oriented task. Two main research questions were addressed: (1) To what extent did the children incorporate academic vocabulary? (2) When children tended to use more academic words in compositions, did the phonological/orthographic uniqueness or the semantic challenge of words in compositions tend to increase?

The sample consisted of 191 first-grade control group children from a prior study (Kim et al., 2021). In the spring of first grade, after hearing a passage about a science topic, students were prompted to write an argumentative composition. Identification of academic words was derived from automated methods involving statistical classification modeling. Five variables were created for each composition. For the main research question, the raw count values for academic words were examined through descriptive statistics. For the secondary research issue, Multilevel

Poisson Regression was conducted. The outcome variable was number of academic words in a composition. Four predictor variables were characteristics of words in the compositions: (a) phonological uniqueness (how similar a word sounds in comparison to other words); (b) orthographic uniqueness (how similar a word's spelling is compared to other words); (c) semantic challenge (the age at which a word is commonly known) and (d) semantic concreteness (how tangible or imageable a word's meaning is). Each of the four predictor variables was operationalized as the average score for all word types in a composition.

## 3 Method

### *3.1 Participants, Schools, and Teachers*

The 191 first-grade students were in 13 classes distributed among 10 schools in one school district in the southeast United States. The students' ethnicity/racial identification according to school records was: 28.27% African American (n = 54), 27.23% Caucasian of European descent (n = 52), 32.98% Latinx (n = 63), 8.38% Asian (n = 16), and 3.14% multiethnic/racial (n = 6). Male students represented 43.46% of the sample (n = 83), and 56.54% were female (n = 108). English-language learners constituted 22.51% of the sample (n = 43), and 4.71% of students were classified as requiring special education (n = 9).

Five of the 10 schools were Title 1 schools. On average, across the 10 schools, 45.87% of students received free/reduced lunch (SD = 20.67%), and the range across the 10 schools was 15.20% to 68.90%. All 13 teachers held elementary grades teaching licenses.

### *3.2 The District's Literacy Curriculum*

The teachers were expected to follow the district's scope and sequence for first-grade language arts. The district curriculum did not specify the extent to which teachers were expected to attend to academic vocabulary. Typical literacy instruction in first grade was considered to be a balanced program which included word study, guided reading, and writing activities. Small group and whole class teacher-directed instruction was conducted and independent reading was included in the curriculum. Teachers followed a district-created calendar to support instructional decisions and pacing. For example, during a spring 10-week period, teachers involved students in a fiction unit conducted in a readers' workshop format. The workshop focused on retelling, asking and answering questions, and identifying who was telling the story at various points in time. For writing workshop during the same time period, teachers led a unit on persuasive writing.

### 3.3 The Composition Task and Transcription

Students were asked to write an argumentative composition in response to hearing a passage about rainforests. The examiner read the rainforest passage aloud to students as a group. The writing prompt was, "Should people be allowed to cut down trees in the rainforest?" Students were specifically directed to answer the question by making an argument. The directions also stated that a good argument provides an opinion, reasons, explanations, and a conclusion. Three minutes were provided for planning and 40 min were allotted to write the composition.

Because the rainforest passage was an informational passage, it was expected that students would either already have some relevant background knowledge for the writing task, or they would glean some knowledge from the orally presented passage. Also, because of the disciplinary nature of the topic and the informational genre of the rainforest passage, it was expected that students would need to use at least some small modicum of academic words in the compositions. Examples of academic words in the passage are: *rainforest*, *weather*, *crop*, and *organism*.

Compositions were transcribed to represent children's semantic intentions. That is, surface errors were corrected (misspelling, incorrect capitalization, incorrect punctuation). Transcriptions were then digitized.

### 3.4 Automated Identification of Academic Words

Academic words were identified in students' compositions using an automated statistical classification method. The conceptual reasoning that guided the automated statistical classification method was as follows. The logic was identical to the logic used by prior researchers of *general* academic vocabulary (Coxhead, 2000; Gardener & Davies, 2014). In conceptual terms, the process for identifying academic words mirrored the definition of academic vocabulary presented in the opening of the present manuscript.

In brief, academic words are defined as words that appear more frequently in disciplinary textbooks than in fiction or informal oral language. Domain-specific academic words tend to appear mainly in a particular academic discipline, and general academic words are academic words that tend to appear across a variety of academic disciplines. To capture words that met those definitions, the modeling included predictor variables that assessed: (a) the comparative presence of a given word in disciplinary textbooks versus reading/English language arts (ELA) textbooks or oral corpora; and (b) the comparative presence of a given word within and across disciplinary textbooks. A word was considered to be an academic word if it was more likely to appear in disciplinary texts than in reading/ELA or oral corpora. A word was considered to be domain-specific academic if it was more likely to appear in a particular disciplinary textbook than in other disciplinary textbooks. A general academic word had higher likelihood of appearing across the

disciplinary texts than in a single disciplinary textbook or than in reading/ELA or oral language corpora.

Statistical predictor classification models were constructed to operationalize the logic of word usage. Here is an overview of construction of the predictor models. First, the modeling required a predetermined set of words identified as academic or not. So in a first phase, using words from a separate corpus, human judges provided decisions about whether words were academic or not, and if a word was deemed to be academic, it was classified as one of the four academic word categories (science, mathematics, social studies, or general academic). In a second phase, the judges' determinations from Phase 1 were used as the outcome variable in statistical modeling. Predictor variables were entered to build a statistical model that successfully predicted the judges' prior determinations. Then, the predictor model was employed to identify all academic words in the first-grade children's compositions.

Here are some details for the two phases. In Phase 1, after training, using a sample of 1800 words randomly selected from a separate corpus of 144 first through twelfth grade textbooks, two judges' inter-agreement reliability for whether a word was academic was .85. For words deemed to be academic, the inter-agreement reliabilities were .87, .92, .91, and .89 for science, mathematics, social studies, and general academic words, respectively.

In Phase 2, four Random Forest classification models (Breiman, 2001) were employed (one each for science, mathematics, social studies, and general academic words). The separate corpus of 144 textbooks was again used in Phase 2. The judge's classification decisions for the 1800 words from Phase 1 were the dependent variables in the models. The models consisted of 14 predictors in 5 categories: for each word, (a) relative frequency in each grade in a first- through twelfth-grade textbook corpus (7 predictors, one for each of the Grades 1 through 5, and one for Grades 6 through 8 collectively, and one for Grades 9 through 12 collectively); (b) relative frequency in the textbook corpus in each academic textbook domain compared with the other textbook domains (4, one each for science, mathematics, social studies, and reading textbooks); (c) the word's dispersion across four textbook disciplines and all grades (1) (Gries, 2008); (d) relative frequency in the textbook corpus compared with an oral language corpus (1) (Brysbaert & New, 2009); and (e) the word's abstractness (1) (Brysbaert et al., 2014).

Finally, the validity of the classification models' academic word identification was evaluated. Using the 1800 words originally categorized by the two judges, for determining whether a word was academic or not, classification model determinations were in high accord with each judge's determinations—.90 and .91, respectively. When a word was academic, there was high agreement between classification models and each judge's assignments: .91 and .89 for science, .93 and .93 for mathematics, .93 and .92 for social studies, and .86 and .87 for general academic words.

The final predictor models were then used to classify every word in every student's composition in the present study. Each word was identified as academic or not.

## *3.5 Variables*

### Types and Tokens

A type was a word family that consisted of a base word (e.g., *survive*) and its inflections (*survives*, *survived*, and *surviving*). A token was a word. Whereas *survive*, *survives*, *survived*, and *surviving* were counted as one type, the same words were counted as four tokens.

### Number of Academic Words

Academic word types in a composition were summed. The range in the present sample was 0–5 academic word types in a composition. Academic word tokens were also summed, and the range in the present sample was 0–96 tokens in a composition.

### Phonological, Orthographic, and Semantic Characteristics of Words in Compositions

For each of the following variables, the score for each word type in a composition was then averaged across all words in a composition.

#### Phonological Uniqueness

Phonological Uniqueness of the base word in every word family type in a composition was measured by the Phonological Levenstein Distance (PLD20) metric (Balota et al., 2007). The PLD is intended to capture humans' judgement of how phonologically similar or unique words are from one another. It was the average number of sound changes (additions, substitutions, etc.) that would have to be accomplished to make a word into its auditorily 20 nearest neighboring words. For instance, the PLD20 for *air* (/ay/ /er/) is 1.00, so one sound change in *air* would be required to make another word. For instance, adding /b/ would change the sounds to *bare* (/b/ / ay/ /er/). For *air*, the average number of sound changes when compared to the 20 nearest neighbors was 1.00. On the other hand, the PLD20 for *oxygen* is 3.55. A higher PLD20 indicated that fewer words in the closest 20 words share phonology, or that a word had unique phonology compared to other words. The possible range of scores could be 1.00 to an upper limit bounded by the sample word containing the most sounds.

### Orthographic Uniqueness

Orthographic Uniqueness was measured by the Orthographic Levenstein Distance (OLD20) metric (Yarkoni et al., 2008). It was the number of letters in a word that would have to be changed to spellings of its 20 nearest neighbors. For example, *oxygen*'s OLD20 score is 3.05 meaning that approximately three changes need to be made to turn its spelling into the 20 nearest neighbors. A higher Orthographic Uniqueness score indicated that few words in the closest 20 words shared orthography with the target word, or that a word had unique orthography compared to other words. The final score range could be 1.00 to an upper limit bounded the sample word with the largest number of letters.

### Semantic Challenge

Semantic Challenge was operationalized as a word's age-of-acquisition measure. Age-of-acquisition values for each academic word were obtained from the Kuperman and colleagues' database (Kuperman et al., 2012; updated at the Center for Reading Research, Ghent University website, n.d.). Age-of-acquisition values are ratings on a scale that orders words according to the chronological age at which their meanings are first known. A higher age-of-acquisition value indicated a word that was, on average, learned later. The final score could range from approximately 3 years of age to any higher chronological age.

### Semantic Concreteness

Semantic Concreteness was the degree to which a word referenced a concept that could be seen or imaged (Brysbaert et al., 2014). A higher score indicated a word for a concept that was more concreate, more imageable. The final score range could be 0–7.

## Additional Variables for Participant Descriptions

Two additional scores were used to describe the participants' writing and reading, but the scores were not used in analyses—an argumentative writing score and a reading score. The argumentative writing score was the sum of the presence of a claim, supporting evidence, and closure. The possible score range was 0 to 7. Inter-agreement reliability (Cohen's κ) was .74.

Reading achievement in spring of first grade was assessed by the Measure of Academic Progress (MAP) Primary Grade Reading test (NWEA: Northwest

Evaluation Association, n. d.). The test measures students' comprehension of literature and informational text, vocabulary meaning knowledge, foundational phonics skills and decoding, and English conventions. The score was the total Rasch unit scale score, with a possible range from approximately 100 to 300. Measure of Academic Progress test-retest reliability reported by NWEA was from .89 to .96, depending on grade.

### *3.6 Analyses*

The main analytical model was a Poisson regression model. Poisson regression is used when the outcome is a count variable (in our case, number of academic words). The Poisson regression accounts for the special distribution for count values, including when many zero values are possible. Because the data originated from multiple classrooms nested in multiple schools, we began by considering multi-level Poisson regression. We found that the Variance Partition Coefficients (VPC, Poisson counterparts for Intraclass Coefficients) for class and school were very low (.05 and .01 respectively). The VPCs indicated that very little of the total variance in the academic vocabulary outcome was due to class or school. Consequently, we conducted single-level Poisson regression.

Number of academic words was the dependent variable. Four predictor variables were: phonological uniqueness, orthographic uniqueness, semantic challenge, and semantic concreteness. Control variables were composition length (the total number of tokens in a composition) and school poverty (a school's percent of students receiving free or reduced lunch).

The effect size metric for the Poisson analysis is the Incidence-Rate Ratio (IRR) (Frome & Checkoway, 1985). The IRR is the ratio by which the predicted count of the outcome variable would change given a one-unit increase in the predictor. Incidence-Rate Ratios of less than one indicate a relative decrease in the outcome's incidence, and IRRs of greater than one indicate a relative increase in the outcome's incidence. There is no set of "standard cut points" for magnitude. The IRR can only be interpreted in context.

## 4 Results

To provide context for the reader as analytical results are presented, Table 1 exhibits a few of the children's compositions. Features of the table will be discussed in following sections.

### 4.1 Preliminary Analysis of Data Structure

Descriptive statistics for each variable and correlations among variables in the present study are presented in Tables 2 and 3. The average number of types in compositions was 18.08, and the average number of tokens was 25.72. Not shown in tables are the following: On average, the students' argumentative writing score was 3.15 (*SD* = 1.68, range = 0 to 7). The average reading score was 173.75 (*SD* = 13.24, range = 129 to 208).

When considering all words in a composition (Table 2, column 2), the average phonological uniqueness and orthographic uniqueness measures indicated that, on average, relatively few sounds (1.34) and relatively few letters (1.59) in words would be changed to arrive at the sounds and spellings of the 20 nearest neighboring words. The average semantic challenge measure (age at which the words would commonly be known = 4.63) indicated that, when all words in a composition were considered, the word meanings in the compositions were likely to be commonly known by the first-grade children whose chronological ages were approximately 6 to 7 years old. The average semantic concreteness measure (2.80) suggested that on the whole, the words in the compositions tended very slightly towards abstractness (given that the range of possible concreteness scores was 0–7, with a score of 4 in the center).

As seen in Table 3, intercorrelations among variables tended to follow expected patterns. The number of academic words in compositions correlated moderately positively (and significantly) with the number of tokens (correlation = .38). More academic words in a composition tended to accompany greater presence of words with unique phonology, unique orthography, and greater semantic challenge levels. Surprisingly, number of academic word types was not statistically significantly associated with word type concreteness.

With the exception of correlations with semantic concreteness, the six intercorrelations among the four characteristics of all words in compositions were moderately high to very high and in expected directions. Phonological uniqueness and orthographic uniqueness were most highly correlated (.84).

Finally, writing scores correlated moderately positively with reading scores (.40). Notably, when considering remaining correlations involving writing or reading measures, only composition length (N of tokens in composition) was significantly (positively), but weakly, related to writing scores.

### 4.2 Research Question 1: To What Extent Did the Children Incorporate Academic Vocabulary?

As seen in column 3 of Table 2, the average number of academic word types in compositions was quite low with a mean of .88—just under 1 academic word per composition. As might be expected, the distribution was slightly positively

**Table 1** Composition examples by N of academic word types in compositions with some descriptives

| N of academic word types | Example compositions | Academic words in composition | *M* phonological uniqueness | *M* semantic challenge | Writing score |
|---|---|---|---|---|---|
| 0 | 1. No because the animals need a place to live and the animals will die | NA | 1.33 | 4.34 | 4 |
| | 2. We should not be allowed to cut down trees because animals live there and it makes food. | NA | 1.32 | 4.92 | 7 |
| 1 or 2 | 3. I think people should be allowed to cut down the rainforest | Rainforest | 1.31 | 5.19 | 1 |
| | 4. You should not because monkeys will be endangered. You should not kill them because it's God's creation. You should not because animals live in trees. You should not kill trees because less oxygen. | Oxygen, endanger | 1.60 | 5.47 | 7 |
| 3 to 5 | 5. I don't think you should cut down trees in the rainforest because you're taking air away. You are also destroying habitats. You are also killing animals. I don't think you should cut trees down in the rainforest | Rainforest, air, habitat | 1.49 | 5.21 | 7 |
| | 6. In my opinion, people should not be allowed to cut down trees because plants and animals are getting endangered and becoming extinct. Rainforest are home to unique animals they are very rare animals. There are so many trees in the rainforest | Endanger, extinct, rainforest, unique | 1.52 | 5.20 | 7 |

skewed. Still, there was a wide range for the number of academic word types in compositions, from 0 to 5 types. Approximately 61% of the children incorporated at least one academic word in their compositions. Specifically, 39% of children incorporated no academic word types, 57% incorporated 1 or 2 academic word types, and 4% incorporated 3 to 5 academic word types.

**Table 2** Mean, SD [Range] for N of types and tokens in compositions, all four predictors, school poverty, writing, and reading

| Variable | All words in composition | Academic words in composition | Non-academic words in composition |
|---|---|---|---|
| N of word types | 18.08 (9.12) [1.00–52.00] | 0.88 (0.93) [0.00–5.00] | 17.14 (8.66) [1.00–51.00] |
| N of word tokens | 25.72 (17.09) [1.00–96.00] | 1.12 (1.28) [0.00–7.00] | 24.60 (16.45) [1.00–51.00] |
| | All word types | Academic word types | Non-academic word types |
| Phon uniqueness | 1.34 (0.13) [1.00–1.88] | 2.60 (0.77) [1.00–3.55][a] | 1.31 (0.11) [1.00–1.88] |
| Orthogr uniqueness | 1.59 (0.12) [1.15–2.08] | 2.50 (0.52) [1.45–3.60][a] | 1.57 (0.11) [1.15–2.08] |
| Semantic challenge | 4.63 (0.43) [2.72–5.53] | 7.50 (0.91) [3.94–9.85][b] | 4.48 (0.39) [2.72–5.47] |
| Semantic concrete | 2.80 (0.27) [1.90–3.54] | 4.03 (0.80) [1.96–4.80][b] | 2.74 (0.29) [1.75–3.54] |

*Note.* Some compositions had no academic words, and consequently the number of compositions involved in calculating four variables was fewer than 191

[a]*n* of compositions = 62

[b]*n* of compositions = 117; all other *n* of compositions = 191. Phon Uniqueness = Phonological Uniqueness; Orthogr Uniqueness = Orthographic Uniqueness

**Table 3** Correlations among variables

| | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|
| 1. N of academic words types | .38[a]*** | .16[b]*** | .04[b]*** | .33[b]*** | .06[b] | .15[b] | .10[b] |
| 2. N of tokens in composition | | .13[b] | .01[b] | .29[b]*** | .05[b] | .15[b]** | .13[b] |
| 3. Phonological uniqueness | | | .84*** | .35*** | .02 | .02 | .05 |
| 4. Orthographic uniqueness | | | | .26*** | .01 | .02 | .03 |
| 5. Semantic challenge | | | | | .13 | .08 | .10 |
| 6. Semantic concreteness | | | | | | .02 | −.06 |
| 7. Writing | | | | | | | .40*** |
| 8. Reading | | | | | | | |

*$p < .05$; **$p < .01$; ***$p < .001$; Except as the following superscripts note, correlations are Pearson Correlations

[a]Kendall's Tau

[b]Spearman

### 4.3 Research Question 2: Selected Word Characteristics in Compositions Were Related to N of Academic Words in Composition

On average, as numbers of academic words in compositions increased, compositions were also characterized by words that were more unique in phonology (phonological uniqueness $b = 2.70$, $SE = 1.20$), $p < .05$) and more semantically challenging (semantic challenge $b = 1.44$, $SE = .28$, $p < .001$). (Table 4 provides the sources of variance.) The effect for both predictors was relatively small. On average, a 0.10 increase in a composition's average phonological uniqueness measure was associated with a 31% rise in number of estimated academic words in a composition (IRR = $14.91^{.10} = 1.31$, $p < .05$). On average, a 0.10 increase in semantic challenge (one-tenth of a chronological year) was associated with a 15% increase in the estimated number of academic word types in a composition (IRR = $4.20^{.10} = 1.15$, $p < .001$).

### 4.4 Further Exploration

To better understand the two statistically significant results, we examined the word characteristic scores (using the original raw scores) according to bands for number of academic word types in compositions. Although only phonological uniqueness and semantic challenge were statistically significantly related to number of words in compositions in the modeling, as shown in Table 5, as the bands increased, the average phonological uniqueness, orthographic uniqueness, and semantic challenge of words in compositions also increased. For instance, the average semantic challenge of words (column 4) for compositions with no academic words was 4.34 years of age. It rose to 4.80 for compositions with 1 or 2 academic words, and again rose to 5.02 years of age for compositions with 3 to 5 academic word types.

**Table 4** Sources of variance for the final fitted poisson regression model

| Source | Coefficient (*SE*) | IRR (*SE*) |
|---|---|---|
| Intercept | −2.37 (1.19)* | .09 (.11)*** |
| Phonological uniqueness | 2.70 (1.20)* | 14.91 (17.96)* |
| Orthographic uniqueness | −1.23 (1.31) | .29 (.38) |
| Semantic challenge | 1.44 (0.28)*** | 4.20 (1.17)*** |
| Semantic concreteness | −0.03 (0.33) | .97 (.32) |
| School poverty, $\gamma_{001}$ | −0.01 (0.01) | 1.00 (1.17)*** |
| Composition length, $\gamma_{500}$ | .02 (.01)*** | 1.02 (.01) |

*IRR* = incident rate ratio
*$p < .05$, **$p < .01$, ***$p < .001$

**Table 5** Composition word characteristics *M*, SD, and range by academic words types in composition bands

| | Composition word characteristic M (SD) [Range] | | | |
|---|---|---|---|---|
| Bands for N of academic word types in composition | Phonological uniqueness | Orthographic uniqueness | Semantic challenge | Semantic Concreteness |
| 0 (n = 74) | 1.29 (0.14) [1.00–1.88] | 1.57 (0.14) [1.15–2.08] | 4.34 (0.50) [2.72–5.47] | 2.80 (0.34) [1.90–3.54] |
| 1 or 2 (n = 109) | 1.35 (0.12) [1.05–1.79] | 1.60 (0.11) [1.34–1.95] | 4.80 (0.24) [4.13–5.53] | 2.80 (0.23) [2.12–3.42] |
| 3 to 5 (n = 8) | 1.47 (0.12) [1.29–1.68] | 1.68 (0.10) [1.53–1.82] | 5.02 (0.26) [4.55–5.24] | 2.82 (0.18) 2.54–3.13] |

To assess the extent to which addition of academic words in compositions may have been responsible for the statistically significant association between number of academic words and both phonological uniqueness and semantic challenge, we compared phonological uniqueness and semantic challenge of academic word types to non-academic word types. Looking at the last two columns of Table 2, we can see the observed means for the two significant predictors broken out by academic and non-academic word types. For both predictors, the raw score averages are higher for academic words than for non-academic words. Consequently, by inference, the statistically significant relationships for number of academic words with phonological uniqueness and semantic challenge may be attributable to the addition of academic words in compositions.

Finally, Table 1 provides examples of compositions in three bands for number of academic words. First, the positive correlation mentioned earlier for composition length with number of academic word types in compositions is notable in the table. (Recall that composition length was controlled in the statistical modelling.) Second, if we compare the phonological uniqueness of words in examples 5 and 6 versus 1 and 2, in spite of the length difference, on the whole, words in 5 and 6 tend to be longer, reflecting more sounds in words, and generally, more sounds in a word tends to indicate more phonological uniqueness. For instance, comparing example 6 to example 2, 33.33% of the word types in example 6 are multisyllabic whereas 20% of the types in example 2 are multisyllabic. Third, comparing the semantic challenge across example compositions, some words in 5 and 6 stand out as higher in age of acquisition than words in examples 1 and 2. For instance, in compositions 5 and 6, *endanger*, *extinct*, *unique*, *opinion*, and *habitat* have age of acquisition ratings of 9.4, 8.1, 8.7, 8.8, and 9.4, respectively. In examples 1 and 2, nearly all words have age of acquisition ratings below 6.2, and most are below 4.0.

# 5 Conclusions, Limitations, and Discussion

## *5.1 Conclusions*

The main conclusions were: (a) Given the academically-oriented writing task, slightly more than half of the children (57%) included one or two academic word types in their compositions. A very few children (4%) incorporated three to five academic word types. Slightly more than a third of the children (39%) wrote no academic words. (b) Greater inclusion of academic words in a composition was weakly associated with composition words that were more phonologically unique and more semantically challenging.

The study is, to our knowledge, the first to explore young children's use of academic vocabulary in compositions. We now have evidence that some, if not many, first-grade children did choose to use academic vocabulary, some of which co-occurred in a disciplinary text they had heard. The study is also among the first to explore connections of academic word use with word characteristics in a theoretically-based framework.

## *5.2 Limitations*

Limitations of the present exploratory study set boundaries and context for interpretation of findings. A main limitation of the present study was that only one writing task was accomplished. Topic is known to be significantly related to children's writing production and judgments of the quality of compositions. Inclusion of at least one additional composition accomplished under the same task conditions but for a different science topic would strengthen generalizability of findings. A second limitation was that no explicit information was obtained about the teachers' attention to academic vocabulary in everyday instruction. Although the district literacy curriculum did not specify an expected degree of attention to academic vocabulary, it would be useful to query the teachers about it.

## *5.3 Discussion*

### Children's Use of Academic Words in Compositions

The first-grade children's use of academic vocabulary in compositions was remarkable in that the children had minimal exposure during the first step of the writing task in the present study—they heard academic words in the rainforest passage that was read to them. Even after hearing the passage only once, more than half of the children used at least some academic words in their compositions.

The Goldilocks question comes immediately to mind—"Was the academic vocabulary presence in the compositions just right, not enough, or too much?" It is impossible to answer the question because there is no standard and no historical precedent against which to judge. What we can say is that we now have some evidence that some, if not many, first-grade children *do* choose to use academic vocabulary (at least to some minimal degree) when posed with an academically-oriented composition task.

The reasons for the variation among children in academic word usage are not clear, though some conjecture could be in order. On the one hand, we might expect that some children would employ academic words for the present composition task, especially given the academic nature of the task structure. Also, as we noted earlier, there has been increased policy emphasis on the importance of disciplinary learning and academic vocabulary in the earliest grades of schooling. Also, recently, some core reading program publishers have claimed inclusion of academic vocabulary in the earliest grades (e.g., Houghton Mifflin *Journeys Common Core* [Houghton Mifflin Harcourt, n.d.; Vogt et al., 2014]), and minimal evidence suggests that one publisher's contemporary first-grade core reading program includes more academic words than programs in past years (Fitzgerald et al., 2021). Without evidence of their impact, we can only speculate that policy and publishers' text changes may have impacted first-grade children's lived classroom experiences.

On the other hand, it may not be surprising that some first-grade children would use no academic words in their compositions. In the United States, during the last national assessment of students' writing, 67% of fourth-grade students and 76% of eighth-grade students did not reach "proficiency" on the most recent National Assessment of Educational Progress writing assessment (U.S. Department of Education, Institute of Education Science, National Center for Education Statistics, 2011). Proficiency has its roots in the earliest years. If such a large percentage of fourth-grade and eighth-grade students are not "proficient" writers, then it seems highly likely that a large percentage of first-grade children's acquisition and development of writing processes lags. Moreover, in the present sample, there was a wide range of writing scores for the writing task, with some students scoring well below the average. Learning to write involves a set of complex processes (e.g., MacArthur & Graham, 2016), and word choice is embedded in those complex processes. Given those contexts, one could expect that some first-grade children would not have acquired some composing processes, including word-choice processes, at the same rate as their counterparts.

### When Students Used More Academic Words in Compositions, the Compositions Reflected Increased Word Phonological Uniqueness and Semantic Challenge

Because academic words tend to be characterized as morphologically complex and as abstract, they are also likely to be phonologically/orthographically unique and semantically challenging in comparison to other words. In the present study,

first-grade students' inclusion of a very small number of academic words in compositions, on average, raised phonological uniqueness and semantic challenge of words in compositions. A follow-up examination offered some support for the contention that the increases could be at least partly due to the presence of the academic words. In the context of the lexical quality hypothesis, the finding suggests that children who choose to use even a small number of academic words may tend to have stronger mental representations for phonological and semantic features of words.

The lack of a statistically significant effect for orthographic uniqueness was not surprising given the high correlation between phonology and orthographic. The number of sounds in words is necessarily highly correlated with the number of letters in words. The high correlation between the two variables in the present study is witness to that effect.

The present results may be compared to findings from just one prior study involving middle grades students where no relationships were located for any phonological, orthographic, or semantic feature of academic words with students' attempts to use, or actual use of, taught academic words in compositions (Dobbs & Kearns, 2016). In that study, only a word frequency measure was (positively) related to students' inclusion of academic words in compositions. Reasons for the differences in results are not easily discernable. Differential findings could be due to differences in tasks, scoring, students' age, and/or the fact that the words were taught to students in the prior study but not in the present study.

## Future Research

The findings from the present project were sufficiently positive to support the need for further, rigorous investigation of young children's academic word usage during composing. The possibilities are numerous and could take on a wide variety of directions. To name a few examples: The generalizability of findings across two or more writing samples, and across different types of tasks is warranted; and academic word usage could be compared across a control condition where disciplinary topics and academic vocabulary were not emphasized and a treatment condition in which they were heavily emphasized.

The measures of word phonology and orthography in the present study have been well researched and widely used. However, in the statistical modeling, the definition of a "one-unit" change in both word characteristics was unclear. When associated with the log of a count variable in a Poisson regression, interpretability of the association in original value terms was additionally challenged. An additional issue was the strong correlation between the two variables. In the present study, the aim was solely to ascertain whether there was an association of each of the two predictors and number of academic words in composition. That is, there was little interest in coefficient interpretation. Because our primary goal was only to assess the association, there wasn't a need to reduce collinearity in the regression modeling (cf. Neter et al., 2020). However, in future research, other computational measures might be

considered. For instance, the phonological variable, phoneme count, could be more interpretable than the measure in the current study. Similarly, bigram frequency, an indicator of orthographic complexity, has been used in at least one prior study involving older children (Dobbs & Kearns, 2016). It may be more interpretable.

## References

Andrews, S. (1997). The effect of orthographic similarity on lexical retrieval: Resolving neighborhood conflicts. *Psychonomic Bulletin & Review, 4*(4), 439–461.

Balota, D. A., Yap, M. J., Cortese, M., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*(3), 445–459.

Baumann, J. F., & Graves, M. F. (2010). What is academic vocabulary? *Journal of Adolescent & Adult Literacy, 54*, 4–12. https://doi.org/10.1598/JAAL.54.1.1

Biemiller, A. (2009). *Words worth teaching*. McGraw-Hill, Science Research Associates.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brysbaert, M., & New, B. (2009). Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of new and improved word frequency measures for American English. *Behavior Research Methods, 41*(4), 977–990. https://doi.org/10.3758/BRM.41.4.977

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*(3), 904–911. https://doi.org/10.3758/s13428-013-0403-5

Center for Reading Research, Ghent University. (n.d.). *Age-of-acquisition (AoA) norms for over 50 thousand English words.* Retrieved from http://crr.ugent.be/archives/806

Connor, C. M., Radach, R., Vorstius, C., Day, S. L., McLean, L., & Morrison, F. J. (2015). Individual differences in fifth graders' literacy and academic language predict comprehension monitoring development: An eye-movement study. *Scientific Studies of Reading, 19*(2), 114–134. https://doi.org/10.1080/10888438.2014.943905

Coxhead, A. (2000). A new academic word list. *TESOL Quarterly, 34*, 213–238. https://doi.org/10.2307/3587951

Dobbs, C. L., & Kearns, D. (2016). Using new vocabulary in writing: Exploring how word and learner characteristics related to the likelihood that writers use newly taught vocabulary. *Reading and Writing, 29*(9), 1817–1843. https://doi.org/10.1007/s11145-016-9654-8

Engel, M., Claessens, A., Watts, T., & Farkas, G. (2016). Mathematics content coverage and student learning in kindergarten. *Educational Researcher, 45*(5), 293–300. https://doi.org/10.3102/0013189X16656841

Fitzgerald, J., Elmore, J., Relyea, J. E., & Stenner, A. J. (2020). Domain-specific vocabulary network development in elementary years. *Journal of Educational Psychology, 112*(5), 855–879.

Fitzgerald, J., Relyea, J. E., Elmore, J., & Hiebert, E. H. (2021). Has the presence of first-grade core reading program academic vocabulary changed across six decades? *Reading Research Quarterly, 112*(5), 855–879. https://doi.org/10.1002/RRQ.329

Frome, E. L., & Checkoway, H. (1985). The use of Poisson regression models in estimating incidence rates and ratios. *American Journal of Epidemiology, 121*(2), 309–323. https://doi.org/10.1093/oxfordjournals.aje.a114001

Galloway, E. P., & Uccelli, P. (2019). Beyond reading comprehension: Exploring the additional contribution of core academic language skills to early adolescents' written summaries. *Reading and Writing, 32*(3), 729–759. https://doi.org/10.1007/s11145-018-9880-3

Gardner, D., & Davies, M. (2014). A new academic vocabulary list. *Applied Linguistics, 35*(3), 305–327. https://doi.org/10.1093/applin/amt015

Gries, S. T. (2008). Dispersions and adjusted frequencies in corpora. *International Journal of Corpus Linguistics, 13*(4), 403–437. https://doi.org/10.1075/ijcl.13.4.02gri

Houghton Mifflin Harcourt. (n.d.). *Journeys Common Core: A research-based approach*. Retrieved from https://www.hmhco.com/~/media/sites/home/educators/education-topics/hmh-efficacy/journeys_cc_2014_national_research_base_k-6.pdf

Kim, J. S., Burkhauser, M. A., Mesite, L., Ascher, C., Relyea, J. E., Fitzgerald, J., & Elmore, J. (2021). Improving reading comprehension, science domain knowledge, and reading engagement through a first-grade content literacy intervention. *Journal of Educational Psychology.* https://doi.org/10.1037/edu0000465

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods, 44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 24–40). Guilford.

Meneses, A., Uccelli, P., Santelices, M. V., Ruiz, M., Acevedo, D., & Figueroa, J. (2018). Academic language as a predictor of reading comprehension in monolingual Spanish-speaking readers: Evidence from Chilean early adolescents. *Reading Research Quarterly, 53*(2), 223–247. https://doi.org/10.1002/rrq.192

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91–108. https://doi.org/10.1002/RRQ.011

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for English language arts and literacy in history/social studies, science, and technical subjects*. Authors.

Neter, J., Kutner, M. H., & Nachtshelm, C. J. (2020). *Applied linear statistical models*. McGraw-Hill Education.

Neuman, S. B., & Wright, T. S. (2014). The magic of words: Teaching vocabulary in the early childhood classroom. *American Educator, 38*(2), 4–13.

NWEA: Northwest Evaluation Association. (n.d.). mapGROWTH. https://www.nwea.org/map-growth/

Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357–383. https://doi.org/10.1080/10888430701530730

Perfetti, C. A., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading, 18*(4), 22–37. https://doi.org/10.1080/10888430701530730

Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Erlbaum. https://doi.org/10.4324/9781410610317

Truckenmiller, A. J., & Petscher, Y. (2020). The role of academic language in written composition in elementary and middle school. *Reading and Writing, 33*(1), 45–66. https://doi.org/10.1007/s11145-019-09938-7

U. S. Department of Education, Institute of Education Sciences, National Center for Education Statistics. (2011). *The Nation's Report Card: Writing 2011*. NCES 2003–529. Washington, DC: Author. https://nces.ed.gov/nationsreportcard/pubs/main2011/2012457.asp#section1

Verhoeven, L., van Leeuwe, J., & Vermeer, A. (2011). Vocabulary growth and reading development across the elementary school years. *Scientific Studies of Reading, 15*(1), 8–25. https://doi.org/10.1080/10888438.2011.536124

Vogt, M. E., Hougen, M. C., Jago, C., Palmer, E., Templeton, S., Valencia, S. W., & Fountas, I. (2014). *Journeys common core*. [Grades K-6 series]. Boston: Houghton Mifflin Harcourt.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's *N*: A new measure of orthographic similarity. *Psychonomic Bulletin & Review, 15*(5), 971–979.

# Examining the Impact of an Integrated STEM and Writing-to-Learn Unit on Third Graders' Writing Knowledge and Writing Performance


**Amy Gillespie Rouse, Ashley Sandoval, and Murphy K. Young**

**Abstract** In this chapter, we examined students' knowledge of the writing process and their knowledge of how writing can be used for different purposes in different subject areas. After participating in a 10-day integrated science, technology, engineering, and mathematics (STEM) and writing-to-learn unit, we interviewed 26 third-grade students, asking them to describe the writing process as well as how writing in science class differed from writing in literacy or English/language arts class. We coded students' oral responses to interview questions using a rubric we developed based on Graham's (Educ Psychol 53(4):258–279, 2018) Writers-within-Community model. We examined students' posttest interview responses as well as whether students' posttest writing knowledge was predictive of their performance on a posttest expository essay, controlling for gender and writing abilities. Implications for instruction and next steps for research on writing-to-learn are discussed.

**Keywords** Writing · STEM · Writing to learn · Elementary

This chapter focuses on how writing can be used as a tool to facilitate students' learning of information in science, technology, engineering, and mathematics (STEM). We begin with a brief review of research and theory supporting the use of writing as a tool for learning. The majority of the chapter is devoted to our analysis of the general and STEM-specific writing knowledge and writing performance of a group of third-grade students who participated in an integrated STEM and writing-to-learn unit. We conclude with our hopes for the future of writing-to-learn instruction and research.

We define writing-to-learn as the act of students completing a writing activity intended to help them think and learn about subject matter (Klein et al., 2019). Examples of this type of activity in STEM range from simple recall tasks, such as

A. G. Rouse (✉) · A. Sandoval · M. K. Young
Southern Methodist University, Dallas, TX, USA
e-mail: a.rouse@uq.edu.au



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_6

writing short answers to questions after a science experiment, to more cognitively demanding tasks, such as composing a written argument to defend a point-of-view about a STEM topic. The research on writing-to-learn shows that either type of writing activity may be effective for increasing learning (e.g., Graham et al., 2020), but results are not consistent across studies.

## 1 Writing-to-Learn Research and Theory

Early studies of writing-to-learn instruction mostly involved professors conducting action research projects with their own college-aged students (Durst & Newell, 1989; Klein & Boscolo, 2016), and early reviews of writing-to-learn instruction (e.g., Ackerman, 1993) reported inconclusive findings about its effectiveness. However, beginning in the 2000s, with new research methods such as meta-analysis, a greater consensus emerged supporting positive effects of writing-to-learn instruction in K-12 classrooms (Bangert-Drowns et al., 2004; Graham & Perin, 2007).

However, the context for and types of writing that best facilitate learning have often been debated (Klein & Boscolo, 2016). In their meta-analysis of writing-to-learn studies, Bangert-Drowns et al. (2004) found positive effects of writing on learning for elementary through college-aged students, with longer writing treatments and use of metacognitive prompts during writing resulting in greater effects on learning. Bangert-Drowns et al. also reported that writing-to-learn interventions were less effective with middle school students and less effective when they involved longer writing tasks. Some have discussed how particular writing genres contribute to learning in different ways (e.g., Klein et al., 2019), while others have questioned the relative effectiveness of writing activities that can be applied across disciplines (e.g., journaling, notetaking) compared to discipline-specific writing activities (e.g., The Science Writing Heuristic; Hand, 2008). In the most recent meta-analysis of writing-to-learn interventions, Graham et al. (2020) found positive effects of writing-to-learn for students in grades 1–12; grade level, content area, and the type of writing activity did not moderate these effects.

In addition to research support, cognitive theories of writing frame how writing should contribute to learning. Cognitive models portray writing as a problem-solving activity during which the writer orchestrates many different mental processes to meet writing goals; many of these processes mirror those that occur during learning (Klein & Boscolo, 2016). As writers carry out the writing process and try to meet writing goals, they continually reorganize information and make connections between new and old ideas (Scardamalia & Bereiter, 1987; Flower & Hayes, 1981). Knowledge is constructed, or re-constructed, and learning occurs as writers interpret new information in relation to what they already understand (Boscolo & Mason, 2001). Additionally, the permanence of writing (as opposed to talk) permits writers to revisit and reflect upon their learning (Langer & Applebee, 1987).

Galbraith (1999) explained that a writer's emerging text helps them process and synthesize what to produce next; thus, learning occurs and is dependent upon the

successive generation of written content. Galbraith and Baaijen (2018) clarified that when a writer synthesizes information and produces text that extends or expands upon what is held in their episodic memory, they develop new knowledge, and this learning process is further shaped by the writer's focus on rhetorical goals while writing.

Extending beyond the cognitive processes involved in writing that may contribute to learning, Graham's (2018) Writer(s)-Within-Community Model of Writing is unique in that it blends both cognitive and sociocultural perspectives. Graham's model includes the cognitive aspects and actions of writers that likely contribute to learning, while also accounting for the community in which writing occurs, which also likely impacts writing's effects on learning. For example, the purpose and goals of a writing community must prioritize writing-to-learn; if a writing community focuses on other goals, then learning is less likely to be facilitated by writing (Graham et al., 2020). Also, a collective history of using writing-to-learn likely promotes learning because writers know how to participate in writing-to-learn practices. Success with writing-to-learn may also shape a writer's beliefs about its effectiveness and impact their motivation to use it in a particular setting or to apply it in other settings (Graham, 2018). Graham's model guided our creation of a coding scheme for the study we discuss in this chapter. We provide details about the coding process and rubric development in subsequent sections describing our study.

## 2 Our Work on Writing-to-Learn

### *2.1 The Larger Study*

Data analyzed for this chapter were taken from a larger intervention study (Gillespie Rouse & Rouse, 2019) that examined the impacts of a 10-lesson integrated STEM and writing-to-learn unit on third-grade students' learning. Participating students attended a private, non-religious elementary school in the Southern United States. The school's total enrollment was 440 and participating students were drawn from the three third-grade classrooms in the school.

In that study, we used stratified random assignment by pretest STEM vocabulary assessment to place students in a treatment or comparison condition within their intact classrooms. Students in both conditions participated in hands-on learning about STEM concepts and writing-to-learn activities for the first 45 min of each of the ten lessons in the unit. Treatment students responded in writing to two prompts during the final 15 min of each lesson. Comparison students discussed the same two prompts orally in a small group during this time. In the larger study, we found that students who participated in the unit made statistically significant gains from pre- to posttest on a STEM vocabulary assessment; total words written (TWW), number of different STEM concepts used (NC), and depth of understanding of STEM concepts (DC) in a written essay response; and number of different STEM concepts used in

an oral interview response, regardless of their incoming writing skills and regardless of whether they participated in additional writing or small-group discussion at the end of each unit lesson.

## *2.2 The Current Study*

The purpose of the current study was to examine students' general and STEM-specific writing knowledge after the integrated STEM and writing-to-learn unit as well as to analyze whether students' writing knowledge was predictive of their writing performance at posttest. For the current study, we examined a subset of participating students' oral interview responses from the larger study that we had not previously analyzed. Specifically, we examined students' responses to three writing knowledge questions about: (1) what good writers do when they write, (b) what good writers do when they write in science class, and (3) how writing in science class was different than writing in English/Language Arts (ELA) class. In addition, in the current study, we focused only on responses from the 26 students in the treatment condition, those students who participated in additional writing during each lesson in the larger study (Gillespie Rouse & Rouse, 2019). Sixty-nine percent of this subsample were female and 31% were male. The average age was 9 years (*SD* = 0.39). Students identified as: 61% White, 28% Asian/Indian, and 11% Hispanic.

We chose Graham's (2018) Writer(s)-Within-Community Model of Writing to develop a rubric for coding students' interview responses to the three writing knowledge questions because his model captured both the cognitive and social aspects of writing-to-learn present in our intervention. Specifically, Graham's description of the writing community includes the purposes for writing, which we established with students during our intervention (i.e., writing as a tool for documenting, reflecting upon, and explaining learning to others). Graham's model also captured the context for writing-to-learn in our intervention. Students completed the unit in small learning groups designed to support them in both their STEM and writing-to-learn efforts; this aligned with Graham's theories about feedback and support promoted in writing communities with collaborators. We also included codes on our rubric to capture Graham's depiction of the cognitive mechanisms involved in writing, including long-term memory resources (i.e., knowledge and beliefs) and production processes (i.e., conceptualization, ideation, translation, transcription, and reconceptualization), as well as modulators of those resources and processes (i.e., emotions, personality traits, and physical state). This was important, as students' typical classroom writing instruction focused on the writing process (captured by Graham's production processes) and our intervention focused on increasing students' knowledge (i.e., long-term memory resources) of how to use writing to support their learning in STEM.

## Research Questions and Hypotheses

The current study was guided by two main research questions. We asked:

1. After participating in an integrated STEM and writing-to-learn unit, what do students know about good writing and about writing in different subject areas?
2. Does students' writing knowledge after the integrated STEM and writing-to-learn unit predict their posttest written essay performance?

For research question 1, we had different hypotheses for the three different writing knowledge questions we asked students at posttest. Because the students' school used the process approach to writing instruction, which emphasizes the procedures involved in planning, drafting, revising, editing, and publishing texts, we anticipated responses to writing knowledge question 1 (What do good writers do when they write?) would primarily focus on production processes; this trend has been demonstrated in previous studies of elementary students' writing knowledge (e.g., Gillespie et al., 2013). We hypothesized that students would show understanding of how writing can be used in science class (writing knowledge question 2) and the distinction between writing in science class and writing in English/language arts (ELA) (writing knowledge question 3), as our intervention involved daily use of writing to demonstrate and document learning of STEM concepts and students were taught a mnemonic for including STEM-specific information in their writing. Thus, we anticipated more idea units that involved STEM and ELA content or genres for writing knowledge questions 2 and 3.

For research question 2, we hypothesized that students' knowledge of the operations or action involved in the writing process (coded as production processes) and their knowledge of how writing can be used in different subject areas (coded as long-term memory resources) would predict variance in the length of their expository essays at posttest. We anticipated students' knowledge of how writing can be used in different subject areas (i.e., long-term memory resources) would be a predictor of the number of different STEM concepts included in posttest essays and the depth of understanding of those STEM concepts.

## Procedures in Intervention

Two researchers delivered the 10-lesson STEM and writing-to-learn unit to students during their typical literacy block, following procedural checklists to ensure each lesson was delivered as planned and delivered in the same way for each class. Each lesson lasted 1 h, for a total of 10 h of instruction.

During the unit, students learned STEM concepts in the context of planning a five-page pop-up book. After examining features of commercial pop-up books, students reverse engineered pop-ups from commercial books into transparent books to examine their movement, solved design challenges involving different pop-up folds, and designed prototype pop-ups. Throughout the unit, students also developed working definitions for 12 key STEM concepts (i.e., asymmetric, angle fold,

engineering design process, gutter, optimize, page positions, parallel fold, prototype, reverse engineering, symmetric, trade off, and troubleshooting). After these learning activities, students designed and built five pop-ups for their own books.

Students also used writing throughout the unit to document, support, and extend their learning. In each lesson, students used their Sketchbooks (i.e., a bound notebook with daily writing activities) to record informal written responses to questions aligned with learning activities and to sketch how they solved design challenges. For example, in lesson 2, students were asked to write down issues they experienced when reverse engineering a commercial pop-up into their transparent books and to write down how they troubleshot to fix those issues. In lesson 2, students were also asked to sketch how their pop-up worked and moved within their transparent book. The Sketchbook had additional space for students to record ideas and observations they made during unit lessons.

At the end of each unit lesson, students spent 15 min writing responses to 2–3 prompts (see Table 1) aligned with the learning activities for the day. For example, in lesson 8, students responded in writing to: What are the differences between symmetric and asymmetric parallel-fold pop-ups? Choose one (asymmetric or symmetric parallel-fold pop-up) and write directions for someone to make it. The first author taught students the mnemonic ADDS (i.e., Answer the question using Data Details and Science and engineering vocabulary) during lesson 1 to help students plan and write their responses. During lessons 1 and 2, she modeled how to use ADDS and produced a collaborative response with students to the first prompt for each lesson, while students answered the remaining prompts on their own. After lesson 2, students responded to all writing prompts independently, with the first author reminding them to check ADDS after writing to make sure they had included relevant information in their responses.

Students also produced the text for their five-page pop-up book during the unit. Classroom teachers identified figures of speech (i.e., simile, metaphor, hyperbole, personification, and idiom) as a relevant topic for the book based on what students had recently learned in ELA class. Students completed a graphic organizer to write the text intended for each page of their pop-up book and to sketch their pop-up and background illustrations for each page. For the final pop-up book, students transferred the text they planned and the pop-ups they created into Chunky Bare Books; these books resembled commercial pop-up books because they had thick pages and were made of rigid cardboard.

Although each student produced their own pop-up book and their own written responses during each session, all work during the unit occurred in small groups. Thus, learning and writing were both accomplished with the help of peers in the small group. Students solved design challenges in their groups, discussed ideas for writing in their groups, and relied on peers for feedback and support with their pop-up book text and designs.

**Table 1** Journal prompts for each day of unit instruction

| Lesson: topic | Journal prompts |
|---|---|
| 1: Examining features of a pop-up book | What parts do pop ups need to have to work?<br>How were the pop ups in each book you looked at today the same or different? |
| 2: Reverse engineering parallel-fold pop ups | What does it mean to troubleshoot?<br>What does it mean to reverse engineer a pop up?<br>How were the pop ups you reverse engineered today the same and how were they different (think of how they move and where they were attached to the book)? |
| 3: Parallel-fold pop-up design challenge I | What is a parallel-fold pop up?<br>How does a parallel-fold pop up move when you open and close the book? |
| 4: Parallel-fold pop-up design challenge II | What are the differences between symmetric and asymmetric parallel-fold pop ups?<br>Choose one (asymmetric or symmetric parallel-fold pop up). Write directions for someone to make it. |
| 5: Designing a parallel-fold pop-up prototype | What is a prototype? Why do we make prototype pop ups?<br>What does it mean to optimize a pop up? |
| 6: Reverse engineering angle-fold pop ups | What problems did you have when reverse engineering your pop ups? How did you troubleshoot these problems?<br>How were the pop ups you reverse engineered today the same and how were they different (think of how they move and where they were attached to the book)? |
| 7: Angle-fold pop-up design challenges | What is an angle-fold pop up?<br>How does an angle-fold pop up move when you open and close the book? |
| 8: Designing an angle-fold pop-up prototype | What are the differences between symmetric and asymmetric angle-fold pop ups?<br>Choose one (asymmetric or symmetric angle-fold pop up). Write directions for someone to make it. |
| 9: Authoring and designing a pop-up book | What problems did you encounter while building your pop-up book? How did you troubleshoot these problems? |
| 10: Authoring and designing a pop-up book | What trade-offs did you consider when designing the pops up in your book?<br>Describe the engineering design process for building the pop ups in your book. |

## Measures

Although multiple measures were administered to the entire sample in the larger study (Gillespie Rouse & Rouse, 2019), for the purposes of this chapter, we focused on students' written essay responses at posttest and their posttest videorecorded oral interview responses about writing knowledge and use of writing in different subject areas (i.e., writing knowledge questions 1–3 that had not been previously analyzed). Students completed the written essay response and interview 1 week after the unit was completed. We also used students' scores on the Essay Composition subtest of

the Wechsler Individual Achievement Test-Third Edition (WIAT-III), administered prior to intervention, to account for students' incoming writing abilities in our analyses.

**Written Essay Response** For the written essay response, the first author read aloud the following prompt: What do you know about pop-ups? How do you build them? How do they work? Students were told they could have as much time as needed to write their responses, to write as if explaining to someone who did not know anything about pop-ups, and not to worry about spelling.

All written essay responses were typed into Microsoft Word, with corrections for spelling or grammar errors, as handwriting legibility as well as spelling and grammar mistakes have been shown to negatively affect raters' scores of writing content (Graham, 1999; Graham et al., 2011). The researchers double entered 30% of the responses for reliability purposes. Interrater reliability for typed essay responses was 98% exact agreement, with all disagreements resolved by consensus.

We scored essay responses for: TWW, NC, and DC. We calculated TWW using the word count feature of Microsoft Word; interrater reliability was 100% exact agreement with all responses double coded for TWW. The score for NC (i.e., the 12 STEM and pop-ups concepts students developed working definitions for during the unit) ranged from 0 to 12, as each concept used was only counted once, even if used multiple times in the essay. Importantly, each concept had to be used and defined correctly to be counted in this score. We double coded all responses for NC and interrater reliability was 99% exact agreement, with all disagreements resolved by consensus.

The researchers scored DC holistically using writing anchors. Holistic scoring involves rating a writing sample on a numerical scale based on an overall impression of its quality (as opposed to scoring separately for specific features) (Myers, 1980). The writing anchors we chose came from our own sample. Thus, rather than impose a scoring system developed for another type of writing or another type of writing prompt, the anchors and rubric were grounded in the actual writing samples produced in our study.

To develop the writing anchors to serve as examples or benchmarks for each level (0, 1, 2, and 3) on our scoring rubric, the researchers independently ordered a random sample of students' pre- and posttest written essay responses from lowest to highest, based on depth of understanding of STEM concepts. Next, we compared our rank-ordered responses and resolved any differences. From our agreed upon rank ordering, we decided on two anchor papers each to represent no (score = 0), minimal (score = 1), intermediate (score = 2), and deep (score = 3) understanding of STEM concepts. Using these anchors, the researchers each scored 50% of the written essay responses and double scored 30% for reliability. Inter-rater reliability was 94% exact agreement with all disagreements resolved by consensus.

**Writing Knowledge Interview** For the writing knowledge interview, an examiner met individually with each student to videorecord their oral responses to three questions; the first question was replicated from previous studies of elementary students'

writing knowledge (Gillespie et al., 2013; Graham et al., 1993; Olinghouse et al., 2015) while the remaining questions were developed specifically for this study: (1) What do good writers do when they write? (2) What do good writers do when they write in science class?, and (3) How does writing in science class differ from writing in literacy or ELA class?

Individuals unfamiliar with the purposes of our study transcribed all video-recorded interviews. Then, the second and third authors parsed each interview response into individual idea units. We defined an idea unit as previous writing researchers had: each unique and specific idea in a student's response (Gillespie et al., 2013; Olinghouse & Graham, 2009). For example, "They try to write a short story and they try to get every single detail in it" was counted as two idea units (idea unit 1: write a short story and idea unit 2: get every single detail in it). Idea units that were repeated were only counted once and idea units that elaborated on another idea unit but did not add new or unique information were only counted once. For example, "In science class, you make a lot of drafts, like two or three drafts. You make many drafts," was counted as one idea unit referencing making drafts. The second and third authors double coded one-third of students' responses for reliability of parsing idea units. Interrater reliability for idea unit coding was calculated using the bivariate correlation between their two scores for total idea units for each double-coded response. Reliability was 0.92. All disagreements were discussed and resolved by consensus.

**Rubric for Categorizing Idea Units** To develop a rubric for categorizing the idea units in students' responses, we consulted Graham's (2018) Writers-within-Community Model described in our chapter introduction. We coded one-third of students' responses with our first rubric that contained codes for the cognitive aspects of writing outlined in Graham's (2018) model as well as aspects of students' writing community. From this initial coding, we added or modified codes as necessary to capture what was present in students' responses in a revised rubric. Importantly, some aspects of Graham's model were combined into one code on our rubric (e.g., translation and transcription) while we split ideation into two codes (i.e., internal and external ideation). We also added a code for writing and drafting texts that differed from translation and transcription. These decisions were based on discussions between the three authors related to the types of idea units produced in our sample.

Because our interview questions focused on what good writers do, or the actions and behaviors of good writers, many of the codes on our final rubric incorporated aspects of Graham's (2018) model related to production processes, or "the mental and physical operations writers apply to produce text," (Graham, 2018, p. 269). Our production processes codes were: (a) conceptualization, (e.g., use your imagination, make pictures in your mind, plan it out in their head); (b) internal ideation (e.g., pick a topic, think of all of the facts they know); (c) external ideation (e.g., look at more than one source for information, research; get ideas from other stories); (d) translation and transcription (e.g., put it into their own words, spelling, capitalization); (e)

writing and drafting (e.g., write down the details of it, make drafts); and (f) reconceptualization (e.g., go back and fix it to make it better, make sure it makes sense, edit their work). We also included codes related to the writer's community and long-term memory resources (Graham, 2018). Our community codes were: (a) audience/purpose (e.g., invite the reader into the book, entertain, inform); and (b) seeking assistance (e.g., ask for advice from other people, ask a science teacher). Our long-term memory resources codes were: (a) ELA content or genres (e.g., writing stories, similes, plot, fantasy), and (b) STEM content or genres (e.g., draw graphs, write number stories in math, do experiments, collect data). The rubric also included a code for idea units that were related to the interview questions but not captured in other categories (e.g., they're different subjects, make sure the information is from that year, use ADDS) and a code for unrelated idea units (e.g., put forward the total time it took). Codes for modulators (e.g., emotions, physical state) were dropped after coding our final dataset, as no students produced responses in these categories.

The second and third author coded all idea units within students' interview responses using the final rubric, double coding one-third of responses for reliability. Percent exact agreement for coding idea units ranged from 82% (only one response) to 100% across the double-coded responses, with an average of 91% exact agreement. All coding disagreements were discussed and resolved by consensus.

## Data Analysis

To answer research question 1 (After participating in the integrated STEM and writing-to-learn unit, what do students know about good writing and about writing in different subject areas?) we calculated means and standard deviations (see Table 2), and we discuss percentages of the types of responses students gave to the writing knowledge interview questions at posttest.

To answer research question 2, we conducted three hierarchical regression analyses, one with each of the posttest essay measures (i.e., TWW, NC, and DC) as our dependent variable. To generate our independent, or predictor, variables, we combined students' production processes responses across all questions (i.e., total production processes) and combined students' long-term memory resources responses (i.e., total long-term memory responses) across all questions. We decided against using responses coded as community, as these represented only 3% of the responses students produced across writing knowledge questions.

Each regression analysis consisted of three steps to analyze the impact of predictor variables. For each analysis, we controlled for students' incoming writing skills (as measured on the WIAT-III Essay Composition subtest) and gender, entering this as our first block of predictor variables. For step 2, we entered total production processes, followed by total long-term memory responses in the third block of predictor variables.

**Table 2** Categorization of idea units in students' oral interview responses

| | Question 1 | | Question 2 | | Question 3 | |
|---|---|---|---|---|---|---|
| Response type | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| Community | 0.31 | (0.84) | 0.04 | (0.20) | 0.00 | (0.00) |
| Audience/purpose | 0.12 | (0.33) | 0.04 | (0.20) | 0.00 | (0.00) |
| Seeking assistance | 0.19 | (0.80) | 0.00 | (0.00) | 0.00 | (0.00) |
| Long-term memory resources | 0.69 | (1.16) | 0.88 | (1.03) | 3.65 | (1.81) |
| ELA content/genres | 0.35 | (0.63) | 0.00 | (0.00) | 1.96 | (1.37) |
| STEM content/genres | 0.35 | (1.09) | 0.88 | (0.99) | 1.69 | (0.93) |
| Production processes | 3.15 | (1.74) | 2.12 | (1.28) | 0.54 | (1.30) |
| Conceptualization | 0.38 | (0.64) | 0.23 | (0.71) | 0.00 | (0.00) |
| Internal ideation | 0.15 | (0.37) | 0.12 | (0.33) | 0.04 | (1.96) |
| External ideation | 0.50 | (0.91) | 0.50 | (0.86) | 0.04 | (1.96) |
| Translation and transcription | 0.58 | (0.95) | 0.35 | (0.75) | 0.31 | (1.19) |
| Writing and drafting | 1.12 | (1.07) | 0.88 | (0.99) | 0.15 | (0.37) |
| Reconceptualization | 0.42 | (0.95) | 0.04 | (0.20) | 0.00 | (0.00) |
| Related other | 0.58 | (1.07) | 0.85 | (0.97) | 0.35 | (0.80) |
| Unrelated other | 0.00 | (0.00) | 0.00 | (0.00) | 0.08 | (0.39) |
| Total idea units | 4.73 | (2.34) | 3.88 | (2.14) | 4.62 | (2.42) |

# 3 Results

## *3.1 Research Question 1*

Across all three writing knowledge posttest questions, students produced a total of 344 unique idea units ($M$ = 13.23, $SD$ = 5.37, range = 6 to 31 total idea units per student). The means and standard deviations for types of responses produced for each writing knowledge question are presented in Table 2.

### Writing Knowledge Question 1: What Do Good Writers Do When They Write?

On average, students produced 4.73 idea units ($SD$ = 2.34) when responding to writing knowledge question 1. A majority of students' responses (67%) involved production processes. Of these, 35% of responses were related to what good writers do when they write and draft their texts (e.g., you need to add detail, write a rough draft). The remaining production processes responses referenced translation and transcription (19%; e.g., they write neatly, usually in cursive); external ideation (16%, e.g., they look at other books, they get ideas from websites); reconceptualization (13%, e.g., edit their work, if they see something is not right they cross it out and do it over again); conceptualization (12%; e.g., think about what they're going to write, they usually plan it out in their head), and internal ideation (5%, e.g., pick a topic to write about, come up with the main idea).

Approximately 15% of students' responses to writing knowledge question 1 involved long-term memory resources, with half of those responses about ELA content or genres (e.g., write a short story, use figurative language) and the other half about STEM content or genres (e.g., they answer the question using data, they use science and engineering vocabulary). The remaining responses focused on community (7%), including audience or purpose (e.g., write so the reader understands it) and seeking assistance (e.g., ask for help), or they were responses coded as related other (11%; e.g., there's also an illustrator, they use ADDS).

### Writing Knowledge Question 2: What Do Good Writers Do When They Write in Science Class?

Students produced about 3.88 idea units (*SD* = 2.14) when responding to writing knowledge question 2. When asked about what good writers do when writing in science class, students mostly referenced production processes (54% of responses), just as they had done for writing knowledge question 1. Forty-two percent of the references to production processes involved writing and drafting texts (e.g., they make drafts, write what they think is a good idea). The remaining production processes responses focused on external ideation (24%; e.g., they research their topic, they might have to actually go learn more about it), translation and transcription (16%; e.g., they capitalize, they don't really care about spelling), conceptualization (11%; e.g., they plan it out on the computer, most writers plan ahead), internal ideation (5%; e.g., they have to know about science so they know what to write about, they think of everything they have learned in science), and reconceptualization (2%; e.g., go back and make sure stuff is very accurate).

Students also referenced long-term memory resources in about 23% of their responses to writing knowledge question 2. It was not surprising that all of these references involved STEM content or genres (e.g., they put in science vocabulary, in science they keep a notebook with all of their data), as writing knowledge question 2 specifically related to science. There was only one response (1%) related to community (e.g., describe it well because some people may not have seen it). The remaining responses to writing knowledge question 2 were coded as related other (22%; e.g., you could take a picture for your book, ADDS starts with answering the question).

### Writing Knowledge Question 3: How Does Writing in Science Class Differ from Writing in Literacy or ELA Class?

Students averaged 4.62 idea units (*SD* = 2.42) when responding to writing knowledge question 3. When describing the differences in writing for science class compared to literacy class, students mostly focused on long-term memory resources (79% of responses). These references to long-term memory resources were fairly evenly split, with 54% involving ELA content or genres (e.g., in literacy we write

about our own adventures, we write stories, use personification) and 46% involving STEM content or genres (e.g., in science we write about the moon and solar system, a chapter on where energy comes from, in science it has to all be true).

Twelve percent of students' responses to writing knowledge question 3 involved production processes, with more than half of those responses referencing translation and transcription (57%; e.g., use pronouns, put more punctuation). The remaining production processes responses focused on writing and drafting texts (29%; e.g., write about different things, writing freely on the page) and ideation, with an equal number responses (7% each) for external ideation (e.g., we don't usually look at other books) and internal ideation (e.g., use the ideas you already know). No responses to writing knowledge question 3 involved community and less than 10% were coded as related other (8%; e.g., that's the one big difference between them) or unrelated other (1%; e.g., put forward the total time it took).

## *3.2 Research Question 2*

To examine if students' writing knowledge, specifically their knowledge of production processes and knowledge related to long-term memory resources (i.e., ELA content or genres and STEM content or genres) made significant contributions to predicting their performance on a posttest essay, we first entered gender and WIAT-III Essay Composition subtest scores into our regression models.

On average students scored an 8.67 ($SD$ = 2.45) on the WIAT-III. Our predictor variables were total production processes responses across questions 1–3 ($M$ = 5.81, $SD$ = 2.53) and total long-term memory resources responses across questions 1–3 ($M$ = 5.23, $SD$ = 3.19). On posttest essay measures, used as our dependent variables, students produced about 194 TWW ($SD$ = 93.97; range 50 to 392 TWW), with, on average, 4.62 ($SD$ = 2.91; range 0 to 12) different STEM concepts in their essays (i.e., NC). Students had a mean score of 1.42 ($SD$ = 1.09) for posttest depth of understanding of STEM concepts (i.e., DC) on a rubric from 0 to 3.

Table 3 shows correlations among the regression variables. There were statistically significant correlations between total production processes and all three posttest essay outcomes, our dependent variables in the regression models, TWW, NC, and DC (all $p < 0.05$). The three dependent variables were also all significantly correlated with each other (all $p < 0.01$).

### Regression 1: Total Words Written (TWW)

Table 4 displays the results from our regression analyses. For TWW, the first step with only our control variables, gender and WIAT-III score, was not a statistically significant model, $F(2, 25) = 2.02$, $p = 0.16$. With the addition of total production processes responses in step 2, the regression model was statistically significant, $F(3, 25) = 5.02$, $p < 0.01$, and both gender ($p < 0.05$) and total production processes

**Table 3** Correlations among regression variables

| Variable | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1. Gender | – | | | | | | |
| 2. WIAT-III | 0.31 | – | | | | | |
| 3. Total PP | −0.15 | −0.06 | – | | | | |
| 4. Total LTM | −0.06 | 0.17 | −0.04 | – | | | |
| 5. TWW | 0.38 | 0.18 | 0.44* | 0.21 | – | | |
| 6. NC | 0.23 | −0.03 | 0.46* | 0.09 | 0.71** | – | |
| 7. DC | 0.30 | 0.09 | 0.43* | 0.14 | 0.80** | 0.91** | – |

*Note*. WIAT-III = score on the Essay Composition subtest of the Wechsler Individual Achievement Test-Third Edition; *PP* production processes responses, *LTM* long-term memory resources responses, *TWW* total words written on posttest essay, *NC* number of different STEM concepts included in posttest essay, *DC* depth of understanding of STEM concepts in posttest essay
*$p < 0.05$, **$p < 0.01$

responses ($p < 0.01$) were statistically significant predictors of TWW at posttest. $R^2$ change from step 1 to step 2 was 0.26. With the addition of total long-term memory resources responses in step 3, the regression model remained statistically significant, $F(4, 25) = 4.63$, $p < 0.01$, with gender ($p < 0.05$) and total production processes responses ($p < 0.01$) remaining as statistically significant predictors of TWW at posttest. $R^2$ change from step 2 to step 3 was 0.06. The third model explained 47% of the variance in TWW at posttest with all predictors in the model; however, total long-term memory resources was not a statistically significant predictor of TWW at posttest as we had hypothesized.

### Regression 2: Number of Different STEM Concepts (NC)

The control variables, gender and WIAT-III score, were not statistically significant predictors of NC at posttest in any of the steps for this regression analysis. The addition of total production processes responses in step 2, resulted in a statistically significant increase in variance explained, $F(3, 25) = 3.31$, $p < 0.05$, and total production processes responses ($p < 0.05$) was a statistically significant predictor of NC at posttest in the regression model. $R^2$ change from step 1 to step 2 was 0.24. With the addition of total long-term memory resources responses in step 3, the regression model was no longer statistically significant, $F(4, 25) = 2.64$, $p = 0.06$, although total production processes responses ($p < 0.05$) remained a statistically significant predictor in the model. $R^2$ change from step 2 to step 3 was 0.02. The third model explained 33% of the variance in NC at posttest with all predictors in the model. As with TWW, total long-term memory resources was not a statistically significant predictor of NC at posttest as we had hypothesized.

**Table 4** Regression results for predicting posttest essay outcomes from writing knowledge

| Step | Variable | *B* | SE *B* | *β* | *t* | $R^2$ | F |
|---|---|---|---|---|---|---|---|
| TWW | | | | | | | |
| 1 | Gender | 71.34 | 40.37 | .357 | 1.77 | 0.15 | 2.02 |
| | WIAT-III | 2.82 | 7.77 | 0.07 | 0.36 | | |
| 2 | Gender | 86.64 | 34.84 | 0.43 | 2.49* | 0.41 | 5.02** |
| | WIAT-III | 3.01 | 6.64 | 0.08 | 0.45 | | |
| | Total PP | 19.05 | 6.17 | 0.51 | 3.09** | | |
| 3 | Gender | 93.26 | 34.00 | 0.47 | 2.74* | 0.47 | 4.63** |
| | WIAT-III | 0.93 | 6.56 | 0.02 | 0.14 | | |
| | Total PP | 19.49 | 5.99 | 0.53 | 3.26** | | |
| | Total LTM | 7.52 | 4.80 | 0.26 | 1.57 | | |
| NC | | | | | | | |
| 1 | Gender | 1.64 | 1.31 | 0.27 | 1.25 | 0.07 | 0.80 |
| | WIAT-III | −0.13 | 0.25 | −0.11 | −0.53 | | |
| 2 | Gender | 2.11 | 1.16 | 0.34 | 1.81 | 0.31 | 3.31* |
| | WIAT-III | −0.13 | 0.22 | −0.11 | −0.57 | | |
| | Total PP | 0.58 | 0.21 | 0.50 | 2.81* | | |
| 3 | Gender | 2.23 | 1.18 | 0.36 | 1.89 | 0.33 | 2.64 |
| | WIAT-III | −0.17 | 0.23 | −0.14 | −0.73 | | |
| | Total PP | 0.59 | 0.21 | 0.51 | 2.83* | | |
| | Total LTM | 0.14 | 0.17 | 0.16 | 0.86 | | |
| DC | | | | | | | |
| 1 | Gender | 0.71 | 0.49 | 0.30 | 1.46 | 0.09 | 1.16 |
| | WIAT-III | −0.00 | 0.09 | −0.01 | −0.03 | | |
| 2 | Gender | 0.88 | 0.43 | 0.38 | 2.02 | 0.32 | 3.48* |
| | WIAT-III | −0.00 | 0.08 | −0.00 | −0.01 | | |
| | Total PP | 0.21 | 0.08 | 0.49 | 2.74* | | |
| 3 | Gender | 0.93 | 0.44 | 0.40 | 2.15* | 0.36 | 2.92* |
| | WIAT-III | −0.02 | 0.08 | −0.04 | −0.23 | | |
| | Total PP | 0.21 | 0.08 | 0.50 | 2.80* | | |
| | Total LTM | 0.07 | 0.06 | 0.19 | 1.08 | | |

*Note. TWW* total words written on posttest essay, *WIAT-III* score on the Essay Composition subtest of the Wechsler Individual Achievement Test-Third Edition, *Total PP* production processes responses across questions 1–3, *Total LTM* total long-term memory resources responses across questions 1–3, *NC* number of different STEM concepts included in posttest essay, *DC* depth of understanding of STEM concepts in posttest essay

*$p < 0.05$, **$p < 0.01$

### Regression 3: Depth of Understanding of STEM Concepts (DC)

For DC at posttest, in the first step with only our control variables (i.e., gender and WIAT-III score), the model was not statistically significant, $F(2, 25) = 1.16$, $p = 0.33$. The addition of total production processes responses in step 2, resulted in a statistically significant increase in variance explained by the model, $F(3, 25) = 3.48$,

$p < 0.05$, and total production processes responses ($p < 0.05$) was a statistically significant predictor of DC at posttest. $R^2$ change from step 1 to step 2 was 0.23. With the addition of total long-term memory resources responses in step 3, the regression model remained statistically significant, $F(4, 25) = 2.92$, $p < 0.05$, with both gender and total production processes responses (both $p < 0.05$) as statistically significant predictors of DC at posttest. $R^2$ change from step 2 to step 3 was 0.04. The third model explained 36% of the variance in DC at posttest with all predictors in the model; however, as with the other posttest outcomes, total long-term memory resources was not a statistically significant predictor of DC at posttest as we had hypothesized.

## 4 Discussion

Third-grade students in our study answered three writing knowledge interview questions and wrote posttest essays to demonstrate their knowledge of STEM concepts after participating in a 10-day integrated STEM and writing-to-learn unit. We found students expressed different types of writing knowledge depending on the interview question. We also found that students' writing knowledge was predictive of their writing performance on the posttest essay.

### *4.1 What Do Students Know About Good Writing and About Writing in Different Subject Areas?*

Aligned with our hypothesis for research question 1, a majority of students' responses to the first writing knowledge interview question (i.e., What do good writers do when they write?) related to the production processes involved in writing. This finding was consistent with previous research using the same writing knowledge question with upper-elementary and middle school students (Gillespie et al., 2013; Graham et al., 1993). We were not surprised that more than one-third of students' production processes responses related to writing or drafting texts, as students' classroom instruction involved considerable time for writing each day and instruction in how to carry out the writing process.

The remaining production processes responses covered the range of processes involved in skilled writing and captured by our rubric, but students' responses weren't particularly refined in terms of referencing how these processes were integrated or recursive. Students tended to discuss a linear process of writing (e.g., Good writers come up with an idea, write it down, edit it over and over again, and then put it in a real book) or discuss only one or two components of the writing process (e.g., they search for ideas, get lots of ideas, and then write them down and keep writing until they get all of the details down). Future research and instruction

could focus on teaching students about the non-linear enactment of the writing process, with modeling of how skilled writers participate in multiple aspects of the writing process simultaneously and frequently move out of the typical sequence in which the writing process is taught (e.g., a writer may be simultaneously writing and editing or revising a text or they may revisit plans or ideas for writing while they are drafting the text).

Although not part of our hypothesis, a chief reason we used Graham's (2018) Writer(s)-Within Community Model of Writing to develop the rubric for categorizing students' interview responses was because of its emphasis on the writing community and collaborators. This focus on social aspects of writing not only aligned with our own personal beliefs and experiences teaching writing, but it also aligned with the ways students participated in the activities in our intervention, in small, collaborative learning groups. Because of this, it was interesting that the smallest number of students' responses to the first writing knowledge question, less than 10%, involved community. Few students mentioned that good writers think of the audiences for their writing and few discussed that good writers seek assistance from those in their writing community.

Reflecting on this finding, we plan to directly discuss the writing community as well as audience and/or purpose for writing in future iterations of our intervention. Although we posed the essay prompt for students to write to someone who did not know much about the topic, and we encouraged small-group collaboration on STEM and writing tasks, students would benefit from direct modeling and guidance on how to consider their audience when writing. Posttest essays, on average, were rated as demonstrating minimal depth of understanding of STEM concepts. Yet, in the larger study, we found that students could match most STEM concepts to their definitions and explain most of them correctly during oral interview responses at posttest. A focus on thinking about audience for writing may help students better demonstrate their understanding of STEM concepts in their posttest writing so that it is clear and easily understood by a hypothetical reader. Direct instruction in how to participate in and collaborate with peers in one's writing community may also encourage students to rely on collaborators for writing feedback and assistance as they compose. We anticipate this focus on the writing community will result in stronger posttest essay responses that are more closely aligned with students' STEM knowledge, but, of course, these propositions will need to be tested directly in future research.

For writing knowledge questions 2 and 3 (i.e., What do good writers do when they write in science class? and How does writing in science class differ from writing in literacy or ELA class?) we anticipated students' responses would be predominantly about STEM and ELA content or genres, which we coded as long-term memory resources. This was not the case for writing knowledge question 2. Only about one-fourth (23%) of students' responses to this question involved STEM content or genres (e.g., they do an experiment and write how close their prediction was to the actual thing, answer the question using data). More than half of students' responses to writing knowledge question 2 focused on production processes, with most of these responses involving writing or drafting texts. Thus, students' responses

to writing knowledge question 2 followed a similar pattern to their responses to writing knowledge question 1.

Students' failure to articulate many disciplinary-specific practices for science writing for writing knowledge question 2 may have been due to our intervention and question wording. Our intervention did not solely focus on science, but rather we used the term STEM to capture the engineering and mathematics concepts (technology was less of a focus) that were also emphasized throughout the unit lessons. Other researchers have demonstrated students' abilities to learn and use the writing practices specific to the scientific discipline (e.g., Hand, 2017), but STEM writing incorporates practices from multiple disciplines. We are unsure that this distinction truly impacts our work, as our use of writing-to-learn also involved types of writing that could be applied across disciplines; thus, our participants may not have viewed what they were doing as "science" writing, or "STEM" writing for that matter. Future work could explore the practices that are specific to the multiple disciplines in STEM and teach students to understand distinctions between the ways in which scientists, for example, may use writing differently from engineers.

Students' responses for writing knowledge question 3 did align with our expectations. More than three-fourths of students' responses referenced long-term memory resources, split between idea units related to ELA content or genres (54%) and STEM content or genres (46%). Thus, although students did not seem to articulate practices for science writing for writing knowledge question 2, they were better able to contrast writing in science class with writing in ELA. Perhaps the question wording facilitated students' responses, as they were asked to directly contrast writing in the two subjects and knowledge of writing in ELA may have prompted recall of additional knowledge of writing in science. For example, when a student stated, "In literacy you can really put your imagination into it", they contrasted with the following ideas, "but in science it all has to be true and they usually give you a specific topic to write about."

Of importance, students' conceptions of what was involved in writing in ELA and writing in science were not sophisticated. Students often referenced particular topics in literacy that were the focus of their figurative language unit and the topic for their pop-up book (e.g., personification, simile) and particular topics they had studied in science class (e.g., Mars, energy). They also frequently talked about differences in fiction and non-fiction writing, but no students identified that both fiction and non-fiction could be used in ELA and both (although perhaps less commonly with fiction) could be used in science. There was a tendency for students to state that one could only write fiction in ELA and only write non-fiction in science class. There was little to no discussion of text structures or disciplinary-specific writing practices and there were zero references to audience or purpose for writing knowledge question 3.

Research shows that elementary students' discourse knowledge, which includes their knowledge about how to write, how to write in specific genres, and how to enact schemas and procedures for specific writing tasks, is a significant predictor (beyond their topic knowledge) of their writing quality in three commonly-taught elementary genres: narrative, persuasive, and informational (Olinghouse et al.,

2015). Thus, future work could explore ways to bolster students' understanding of the distinctions in the genres and text structures commonly used in ELA compared to those commonly used in science, or more broadly, in STEM. It would be interesting to discover if direct instruction in these differences impacted students' abilities to express their STEM knowledge in writing at posttest. We assume with a greater focus on helping students develop a deeper understanding of what is included in a scientific explanation, like we asked them to do for our posttest essay prompt, that students would produce higher quality written responses.

### *4.2 Does Students' Writing Knowledge Predict Their Posttest Written Essay Performance?*

Results for research question 2 were somewhat aligned with our hypotheses and somewhat unexpected. We anticipated total production processes would be a statistically significant predictor of TWW (and it was) but did not anticipate it would also predict NC and DC because we anticipated these measures would be predicted by long-term memory resources, or content-specific knowledge (but they were not). Upon further consideration and reflection on Graham's (2018) model of writing, knowledge of mental and physical aspects or processes involved in text production (production processes) is presumed to draw upon a writer's long-term memory resources about a topic. Perhaps then, students with greater knowledge of what is involved in producing a text (production processes) knew to include more topic-specific, or STEM-specific, concepts in their writing and to explain those concepts in greater detail to create a text that met the demands of the writing prompt and the intended audience. As with other conclusions in this chapter, this would need to be directly tested in future studies perhaps with more specific questions about writing knowledge related to long-term memory resources or with think-aloud procedures to ascertain students' intentions while writing, specifically why they include and describe content-specific concepts in their essay responses.

Total long-term memory resources was not a statistically significant predictor of any of the dependent measures, which did not align with our hypotheses. We assumed that students who referenced more STEM and ELA content or genres in their interview responses would better understand what is expected and involved in science writing and thus include more STEM concepts (NC) and describe those concepts in more depth (DC) in their posttest essays. As we mentioned with research question 1, students did not have sophisticated conceptions of disciplinary-specific writing practices for science. Thus, our interview response code for long-term memory resources included more idea units about possible writing topics in science (e.g., Mars, energy) and less about genre-specific differences in science or STEM writing compared to ELA. In future research, a greater focus on disciplinary writing (students only practiced using our ADDS mnemonic, which focused on writing scientific explanations, for 25% of intervention) or development of a strategy for writing

STEM responses may prove more effective in not only developing students' understanding of discipline-specific writing differences but also helping them create texts that include discipline-specific elements.

### 4.3 Limitations

There are several limitations that impact interpretation of our results. First, our sample was relatively small ($n = 26$), which limits the generalizability of our findings. Our small sample size also impacted our regression analyses; with greater power, we may have discovered additional variables that predicted variance in posttest essays. Second, it is possible that young students, like the third graders in our study, are unable to articulate what they know about good writing and writing in different subject areas. Thus, our interview questions may not have captured students' full understandings. More specific follow-up questions to students' interview responses or the use of think-aloud prompts during writing may have better captured students' writing knowledge and how they applied that knowledge. Our third limitation is also related to our writing knowledge interview questions. As we explained, our questions could have been worded more precisely to reflect the interdisciplinary nature of the STEM intervention we implemented. It is quite possible our students did not make the connection that the writing they were doing was related to "writing in science", which was the terminology we used in our interview questions.

### 4.4 The Future: Investing in Writing as a Tool for Learning in STEM

We believe in and have demonstrated the power of writing-to-learn for students to develop greater understanding of STEM concepts. Future research is critically important to test underlying assumptions of writing-to-learn and to develop instructional practices that allow students to carry out writing-to-learn activities in STEM in ways that maximally impact their learning. STEM is at the forefront of educational initiatives, as our nation tries to prepare a STEM-literate workforce (National Science and Technology Council, 2018) that includes individuals who can engage in the specialized ways scientists and engineers talk and write (National Research Council, 2012). To meet these calls to action and better prepare K-12 students for postsecondary success in STEM fields, we need continued research funding for studies of writing-to-learn in STEM. We have identified several next steps for research focusing on both writing-to-learn practices and students' disciplinary-specific writing knowledge. Future research should also involve teacher training that includes a focus on using writing-to-learn in STEM, beginning in the elementary grades.

## References


Ackerman, J. M. (1993). The promise of writing to learn. *Written Communication, 10*(3), 334–370.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*(1), 29–58.

Boscolo, P., & Mason, L. (2001). Writing to learn, writing to transfer. In *Writing as a learning tool* (pp. 83–104). Springer.

Durst, R. K., & Newell, G. E. (1989). The uses of function: James Britton's category system and research on writing. *Review of Educational Research, 59*(4), 375–394.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32*(4), 365–387.

Galbraith, D. (1999). Writing as a knowledge-constituting process. In M. Torrance & D. Galbraith (Eds.), *Studies in writing* (Knowing what to write: Conceptual processes in text production) (Vol. 4, pp. 139–160). Amsterdam University Press.

Galbraith, D., & Baaijen, V. M. (2018). The work of writing: Raiding the inarticulate. *Educational Psychologist, 53*(4), 238–257.

Gillespie Rouse, A., & Rouse, R. (2019). Third graders' use of writing to facilitate learning of engineering concepts. *Journal of Research in Science Teaching, 56*(10), 1406–1430.

Gillespie, A., Olinghouse, N. G., & Graham, S. (2013). Fifth-grade students' knowledge about writing process and writing genres. *The Elementary School Journal, 113*(4), 565–588.

Graham, S. (1999). Handwriting and spelling instruction for students with learning disabilities: A review. *Learning Disability Quarterly, 22*(2), 78–98.

Graham, S. (2018). A revised writer (s)-within-community model of writing. *Educational Psychologist, 53*(4), 258–279.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476.

Graham, S., Schwartz, S., & MacArthur, C. (1993). Learning disabled and normally achieving students' knowledge of writing and the composing process, attitude toward writing, and self-efficacy for students with and without learning disabilities. *Journal of Learning Disabilities, 26*(4), 237–249.

Graham, S., Harris, K., & Hebert, M. (2011). *Informing writing: The benefits of formative assessment*. A report from Carnegie Corporation of New York.

Graham, S., Kiuhara, S. A., & MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research, 90*(2), 179–226.

Hand, B. (2008). Introducing the science writing heuristic approach. In B. Hand (Ed.), *Science inquiry, argument and language: A case for the science writing heuristic* (pp. 1–11). Sense Publishers.

Hand, B. (2017). Exploring the role of writing in science: A 25-year journey. *Literacy learning: The middle years, 25*(3), 16–23.

Klein, P. D., & Boscolo, P. (2016). Trends in research on writing as a learning activity. *Journal of Writing Research, 7*(3), 311–350.

Klein, P. D., Haug, K. N., & Bildfell, A. (2019). Writing to learn. In S. Graham, C. A. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed., pp. 162–184). The Guilford Press.

Langer, J. A., & Applebee, A. N. (1987) *How writing shapes thinking: A study of teaching and learning* (NCTE research report No. 22). National Council of Teachers of English.

Myers, M. (1980). *A procedure for writing assessment and holistic scoring*. ERIC.

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*. The National Academies Press.

National Science and Technology Council. (2018). *Charting a course for success: America's strategy for STEM education*. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2018/12/STEM-Education-Strategic-Plan-2018.pdf

Olinghouse, N. G., & Graham, S. (2009). The relationship between the discourse knowledge and the writing performance of elementary-grade students. *Journal of Educational Psychology, 101*(1), 37–50.

Olinghouse, N. G., Graham, S., & Gillespie, A. (2015). The relationship of discourse and topic knowledge to fifth graders' writing performance. *Journal of Educational Psychology, 107*(2), 391–406.

Scardamalia, M., & Bereiter, C. (1987). Knowledge telling and knowledge transforming in written composition. *Advances in Applied Psycholinguistics, 2*, 142–175.

# Preparing for Adaptive Spelling Instruction During Web-Based Writing Tutoring


**Amanda L. Lindner, Kausalai Wijekumar, and Debra McKeown**

**Abstract** While spelling is a building block of writing and has been shown to impact writing quality in addition to reading development and abilities, the presence of technology in schools has led many to rely on spell check programs while using a computer to compose written text even though these programs may be inconsistent and unreliable. This study investigated the spelling of 163 fifth grade students in the United States utilizing writing samples composed by the participants. Spelling errors were analyzed using latent class analysis to discover hidden categories within the data using nine spelling error categories: (a) Vowel Omission; (b) Vowel Addition; (c) Vowel Substitution; (d) Vowel Sequence; (e) Consonant Omission; (f) Consonant Addition; (g) Consonant Substitution; (h) Consonant Sequence; (i) Vowel-Consonant Sequence. Latent class analysis did not result in any hidden categories among the data. The most frequently occurring errors were Vowel Omission, Vowel Substitution, and Consonant Omission. The results of this study provide insights to the types of errors students make while writing, which will allow for feedback and spelling instruction to be provided to students while utilizing web-based tools such as Intelligent Tutoring Systems to improve their writing.

**Keywords** Intelligent tutoring systems · Spelling · Web-based tools · Writing

Writing quality among elementary children is influenced by their spelling ability (Berninger, 1999; Satangelo & Graham, 2015). Recent advances in writing instruction focus on web-based supports (e.g., Wijekumar et al., 2019). The increasing use of technology in classrooms leads many to rely on spell check programs to catch spelling errors (Peterson & McClay, 2012). However, spell check programs are not

A. L. Lindner (✉)
Texas A&M University–San Antonio, San Antonio, TX, USA
e-mail: alindner@tamusa.edu

K. Wijekumar · D. McKeown
Texas A&M University, College Station, TX, USA
e-mail: k_wijekumar@tamu.edu; debramckeown@tamu.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_7

without fault; accurate suggestions from spell check programs require the initial attempt to be similar to the target word. Additionally, the writer must know which of the suggested words is the correct spelling of their target word, or they may choose the incorrect word from the list of suggested words. For example, if the writer intends to include the word *receipt*, but misspells it as *resiet*, spell check may suggest a list of words including *reset*, *respite*, and *recite*, along with the correct target word. Spell check programs may also miss grammatical errors. For example, if the writer included the word *were* instead of the target word *where*, the spell check program will not indicate *were* as being misspelled as it is a correctly spelled word. Thus, the ability to correctly spell is important when writing both by hand and on a computer. However, these web-based tools can be harnessed to improve spelling, and writing if the program can detect the types of errors students frequently commit while completing writing tasks on the computer.

## 1 Impact of Spelling

Spelling has been identified as a building block of writing development according to the Simple View of Writing (Berninger et al., 2002; Satangelo & Graham, 2015). As one's spelling becomes more fluent, the cognitive demand is reduced, allowing attention to be devoted to higher-order writing processes. Laboring over spelling may distract the writer from their overall goal of the written work and result in shorter and lower quality written works (Berninger, 1999). In a potential effort to limit struggling to spell while writing, writers may limit their written vocabulary to include only words they know how to spell (Moats, 2005).

Considering the impact of spelling on writing, Wijekumar et al. (2019) examined the impact of writing knowledge, motivation, strategic behavior, and skills on writing outcomes, including quality and length (number of words written), of students in fifth grade. The results showed that spelling made a unique and statistically significant contribution to predicting the writing quality and the length of written text among the writing samples. These results confirm the impact of spelling on writing in regards to both quality and quantity.

In addition to impacting one's ability to produce strong written work, inaccurate spelling also impacts others' perceptions of the writer's knowledge and abilities, phenomenon known as Presentation Effect (Graham et al., 2011). When given multiple written works containing the same content, individuals scoring the written works tend to assign a lower grade to the text with more spelling errors regardless of the content (Satangelo & Graham, 2015). Teachers rating written works have also been found to only partially read text that contains spelling errors, resulting in a grade assigned based on only a portion of the writing (Satangelo & Graham, 2015).

In addition to the impact on writing, spelling also plays a key role in an individual's ability to read. Spelling abilities are indicative of an individual's knowledge of the alphabetic principle. Having knowledge of the relationship between letters and sounds allows readers to decode unfamiliar words. As an individual's ability to

decode becomes more automatic, and their reading more fluent, they are able to devote attentional resources to comprehension, according to the Theory of Automaticity (LaBerge & Samuels, 1974).

Beyond indicating one's ability to decode novel words, spelling has been found to predict reading development and later reading abilities (e.g., Abbott et al., 2010; Bahr et al., 2015; Desimoni et al., 2012; Ehri, 2000). Research has shown that appropriate, formal spelling instruction increases students' reading abilities across all grade levels (Graham & Satangelo, 2014). The finding of spelling instruction increasing reading performance across all grade levels is notable as spelling instruction may typically be thought of as exclusive to younger grades.

Examination of the spelling errors students make is critical to determining appropriate spelling instruction. A previous study conducted in the same area as the present study examined the types of spelling errors committed by English Language Learners in Grades 4–6 in writing samples utilizing latent class analysis (Lindner et al., 2020). The results showed that students made errors indicative of both the English and Spanish orthographies, highlighting the difficulty these students face when spelling in a deep orthography such as English while having a transparent first language such as Spanish. Understanding the types of spelling errors students frequently commit will provide insight to the letter-sound correspondences that students have difficulty grasping as well as spelling generalizations or guidelines with which students may be familiar but have not yet mastered. For instance, the findings by Lindner et al. (2020) highlighted the need for spelling instruction which focuses on explicitly examining the differences and similarities between English and Spanish to reduce the students' use of one language's spelling patterns when spelling in the other language. Additionally, knowledge of students' spelling errors can be beneficial to improving students' written vocabulary as they may include higher-level words in their writing if they are able to spell the words accurately.

## 2 The Present Study

The purpose of the present study was to expand on a previous study (Graham et al., 2019; Wijekumar et al., 2019) which found spelling made a unique and statistically significant contribution to students' writing by analyzing the types of spelling errors made by the students. We sought to analyze spelling errors committed by upper-elementary students in order to provide feedback aimed at improving writing skills among these students. With knowledge of the types of errors the students are making, appropriate and immediate feedback may be provided using web-based writing tutoring systems such as We Write (Wijekumar et al., 2016). We hope to utilize the classification of spelling errors and frequency to present spelling instruction within We Write to teach students how to spell. We also plan to introduce scaffolding and feedback on other types of errors if they appear to be sporadic. The research question that guided the present study are: What classifications of spelling errors do students make when writing essays?

# 3 Methods

## *3.1 Participants*

A total of 163 fifth-grade students from three elementary schools (two public schools, one private school) in a single state in the United States were included in the present study. The public elementary schools enrolled a total of 636–694 students and the private school enrolled a total of 174 students. Over 57.5% of children in the public schools qualified for free or reduced lunch at school, and 58.5% were minority students. Data on free or reduced lunch and minority enrollment was not available for the private school.

## *3.2 Measures*

As part of a one-group study (Wijekumar et al., 2019), participants completed up to five writing samples: two persuasive essays and three reading recall written responses. The writing samples were completed by the students at two time points in the school year: once during the fall semester (October) and once during the spring semester (May).

The first persuasive essay (Essay 1) was on the topic of bike safety (e.g., importance of wearing a helmet), and the second persuasive essay (Essay 2) was on the topic of water conservation (e.g., the importance of not wasting water). Students wrote each essay after reading an informational article on the same topic. Each of the informational articles students read were approximately 300 words in length (a single page) and were written at a fourth grade reading level (Wijekumar et al., 2019). Students were given 35 min to read the article, plan their writing, and write their persuasive essay without any assistance from test administrators.

The reading recall written responses were on the topics of fish (Recall 1), rats (Recall 2), and dogs (Recall 3). These passages were equivalent to one another in regards to reading level and length. For each of the reading recall written responses, students read a passage on the topic, placed the reading passage in a folder so they could not look back at it, and wrote what they recalled from the reading passage.

In the previous study (Wijekumar et al., 2019), students participated in a writing intervention which did not teach them spelling. However, spelling may have been reviewed as part of the editing process during their writing instruction.

# 4 Procedures

## *4.1 Scoring Spelling Errors*

Writing samples were examined for spelling errors, and the spelling errors were analyzed by the authors using latent class analysis in Mplus version 8.5 (Muthén & Muthén, 2009) to discover hidden classes among the data as well as likelihood of membership for each category. As latent class analysis requires binary coding of the data, spelling errors were categorized using nine specific categories based on previous research (Lindner et al., 2020): (a) Consonant Omission, (b) Consonant Addition, (c) Consonant Substitution, (d) Consonant Sequence, (e) Vowel Omission, (f) Vowel Addition, (g) Vowel Substitution, (h) Vowel Sequence, and (i) Vowel-Consonant Sequence. These errors categories were chosen as they have been found to be frequently committed by upper-elementary students (e.g., Lindner, 2018; Lindner et al., 2020), and the present study utilized non-standardized writing samples rather than a standardized list of target spelling words. An omission error is the result of the writer omitting a letter in the spelling of the target word (e.g., *realy* for *really*, omitting an *l*). An addition error occurs when the writer includes a letter that does not belong in the word (e.g., *differrent* for *different*, adding an *r*). A substitution error occurs when the writer substitutes a letter in the target word with the incorrect letter (e.g., *coll* for *call*, including an *o* instead of an *a*). A sequence error occurs when the writer includes the correct letters, but places them in the incorrect position (e.g., *macth* for *match*, reversing the order of the *c* and *t*). Each of these errors were categorized as either vowel or consonant depending on whether the incorrect letter was a vowel or consonant. The vowel-consonant sequence error is the result of the writer including the correct vowel and the correct consonant, but they are placed in the incorrect sequence (e.g., *saftey* for *safety*, reversing the order of the *e* and *t*). Each error type may be committed more than once in a given word; for example, if the writer spelled the target word *different* as *diferen*, the writer omitted two consonants (*f* and *t*) in one word. Additionally, it was possible for students to commit the same error type more than once throughout the writing samples submitted. To account for these cases, the total number of each error type was counted and recorded. Then, in order to run latent class analysis, the errors which occurred one or more times were scored as 1 to indicate the occurrence of the error while the error types that were not committed were scored as 0.

## *4.2 Latent Class Analysis*

Latent class analysis was utilized in order to discover hidden classes in the data and gain insight as to which groups of students may make different types of errors (using the error types previously described). Latent class analysis was run multiple times using different numbers of classes to determine the best model fit using information

criteria-based fit statistics, entropy, and model comparisons likelihood ratio tests. Information criteria-based fit statistics used in the analysis included Bayesian information criterion (BIC), Sample-size adjusted BIC (SSA-BIC), and Aikaike information criterion (AIC). For each of these, smaller values indicate a better model fit (Singer & Willet, 2003). Model comparisons likelihood ratio tests used included Lo-Mendell-Rubin likelihood-ratio tests (LMR; Lo et al., 2001) and the Bootstrapped likelihood ratio test (BLRT); these tests compared neighboring models (i.e., a model with K classes compared to a model with K-1 classes). A result of a non-significant value on the model comparisons likelihood ratio tests indicates the model with fewer classes (i.e., K-1) is the best fit for the data (Muthén & Muthén, 2009). Finally, entropy (ranging from 0 to 1) indicates the accuracy of membership designation; a high entropy value indicates that the classification of individuals is more accurate than a low entropy value (Muthén & Muthén, 2009). Additionally, descriptive statistics were run using SPSS.

## 5 Results

In total across all of the writing samples, 2599 spelling errors were made ($M = 15.94$; $SD = 10.704$). The total number of vowel-based errors was 1406 with Vowel Omission 542 errors ($M = 3.33$; $SD = 3.235$), 289 Vowel Addition errors ($M = 1.77$; $SD = 1.984$), 512 Vowel Substitution errors ($M = 3.14$; $SD = 2.655$), and 63 Vowel Sequence errors ($M = 0.39$; $SD = 0.697$). There was a total of 1105 consonant-based errors with 484 Consonant Omission errors ($M = 2.97$; $SD = 2.812$), 319 Consonant Addition errors ($M = 1.96$; $SD = 2.068$), 255 Consonant Substitution errors ($M = 1.56$; $SD = 1.785$), and 47 Consonant Sequence errors ($M = 0.29$; $SD = 0.636$). Finally, there were a total of 88 Vowel-Consonant Sequence errors ($M = 0.54$; $SD = 1.306$).

The overall total number of errors made in Essay 1 was 791 whereas the total number of errors in Essay 2 was 449, a 43% decrease in total errors. The most frequently occurring error types in Essay 1 were Vowel Omission ($n = 138$), Vowel Substitution ($n = 149$), and Consonant Addition ($n = 146$), and the least frequently occurring error type was Consonant Sequence ($n = 9$). The most frequently occurring error types in Essay 2 were Vowel Omission ($n = 92$) and Consonant Omission ($n = 93$) while Vowel Sequence and Consonant Sequence occurred least frequently ($n = 5$ & $n = 4$, respectively). Table 1 shows the number of each error type, and Table 2 shows the percentage of each error type among writing essays.

The overall total number of errors made in the reading recall written responses was 1359. Twenty-nine percent of the errors were made in Recall 1 writing samples ($n = 397$), 43% of the errors were made in Recall 2 writing samples ($n = 589$), and 27% were made in Recall 3 writing samples ($n = 373$). The most frequently occurring errors in the reading recall written responses were Vowel Omission ($n = 312$), Vowel Substitution ($n = 294$), and Consonant Omission ($n = 275$). Similar to the writing essays, sequence errors were the least frequently occurring errors among the

**Table 1** Errors by type for writing essays

| Writing essay | Error type | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Vowel omission | Vowel addition | vowel substitution | Vowel sequence | Consonant omission | Consonant addition | Consonant substitution | Consonant sequence | Vowel-consonant sequence |
| Essay 1 | 138 (18.2%) | 48 (6.3%) | 149 (19.7%) | 19 (2.5%) | 116 (15.3%) | 146 (19.3%) | 95 12.6%) | 9 (1.2%) | 35 (4.6%) |
| Essay 2 | 92 (20.5%) | 70 (15.6%) | 69 (15.3%) | 5 (1.1%) | 93 (20.7%) | 49 (10.9%) | 46 (10.2%) | 4 (0.9%) | 21 (4.6%) |

**Table 2** Errors by type for reading recall written responses

| | Error type | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Reading recall passage | Vowel omission | Vowel addition | Vowel substitution | Vowel sequence | Consonant omission | Consonant addition | Consonant substitution | Consonant sequence | Vowel-consonant sequence |
| Fish | 99 (24.9%) | 45 (11.3%) | 86 (21.6%) | 3 (0.75%) | 76 (19.1%) | 38 (9.5%) | 38 (9.5%) | 3 (0.75%) | 9 (2.26%) |
| Rats | 81 (13.7%) | 60 (10.2%) | 137 (23.2%) | 24 (4.0%) | 147 (24.9%) | 53 (8.9%) | 46 (8.9%) | 29 (4.9%) | 12 (2.0%) |
| Dogs | 132 (35.3%) | 30 (8.0%) | 71 (19.0%) | 12 (3.2%) | 52 (13.9%) | 33 (8.8%) | 30 (8.8%) | 2 (0.53%) | 11 (2.9%) |

**Table 3** Latent class analysis models

| Number of Classes | Log likelihood | AIC | BIC | SSA-BIC | VLMR | Entropy |
|---|---|---|---|---|---|---|
| 1 | −802.214 | 1622.428 | 1650.272 | 1621.779 | $p = 1.000$ | 1 |
| 2 | −757.702 | 1553.405 | 1612.186 | 1552.035 | $p = 1.000$ | 0.777 |
| 3 | −749.441 | 1556.881 | 1646.600 | 1554.790 | $p = 1.000$ | 0.881 |
| 4 | −741.704 | 1561.408 | 1682.064 | 1558.595 | $p = 1.000$ | 0.837 |
| 5 | −734.507 | 1567.014 | 1718.608 | 1563.481 | $p = 1.00$ | 0.795 |
| 6 | −726.180 | 1570.360 | 1752.891 | 1566.106 | $p = 1.000$ | 0.811 |

reading recall written responses (Vowel Sequence: $n = 39$; Consonant Sequence: $n = 34$; Vowel-Consonant Sequence: $n = 32$). Table 3 shows the number of each error type, and Table 4 shows the percentage of each error type among reading recall written responses.

Latent class analysis did not result in any significant hidden classes among the data. The model comparison tests did not have significant results, indicating that the data were homogeneous and there were no hidden classes based on spelling error types. The latent class analysis models are depicted in Table 3.

## 6 Discussion

Whereas previous research in schools in the same area of the state in which this study was conducted found specific latent classes in the spelling error categories (Lindner et al., 2020), this study did not result in the same latent classes. One possible explanation for this may be that the previous study conducted by Lindner et al. (2020) focused on students who were English Language Learners, and that was not a specified qualification for participants in the present study. Even though the data did not show hidden latent classes, the results of the spelling errors cannot be ignored. As spelling has been found to be an essential building block of writing (Berninger et al., 2002; Satangelo & Graham, 2015), addressing spelling errors made by students is of importance in the classroom to improve their writing. Additionally, strong spellers are typically strong readers as encoding is more difficult than decoding, and spelling has been found to be predictive of later reading development and abilities among both monolingual and bilingual students (e.g., Abbott et al., 2010; Bahr et al., 2015; Caravolas et al., 2001; Chua et al., 2016; Desimoni et al., 2012; Ehri, 2000).

The results of this study show that students understand or have experience with certain spelling patterns, but may not have a full grasp of when each spelling pattern occurs based on the error types made by the students. For example, the high number of Consonant Addition and Consonant Omission errors in the writing samples shows a familiarity with spelling patterns such as the doubling rule, where the consonant needs to be doubled in certain circumstances (e.g., when adding a vowel suffix to a one-syllable word that ends with one consonant). However, the

inconsistent application of this spelling generalization may result in the inappropriate doubling of the consonant, resulting in a Consonant Addition error, or omission of the consonant that should be doubled, resulting in a Consonant Omission error.

Additionally, the data showed a high number of vowel-based errors, specifically Vowel Omission and Vowel Substitution. These errors are consistent with the deep orthography of the English language as it has many phoneme-grapheme inconsistencies. Thus, with many graphemes to represent a vowel phoneme, students are likely to commit vowel-based errors when spelling in English. Appropriate instruction on phoneme-grapheme correspondences among vowels would be beneficial for these students.

Consequently, we are now poised to present spelling instructional modules embedded within the We Write web-based tutoring system specifically about the rules of vowel-based errors and overcoming spelling challenges. We also intend to develop specific feedback messages to students about the types of errors in writing samples. For example, we can show students a pop-up window that shows the consonant omission errors, and show the student the consonant doubling rule and help them to determine whether the consonant needs to be doubled.

## 7 Limitations & Future Research

The present study focused on a preliminary analysis of spelling errors committed by upper-elementary students in non-standardized writing samples rather than a specific, standardized list of spelling words. Due to this, the analysis was somewhat limited in regards to the types of errors analyzed. Future studies may examine spelling utilizing a specific list of target spelling words, allowing for additional insights to students' spelling including orthographic features and spelling developmental stages. This will allow for a more in-depth understanding of students' spelling abilities.

An additional limitation of using writing samples instead of a specific set of spelling words is the potential for limited words included in the writing samples as research has found that writers' vocabulary may be limited to only the words they know they are able to spell (Moats, 2005). One result of this may be the low number of sequence errors among vowels and consonants. Sequence errors are inherently more difficult to commit as the word must contain a sequence of vowels or consonants, and the writer must know all of the vowels or consonants that are included in the word, only to place them in the incorrect order when spelling the word. However, it is possible that students purposefully avoid using words with difficult sequences if they are unsure of the spelling and instead use a word with which they are more familiar as previous research has found individuals may include only words they are confident spelling in their written work (Moats, 2005). For example, if a student is uncertain of the spelling of the word *receive*, they may choose to use the word *get* in order to avoid misspelling the word, thus limiting their written vocabulary.

Appropriate spelling instruction may help these students to include more robust vocabulary in their written works.

With insight to the types of errors students are committing, web-based tools can be harnessed to provide immediate feedback and instruction to students. For example, upon submission of written works, students using web-based tools such as ITSs (Wijekumar et al., 2016, 2019) can be shown a pop-up window providing immediate feedback on their spelling along with guidelines and instruction regarding the errors they committed and the spelling generalizations or patterns the students may not be grasping.

## References


Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281–298. https://doi.org/10.1037/a0019318

Bahr, R. H., Silliman, E. R., Danzak, R. L., & Wilkinson, L. C. (2015). Bilingual spelling patterns in middle school: It is more than transfer. *International Journal of Bilingual Education and Bilingualism, 18*(1), 73–91. https://doi.org/10.1080/13670050.2013.878304

Berninger, V. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly, 22*, 99–112. https://doi.org/10.2307/1511269

Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities, 35*, 39–56. https://doi.org/10.1177/002221940203500104

Caravolas, M., Hulme, C., & Snowling, M. J. (2001). The foundations of spelling ability: Evidence from a 3-year longitudinal study. *Journal of Memory and Language, 45*(4), 751–774. https://doi.org/10.1006/jmla.2000.2785

Chua, S. M., Rickard Liow, S. J., & Yeong, S. H. M. (2016). Using spelling to screen bilingual kindergarteners at risk for reading difficulties. *Journal of Learning Disabilities, 49*(3), 227–239. https://doi.org/10.1177/0022219414538519

Desimoni, M., Scalisi, T. G., & Orsolini, M. (2012). Predictive and concurrent relations between literacy skills in grades 1 and 3: A longitudinal study of Italian children. *Learning and Instruction, 22*, 340–353. https://doi.org/10.1016/j.learninstruc.2012.02.002

Ehri, L. C. (2000). Learning to read and learning to spell: Two sides of a coin. *Topics in Language Disorders, 20*(3), 19–36. https://doi.org/10.1097/00011363-200020030-00005

Graham, S., & Satangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing: An Interdisciplinary Journal, 27*, 1703–1743. https://doi.org/10.1007/s11145-014-9517-0

Graham, S., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children, 44*(4), 1–12.

Graham, S., Wijekumar, K., Harris, K. R., Lei, P., Fishman, E., Ray, A., & Houston, J. (2019). Writing skills, knowledge, motivation, and strategic behavior predict students' persuasive writing performance in the context of robust writing instruction. *Elementary School Journal, 119*(3), 487–510. https://doi.org/10.1086/701720

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology, 6*, 293–323. https://doi.org/10.1016/0010-0285(74)90015-2

Lindner, A. L. (2018). *Analyzing spelling performance among native Spanish-speaking English language learners utilizing latent class analysis and contextual factors* [Unpublished doctoral dissertation]. Texas A&M University.

Lindner, A. L., Wijekumar, K., & Joshi, R. M. (2020). English spelling performance among Spanish-speaking ELLs. *Journal of Learning Disabilities*. Advance online publication. https://doi.org/10.1177/0022219420982995.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*, 767–778. https://doi.org/10.1093/biomet/88.3.767

Moats, L. (2005). How spelling supports reading. *American Educator, 6*, 12–22.

Muthén, L. K., & Muthén, B. (2009). Categorical latent variable modeling using Mplus: Cross-sectional data [PDF document]. Retrieved from: www.statmodel.com/courses2.shtml

Peterson, S. S., & McClay, J. K. (2012). Assumptions and practices in using digital technologies to teach writing in middle-level classrooms across Canada. *Literacy, 46*(3), 140–146. https://doi.org/10.1111/j.1741-4369.2012.00665.x

Satangelo, T., & Graham, S. (2015). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review, 28*(2), 225–265. https://doi.org/10.1007/s10648-015-9335-1

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis*. Oxford University Press.

Wijekumar, K., Harris, K., Graham, S., & Meyer, B. J. F. (2016). We-write: A teacher and technology supported persuasive writing tutor for upper elementary students. In S. Crossley & D. S. McNamara (Eds.), *Adaptive educational technologies for literacy instruction*. Routledge Publishers – Taylor & Francis Group.

Wijekumar, K., Graham, S., Harris, K. R., Lei, P. W., Barkel, A., Aitken, A., Ray, A., & Houston, J. (2019). The roles of writing knowledge, motivation, strategic behaviors, and skills in predicting elementary students' persuasive writing from source material. *Reading and Writing, 32*, 1431–1457. https://doi.org/10.1007/s11145-018-9836-7

# Concurrent and Longitudinal Relationships Between Written Composition (Length and Quality) and Spelling Errors (Phonographic, Lexical, Morphological, Total) in French Children in Grades 3 and 6 or in Grades 3 and 4

**Michel Fayol, Bernard Slusarczyk, Virginia Berninger, and Pascal Bressoux**

**Abstract** The following chapter aims to explore in an orthographic system like French, if spelling is related concurrently to written composition within specific grade levels (Grade 3, Grade 6, or Grade 4) and longitudinally across grade levels (Grade 3 and Grade 6 or Grade 3 and Grade 4). French spelling is interesting because it includes phonographic irregularities (i.e., inconsistencies), lexical difficulties, and morphological silent markers (e.g., plural noun, adjective, and verb agreement). Pupils were asked in every grade to compose narrative texts either from verbal instructions (Study 1) or from strips (Study 2). Text length, text quality, and three categories of spelling errors were coded and analyzed. Two important results emerged: First, significant concurrent and longitudinal relationships were observed. Both text characteristics (i.e., length and quality) and total spelling errors were significantly correlated longitudinally across grade levels in the two studies. Second, regression analyses provided evidence that the more spelling errors the texts contained, the texts were shorter and rated of lower quality. Further analyzes showed that the errors with the most weight were the lexical errors. This result was unexpected insofar as national assessments in France have reported evidence that morphological errors are the most frequent and the most troublesome in students'

M. Fayol (✉)
Clermont Auvergne University and CNRS, Clermont-Ferrand, France
e-mail: michel.fayol@uca.fr

B. Slusarczyk
Head of Establishment Albert Camus College, Neuilly-sur-Marne, France

V. Berninger
University of Washington, Seattle, Washington, USA

P. Bressoux
Grenoble Alpes University – Institut Universitaire de France, Saint-Martin-d'Hères, France



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_8

written compositions. Future research should continue to investigate the concurrent and longitudinal relationships of spelling and written composition in orthographies for different languages as well as the instructional applications of these differences.

**Keywords** Spelling · Phonological errors · Lexical errors · Morphological errors · Written composition

Written Verbal Production (WVP) is a complex activity involving several components, including spelling, which may pose different problems, depending on the spelling system in a language. Some studies focused on English, a morphophonemic orthography system, in which there are alternative associations between units of speech (i.e., units of phonology, phonemes, P), units of writing (i.e. units of orthography, graphemes G), and units of morphology (i.e., morphemes, M). However, the relationships between G and P are not exactly the same in the reading direction as between P and G in the spelling direction in English (Venezky, 1995). See Nicknames for the reading direction and Substitutes for the spelling direction in the *Reproducible Instructional Materials* for the *Revised PAL Research Based Reading and Writing Lessons* (Berninger & Abbott, 2020).

In addition, research findings for the relationship between spelling and WVP may depend on the language. In English, for example, spelling influences WVP. For example, Juel (1988) found that from Grade 1 to Grade 4 spelling skills were related to written composition and explained 29% of the variance in quality of writing products in Grade 1 and 10% of the variance in quality of writing products in Grade 4, when spelling skill had improved compared to Grade 1. Results of several other studies support the same conclusion in English-speaking children that spelling is related to written composition (Abbott et al., 2010; Berninger et al., 1992, 2011; Graham et al., 1997). However, see Kim et al. (2014) for contrasting results for English.

The contribution of transcription (handwriting and spelling) skills to written text production in the early grades could result from two constraints. On the one hand, writing takes place in a working memory environment, the capacity of which is highly limited, especially in young children. Domain and discourse knowledge and skills involved in text production thus compete for a limited pool of resources. The resources likely to be allocated to the different components of the writing process are restricted by non-automated skills such as handwriting or spelling that require attention and control (Kellogg, 2001; McCutchen, 2000). Consequently, spelling skills can affect children's written text production (Fayol, 2016; Graham et al., 1997; Kim et al., 2011), writing fluency (Kent et al., 2014), and writing quality (Kent et al., 2014; Kim et al., 2015). On the other hand, the English orthographic system is highly complex and difficult to learn and use. It takes a long time for children to abstract the statistical regularities of the orthographic lexicon and memorize a large part of the orthographic lexicon (often referred to as word-specific spelling; Treiman, 2018). However, explicit spelling instruction has significant positive

effects on the writing performance of English-speaking children (Berninger et al., 1998).

By contrast, the Spanish orthographic system, the Italian orthographic system, and the Finnish orthographic system are thought to be easier to learn (Joshi & Aaron, 2005). In so-called transparent systems in which the associations between phonemes and graphemes are regular (highly predictable) spelling interferes less or briefly with written narrative performances (Mäki et al., 2001 in Finnish; Babayigit & Stainthorp, 2010, in Turkish). Italian has a transparent orthography, and Arfé et al. (2016) reported that spelling skills contributed to accuracy and quality of written composition in Italian children in Grade 2 and Grade 3. However, Italian has complex morphology and grammar that could also affect text generation and text production.

French is a Romance language in which words include a lexical component combined with a grammatical one. As with many other orthographies, the French orthographic system for spelling is an alphabetic system of phoneme-grapheme correspondences and it is also closely related to its oral language. It relies on two general principles: the phonographic principle and the semiographic principle (Fayol, 2014; Jaffré & Fayol, 2005). The phonographic principle establishes correspondences between graphemes (G) and phonemes (P). Graphemes are more numerous than phonemes: one phoneme can thus be associated with several graphemes. In reading (from G to P), the G-P associations are relatively consistent, and as a consequence, reading is rather easy to learn in French (Seymour et al., 2003). In spelling, however, the P-G relations (from P to G) are much more complex. For example, the French phoneme /*E*/ can be spelled differently (*è*, *ê*, *ai*, *est*, *et*, etc.). A simple application of phoneme analysis is thus not sufficient to identify the corresponding orthographic unit for spelling the majority of French words. Writers must also refer to lexical orthography (Martinet et al., 2004), statistical regularities, and morphological knowledge (Pacton et al., 2005). Spelling many French words requires mobilization of associations between sound and letter units larger than the PG correspondences and/or access to an orthographic lexicon of words with memorized and directly retrievable orthographic forms.

Regarding the semiographic principle, most words in French are composed of several morphemes (about 75% according to Rey-Debove, 1984). Derivational morphological knowledge helps to spell correctly in a number of cases, for example, when words end with silent final letters (e.g. *grand* `gR@ tall)`. Referring to word families (*grande*—fem tall; *grandeur*—greatness, etc.) leads to putting a final -d despite the lack of any phonological cue. French has a rich written inflectional morphology, but the plural and feminine endings of nouns, adjectives, and verbs are rarely pronounced. For example, *la poule* (hen, singular noun) after et *les poules* (plural hens) are pronounced the same way, as are the singular and the plural forms of the adjective *rousse* (*rousse* versus *rousses*; adj red) as well as the singular and plural forms of most verbs such as *picore* versus *picorent* (is versus are pecking; Largy & Fayol, 2001). All these specificities make the learning of French spelling both a complex and a long process.

Developing writers learning to spell in French have to discover the alphabetic principle; the regularity of the phoneme-grapheme correspondences is the most crucial factor (Catach et al., 1980; Gak, 1976; Veronis, 1986). They also need to learn that written French codes speech, and only indirectly meaning. Letters (graphemes and series of graphemes) have to be matched up with phonemes and sequences of phonemes. This matching avoids misreading phonologically plausible but incorrectly spelled words (e.g. *bato*) for *bateau*. Learners read regular words more accurately and spell them more correctly than irregular words that they will often regularize (e.g., /fam/ is transcribed *fam* instead of *femme* woman). Implausible phonological errors can occur (e.g. *bado* or *badeau* instead of *bateau*).

In the French spelling system, knowledge of PG correspondences is necessary but not sufficient. Two other categories of orthographic knowledge are also necessary to master French spelling. First, lexical knowledge, that is, the knowledge that a certain word is written in a particular way, for instance *enfant* (child), is acquired mainly through repeated encoding of the orthographic word form until it is memorized and entered into the orthographic lexicon and can be directly retrieved from memory. Second, perceiving regularities about frequently co-occurring letters, for example, double consonants, such as *-ll-* or *-nn-*, facilitates spelling, whereas the rarity of *-bb-* doubling often interferes. The acquisition of the correct orthographic forms and the use of orthographic regularities takes time and can remain fragile even in adults who misspell due to lexical errors or orthographic regularities errors. In most cases, these errors are phonologically plausible because they involve sound-spelling correspondences but not for the standard correct spelling.

In addition, the inflectional morphology of written French has unique characteristics that contribute to spelling; but most of the time, the plural or feminine inflections for nouns, adjectives, and verbs have no phonological counterpart. In the spoken language, plurality is mainly signalled by determinants (e.g. *le / la* or un / une in the singular and *les / des* in the plural), whereas the other segments rarely contain any audible mark (Dubois, 1965). This specificity poses problems for both developing writers and most adult spellers. For instance, the national assessment on spelling performances in fifth graders (DEPP, 2016; see also Manesse, & Cogis, 2007) reported a high proportion of errors due to omissions or substitutions of silent marks such as *-e*, *-s* and *-nt* regarding subject-verb agreement and adjective-noun agreement. The same findings have been reported in adolescent spellers (Grade 9; see Bosse et al., 2020) and adult spellers (Lucci & Millet, 1994).

Consistent with Catach (1986)'s work and the Triple Word Theory (Richards et al., 2006), researchers distinguish three categories of errors: phonological (e.g. bado instead of bato—boat), orthographic (e.g. retart instead of retard—late), and morphological (e.g. ils tombes instead of ils tombent - they are falling down) (see Bahr et al., 2012). All three categories of errors are observed early in spelling development. Their presence in all developing writers raises two questions. First, are these different categories of errors associated with quantitative or qualitative measures of written composition within a given period of schooling, for example, Grade 3 or Grade 6 (Study 1 in this chapter) and Grade 3 or Grade 4 (Study 2 in this chapter)? If yes, is this association global or specific to certain errors (e.g.,

morphological)? Second, is the number and nature of errors at one time in schooling (Grade 3) correlated longitudinally with number and nature of errors at another time in schooling (Grade 6, Study 1, or Grade 4, Study 2)?

Relatively few studies have considered the relationship between French spelling and WVP, as both Studies 1 and 2, presented hereafter, did. On the one hand, French spelling often contains many silent letters that have no phonological correspondence. On the other hand, the morphology is complex and the inflected forms are often homophones. Because beginning in Grade 2 French children receive explicit spelling instruction, we investigated the relationship between French spelling and WVP after Grade 2—in Grades 3 and 6 (Study 1) or in Grades 3 and 4 (Study 2). In the first longitudinal study we addressed whether spelling difficulties of French students in Grade 3 are correlated with WVP in Grade 6. In the second longitudinal study, we addressed whether spelling difficulties of French students in Grade 3 are correlated with WVP in Grade 4.

## 1 Study 1

The first study explored the concurrent and longitudinal relationships between spelling and composition in a group of 79 pupils in 7 classes from 7 different schools whose written compositions were collected when they were in Grade 3 and then again when they were in Grade 6. Two questions were addressed. First, are there concurrent relationships between the nature or frequency of spelling errors and the length or quality of written texts in Grade 3 and in Grade 6? Second, are there longitudinal relationships between the nature of spelling errors or total spelling errors and length or quality of written texts in Grade 3 and spelling errors, length, and quality of written texts in Grade 6?

## 2 Method

***Sample, Material and Procedure*** Every year in France, the Ministry of Education assesses pupils' achievement at several grade levels. In Study 1, pupils were tested in Grade 3 and again in Grade 6. Tests were administered to groups during normal school hours. They were administered by teachers who gave the instructions and supervised students in completing the writing task. To protect the anonymity of student participants in the research, the students' productions were assigned a number which made it possible to match the productions of Third and Sixth graders in a given geographic region (the *Département de la Haute-Loire*). The only information available about the students' background was that they are middle class.

In Grade 3, pupils were provided a prompt: "À ce moment là, un bruit étrange lui signala qu'il/elle n'était pas seul/e dans la maison" (At that moment, a strange noise

informed him/her that s/he was not alone in the house). They were told by the teacher that they had to compose a narrative including this sentence, and that this sentence was supposed to describe an important event of their narrative. Pupils had a blank page on which to write and they had 30 minutes to write their text.

In Grade 6, pupils were read a text about "Cats". The text consisted of 23 lines written in the past tense. The instructions required writing about the past and from the first-person perspective. In addition, pupils were asked to continue the story and imagine a possible end to the cat story. They were allowed between 45 and 90 min to compose. Next procedures for scoring are explained.

***Spelling Errors*** All spelling errors were collected in each of the 158 texts and categorized following this classification from Catach, 1986): phonological, orthographic and morphological word forms (see also Daffern et al., 2017; Richards et al., 2009). For each pupil, the proportion of errors in each category was calculated by dividing the number of errors in each category by the total number of words in the text.

***Text Length and Text Quality*** For each text, the number of words was computed to determine the text length. The judges rated the quality of the texts as transcribed by the children in their own handwriting (legible and illegible) and spelling (correct and incorrect). The quality ratings were based on five translation features: creativity; coherence and organization; number of different ideas; quality of the vocabulary; and syntax. Five trained judges were asked to follow guidelines to provide a score (out of 20) about each text on each of the five features. Instead of providing a correlation between the raters' judgments of quality, we reported the mean of the composite score of the five raters, which should be a reliable estimate of quality score across the five judges.

## 3 Results

The means and standard deviations for each score (length and quality of written compositions and phonographic, lexical, and morphological spelling errors and total spelling errors) are reported in Table 1. Within-grade concurrent correlations are also reported between each of those measures for the same 79 pupils when they were in Grade 3 (upper part of Table 1) and then in Grade 6 (lower part of Table 1).

The results reported in Table 1 show that in both Grades 3 and 6, the morphological errors were more frequent than the lexical errors, which were more frequent than the phonographic errors. This result is expected due to the high difficulty of French morphology. In Grade 6, all the categories of errors were significantly correlated with each other, which suggests that pupils who produce errors in one category also produce errors in the other categories. Also, text length and rated text quality were not correlated with each other in Grade 3. Interestingly, neither text length nor text

**Table 1** Descriptive statistics for the different variables and their correlations in study 1

| | Means | SD | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| **Third grade** | | | | | | | |
| 1 Length (words) | 115 | 41 | 0.15 | 0.03 | −0.10 | −0.17 | −0.06 |
| 2 Quality (/100) | 50 | 17 | – | −0.04 | −0.02 | −0.01 | −0.04 |
| 3 Phonographic errors | 4 | 2 | – | – | 0.21 | 0.17 | 0.58** |
| 4 Lexical errors | 8 | 6 | – | – | – | 0.20 | 0.65*** |
| 5 Morphological errors | 12 | 6 | – | – | – | – | 0.72*** |
| 6 Total errors | 24 | 10 | – | – | – | – | – |
| **Sixth grade** | | | 2 | 3 | 4 | 5 | 6 |
| 1 Length (words) | 270 | 59 | 0.53*** | −0.27* | −0.35** | −0.36** | −0.42** |
| 2 Quality (/100) | 57 | 17 | – | −0.33** | −0.44** | −0.31** | −0.44** |
| 3 Phonographic errors | 1.5 | 1.6 | – | – | 0.32** | 0.32** | 0.58*** |
| 4. Lexical errors | 2 | 2 | – | – | – | 0.55*** | 0.75*** |
| 5 Morphological errors | 9 | 4 | – | – | – | – | 0.90*** |
| 6 Total errors | 12 | 7 | | | | | |

quality in Grade 3 was correlated with any of the kinds of spelling errors or the total frequency of spelling errors. In contrast, in Grade 6, on the one hand, text length and text quality were correlated with each other; and, on the other hand, all spelling error categories were significantly, but negatively, correlated with both length and quality. The more the texts included spelling errors, whatever their categories, the shorter the texts and the less their (judged) quality.

To summarize, there was no relationship within Grade 3 between text length or rated quality of text and spelling errors. However, text length and text quality are moderately and significantly correlated within Grade 6, as well as moderately, but negatively, correlated with all types of spelling errors.

The longitudinal correlations between Grade 3 and Grade 6 for the same scores for text length and rated quality and kinds of spelling errors are reported in Table 2. Length (number of words in written texts) in Grade 3 was not correlated with rated quality for the written texts or any of the kinds of spelling errors in Grade 6. Rated quality for the written texts in Grade 3 was significantly, but negatively, correlated only with morphological spelling errors and total spelling errors in Grade 6.

Regarding spelling errors, phonographic errors in Grade 3 were significantly and positively correlated with phonographic, lexical, and morphological errors, and total spelling errors in Grade 6. Lexical errors in Grade 3 were significantly and positively correlated with lexical, morphological, and total spelling errors in Grade 6. Morphological errors in Grade 3 were not significantly correlated with any spelling errors in Grade 6. Total spelling errors in Grade 3 were significantly and positively correlated with phonographic, lexical, morphological, and total spelling errors in Grade 6.

Lexical errors in Grade 3 were significantly, but negatively, correlated with text length in Grade 6; the more lexical spelling errors, the shorter the text. There were modest negative correlations between the total spelling errors in Grade 3 and text

**Table 2** Correlations between written composition and spelling errors in grades 3 and 6

| | Third grade | | | | | |
|---|---|---|---|---|---|---|
| Sixth grade | 1 Length (words) | 2 Quality (/100) | 3 Phono-graphic errors | 4 Lexical errors | 5 Morpho-logical errors | 6 Total errors |
| 1 Length (words) | 0.25** | 0.13 | −0.16 | −0.30** | −0.006 | −0.22* |
| 2 Quality (/100) | 0.01 | 0.39** | −0.08 | −0.22 | −0.08 | −0.23* |
| 3 Phono-graphic errors | −0.05 | −0.02 | 0.32** | 0.19 | 0.18 | 0.46*** |
| 4 Lexical errors | −0.03 | −0.18 | 0.34** | 0.42*** | 0.21 | 0.34** |
| 5 Morpho-logical errors | −0.04 | −0.25* | 0.27* | 0.23* | 0.18 | 0.46*** |
| 6 Total errors | −0.03 | −0.23* | 0.39** | 0.35** | 0.21 | 0.34** |

Reading note: Sixth-Grade text length correlates with Third-Grade text length at 0.25 (significant $p < 0.01$)

length and text quality in Grade 6, which suggests that students who made a higher number of spelling errors in Grade 3 wrote shorter and lower quality texts in Grade 6. Interestingly, and contrary to what could have been expected regarding French orthography, morphological errors were not related to text length and quality.

Three series of regression analyses were conducted to disentangle the effects of predictors. The first one was about whether the different types of spelling errors in Grade 3 were related to the total percent of errors in Grade 6. The results showed that only phonographic ($\beta = 1.04$; $t(75, 4) = 5.88$ $p < 0.0001$) and lexical ($\beta = 0.27$; $t$ (75, 4) = 3.06 $p < 0.003$) errors significantly predicted total percent of spelling errors in grade 6.

The second and third regressions addressed the question of whether pupils producing many spelling errors in Grade 3 tended to compose shorter and lower quality texts in Grade 6. We tested the relationship of Grade 3 spelling errors to the length and quality of the Grade 6 written texts, after controlling for, respectively, the effect of text length and text quality in Grade 3. As for predicting text length in Grade 6, the text length in Grade 3 was significant ($\beta = 0.27$; $t(76,3) = 2.51$ $p < 0.02$) and the total percent of spelling errors was significant ($\beta = -0.24$; $t(76,3) = -2.23$ $p < 0.03$). Spelling errors did not interact with text length and quality at Grade 3 in predicting text length and quality at Grade 6. The model explained 12% of the variance.

Regarding predicting the text quality, we entered in a hierarchical regression first the Grade 3 text quality and then the three kinds of spelling errors (phonographic, lexical, morphological) in Grade 3. The model explained 29% of the variance (adjusted 26%). The Grade 3 text quality was significant ($\beta = 0.41$; $t(75,4) = 3.97$ $p < 0.0002$) and Grade 3 phonographic errors were significant ($\beta = -0.53$; $t(75,4) = -2.73$ $p < 0.008$); lexical errors were marginally significant ($\beta = -0.18$; $t(75,4) = -1.88$ $p = 0.06$). No interactions were significant.

To summarize, the longitudinal approach to relationships between spelling in French and written text production provided evidence for two main findings. First, phonographic and lexical spelling errors in Grade 3 were significant predictors of phonographic and lexical spelling errors in Grade 6. However, unexpectedly, morphological spelling errors in Grade 3 were not significant predictors of morphological spelling errors in Grade 6. Second, although spelling errors in Grade 3 were not significantly related to composition quantity and quality in Grade 3, these same spelling errors were significantly related to both text length and text quality in Grade 6. After controlling for characteristics of Grade 3 written texts, French pupils who were not very good spellers in Grade 3 tended to compose shorter and lower quality texts in Grade 6. These results are only partly consistent with results of previous studies for other languages such as English that reported early relationships of spelling with text production, although of smaller size than for handwriting in the early primary grades (Grades 1–3) (e.g., Graham et al., 1997).

It was surprising that no concurrent correlation was observed between spelling and text production in Third Grade French pupils; future studies need to evaluate if this finding replicates. It is the case that the Graham et al. (1997) study analyzed relationships between normed spelling measures and normed written composition measures in English, whereas Study 1 in this chapter analyzed coded spelling errors and coded text characteristics (length in words and rated quality) rather than normed measures of spelling and written composing.

## 4 Study 2

The first goal of the second study in this chapter was, therefore, to determine if the finding that coded spelling errors were not related to coded text characteristics (length and rated quality) in Grade 3 French pupils replicated in another sample. The second goal was to analyze longitudinal relationships from Grade 3 to Grade 4 for different kinds of spelling errors and coded text characteristics (length and rated quality) for French pupils.

## 5 Method

***Sample*** An initial sample of 263 pupils from 18 classrooms (14 schools in the French *Département de la Haute Loire*) was selected to be followed from the beginning of Grade 3 to the end of Grade 4. Due to changes in teachers and absences of pupils during the two school years, only 173 pupils (87 girls) from 15 classrooms participated in all the tests. Analyses revealed that the sample of excluded students was very similar to the sample of remaining students. The only significant difference was in the number of phonological errors at Grade 3, which was slightly lower for remaining students (M = 1.8) than for excluded students (M = 3.5) ($p < 0.05$).

***Material and Procedure*** Two tests were used. The first one (Story 1 at Time 1 T1), at the beginning of Grade 3, was the production part of the standardized ECL-Collège test (Khomsi et al., 2005); the second one (Story 2 at Time 2 T2), at the end of Grade 4, was developed by the French authors of this chapter following the model of the ECL test. We wanted to avoid the students having to compose twice from the same material. In both cases, a story in 6 images was presented on a page. To facilitate the planning of the content of the text, the pupils had first to describe in writing each of the images, using some space under each of them. Second, they had to write the story told by the comic strip on a specific place at the bottom of the same page. There were no time limits for either of the two writing tasks. Only the second composed text was analyzed and rated. The same scoring principles as used in Study 1 were also used in Study 2 for the number of words (length) and the three categories of spelling errors (Catach, 1986). All scores were established by only one judge (the second author of this chapter) because the number of texts (more than 350) was too high to require expert help. Text quality was scored using the ECL rating grid (2005): Thirteen pieces of information were considered essential: 9 describing facts (e.g., there are two characters; two phones; uses tools etc.) and 4 constituting inferences (e.g., he is jealous; he is happy etc.). The same analysis grid with 13 items (9 + 4) was established for the second story.

## 6 Results

The means, standard deviations, and concurrent correlations are reported in Table 3 for the same 173 pupils when they were in the beginning of Grade 3 (upper part of Table 3) and then at the end of Grade 4 (lower part of Table 3). The length and quality of the text were significantly and positively correlated in both Grade 3 and Grade 4. The longer the composition, the higher the quality. However, this result may be a consequence of how quality was rated, that is, by the number of pieces of information.

All the categories of spelling errors were significantly and positively correlated with each other in Grade 3 (range $r = 0.22$ to $r = 0.74$) and in Grade 4 (range $r = 0.37$ to $r = 0.86$). These findings suggest that pupils who produce errors from one category of spelling errors also produce errors from the other categories of spelling errors. Finally, in Grade 3, the total number of spelling errors was not correlated with the length and quality of texts, thus, replicating the finding for Grade 3 in the Study 1. Then, in Grade 4, the correlations between the number of spelling errors and written text production became negative and significant: $r = -0.17$ for text length and $r = -0.32$ for text quality. The more frequent the spelling errors, the shorter the text and the lower its quality. Clearly the relationship between spelling and written text production was different within Grade 3 than within Grade 4 in the French pupils. See Table 3.

**Table 3** Descriptive statistics for the different variables, and their correlations in study 2

| | Means | SD | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| **Third grade** | | | | | | | |
| 1 Length (words) | 60 | 22 | 0.25** | −0.06 | −0.11 | −0.007 | −0.08 |
| 2 Quality (/13) | 5.7 | 1.9 | – | 0.06 | −0.04 | −0.10 | −0.05 |
| 3 Phonographic errors | 1.8 | 3.5 | – | – | 0.22** | 0.23** | 0.55*** |
| 4 Lexical errors | 9.6 | 6.4 | – | – | – | 0.32** | 0.80*** |
| 5 Morphological errors | 9.9 | 5.5 | – | – | – | – | 0.74*** |
| 6 Total errors | 21.2 | 11.1 | – | – | – | – | – |
| **Fourth grade** | | | 2 | 3 | 4 | 5 | 6 |
| 1 Length (words) | 122 | 43.5 | 0.35*** | −0.08 | −0.20** | −0.10 | −0.17* |
| 2 Quality (/13) | 7 | 2.6 | – | −0.21** | −0.23** | −0.31** | −0.32** |
| 3 Phonographic errors | 0.7 | 1.9 | – | – | 0.37*** | 0.42*** | 0.61*** |
| 4 Lexical errors | 5.4 | 2 | – | – | – | 0.54*** | 0.85*** |
| 5 Morphological errors | 8.4 | 4.5 | – | – | – | – | 0.86*** |
| 6 Total errors | 14.6 | 8.9 | | | | | |

**Table 4** Correlations between written composition and spelling errors in grades 3 and 4

| | Third grade | | | | | |
|---|---|---|---|---|---|---|
| Fourth grade | 1 Length (words) | 2 Quality (/13) | 3 Phonographic errors | 4 Lexical errors | 5 Morphological errors | 6 Total errors |
| 1 Length (words) | 0.22** | 0.15* | −0.10 | −0.19* | −0.09 | −0.18* |
| 2 Quality (/13) | 0.13 | 0.21** | −0.08 | −0.22 | −0.08 | −0.23* |
| 3 Phonographic errors | −0.05 | −0.02 | 0.57*** | 0.19** | 0.11 | 0.35*** |
| 4 Lexical errors | −0.03 | −0.18 | 0.35*** | 0.45*** | 0.12 | 0.43*** |
| 5 Morphological errors | −0.04 | −0.25* | 0.35*** | 0.12 | 0.20** | 0.46*** |
| 6 Total errors | −0.03 | −0.23* | 0.48*** | 0.43*** | 0.46*** | 0.52*** |

***Longitudinal Correlations for Written Text Measures*** Table 4 reports the correlations between Grade 3 and Grade 4 on the two written text composition measures (length in words and rated quality) and each of three kinds spelling errors (phonographic, lexical, and morphological) as well as total spelling errors. As shown in Table 4, text length for Grade 3 was significantly and positively correlated with text length for Grade 4, and Grade 3 rated quality of text was significantly and positively correlated with Grade 4 rated quality of text. However, text length in Grade 3 was not significantly correlated with text quality in Grade 4.

***Longitudinal Correlations for Kinds of Spelling Errors and Total Spelling Errors*** Phonographic errors in Grade 3 were significantly and positively correlated with phonographic, lexical, and morphological spelling errors and total spelling

errors in Grade 4. Lexical spelling errors in Grade 3 were significantly and positively correlated with phonographic, lexical, and total spelling errors in Grade 4 (but not morphological spelling errors in Grade 4). Morphological spelling errors in Grade 3 were significantly and positively correlated with morphological spelling errors and total spelling errors in Grade 4. Total spelling errors in Grade 3 were significantly and positively correlated with phonographic, lexical, morphological, and total spelling errors in Grade 4.

***Longitudinal Correlations for Spelling Errors (Kinds and Total Number) and Written Text Measures*** The length of the composed text in Grade 3 was not correlated to the total number of spelling errors in Grade 4, but the quality of the composed text in Grade 3 was negatively and modestly correlated with the total number of spelling errors in Grade 4. Producing low quality texts in Grade 3 predicted morphological spelling errors and total spelling errors in Grade 4; these correlations were statistically significant and negative. Those with the higher text quality in Grade 3 produced fewer morphological spelling errors or total spelling errors in Grade 3. Reciprocally, total spelling errors in Grade 3 were significantly but negatively correlated with text length and text quality in Grade 4. Students who produced more spelling errors in Grade 3 wrote shorter texts that were of poorer text quality in Grade 4.

Three regression analyses were conducted. The first one was about whether the different types of spelling errors in Grade 3 predicted the total percent of spelling errors in Grade 4. Only phonographic ($\beta = 0.39$; $t(170, 4) = 6.36$ $p < 0.0001$) and lexical ($\beta = 0.41$; $t(170, 4) = 6.63$ $p < 0.0001$) were significant; and the model explained 39% of the variance. The second one tested whether, once Grade 3 text length was controlled for, the number of spelling errors in Grade 3 would predict text length at the end of Grade 4. The model explained 8% of the total variance. Grade 3 text length had the more important weight ($\beta = 0.21$; $t(171, 2) = 2.91$ $p < 0.005$) followed by the total number of spelling errors ($\beta = 0.17$; $t(171, 2) = -2.29$ $p < 0.03$); there were no significant interactions. The last regression analysis investigated whether, once text quality in Grade 3 was controlled for, the different spelling errors in Grade 3 predicted text quality in Grade 4. The model explained 14% of the total variance: the text quality in Grade 3 had the more weight ($\beta = 0.21$; $t(168, 4) = 2.85 < 0.005$), followed by lexical errors ($\beta = -0.15$; $t(168, 4) = -2.01$ $p < 0.05$). Phonographic errors ($\beta = -0.3$; $t(168, 4) = -1.79$ $p = 0.07$), and morphological errors ($\beta = -0.13$; $t(168, 4) = -1.58$ $p = 0.09$). had no significant contribution. There were no significant interactions.

## 7 General Discussion

Our main objective was to explore in an orthographic system like French, if spelling is related concurrently to written composition within specific grade levels (Grade 3, Grade 6, or Grade 4) and longitudinally across grade levels (Grade 3 and Grade 6 or

Grade 3 and Grade 4). French spelling is interesting because it includes phonographic irregularities (i.e., inconsistencies), lexical difficulties, and numerous morphological silent marks (e.g., plural noun, adjective, and verb agreement). One of the longitudinal analyses was conducted on a medium-sized sample ($N = 79$) followed from Grade 3 to Grade 6; the second longitudinal analysis was conducted on a larger-sized sample ($N = 173$) from the beginning of Grade 3 to the end of Grade 4. In each longitudinal study, pupils were asked in every grade to compose narrative texts either from verbal instructions (Study 1) or from strips (Study 2). Text length, text quality, and three categories of spelling errors were coded and analyzed. Scores in Study 1 were collected from five expert teachers trained to assess children's text productions; this was possible due to the low number of texts to score (about 16 per judge). By contrast, all scores in Study 2 were established by only one judge. The high number of texts also justified using a common frame applied to different topics to prompt children's production. Despite these differences between Studies 1 and 2, the same two important results emerged.

First, both significant concurrent and longitudinal relationships were observed, as in Abbott et al. (2010) for English spelling and composing. Both text characteristics (i.e., length and quality) and total spelling errors were significantly correlated with the corresponding measures longitudinally across grade levels, in Study 1 between Grade 3 and Grade 6, $r = 0.25$, $r = 0.39$ and $r = 0.34$ respectively; and in Study 2 between Grade 3 and Grade 4, $r = 0.22$, $r = 0.21$, and $r = 0.52$ respectively (see Tables 2 and 4 for details). These findings replicated despite different samples, different scoring processes and some differences in study designs. In addition, the phonographic and lexical spelling errors were correlated, but not with morphological spelling errors between Grades 3 and 6 (Study 1); but phonographic, lexical, and morphological spelling errors were correlated between Grade 3 and Grade 4 (Study 2). The nature of morphology learning for spelling may change more from Grade 3 to Grade 6 than from Grade 3 to Grade 4 (Abbott et al., 2016).

Second, when performances in text composition (i.e., length and quality) in Grade 6 (Study 1) and in Grade 4 (Study 2), were regressed on Grade 3's text characteristics and spelling errors, text length and text quality in Grade 3 explained the largest part of variance in text length and quality in Grades 6 or 4, respectively, for the first or second study. Once the autoregressors were entered, total spelling errors contributed significantly and negatively to composition performances. Students who made more spelling errors in grade 3 wrote shorter and lower quality texts in later grades. Further analyses showed that the errors with the most weight were the phonographic and lexical errors. Again, this was an unexpected result insofar as national assessments in France (Andreu & Steinmetz, 2016) have reported evidence that morphological errors are the most frequent and the most troublesome in students' written compositions.

Future research should continue to investigate the concurrent and longitudinal relationships of spelling and written composition in orthographies for different languages as well as the instructional applications of these differences. Whereas normed measures allow comparisons of developing writers to others of the same age or grade on spelling and written compositions skills (Graham et al., 1997), use of

coded spelling errors and written text length and quality, as used in Study 1 and Study 2, is based on the kind of assessment classroom teachers can perform on classroom writing tasks and use to tailor instruction to individual students to optimize their writing development.

## References

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281–298. https://doi.org/10.1037/a0019318

Abbott, R. D., Fayol, M., Zorman, M., Casalis, S., Nagy, W., & Berninger, V. W. (2016). Relationships of French and English morphophonemic orthographies to word reading, spelling, and reading comprehension during early and middle childhood. *Canadian Journal of School Psychology, 31*, 305–321. https://doi.org/10.1177/0829573516640336

Andreu, S., & Steinmetz, C. (2016). *Les performances en orthographe des élèves en fin d'école primaire (1987, 2007, 2015)*. Note d'Information de la DEPP, n°28.

Arfé, B., Dockrell, J. E., & De Bernardi, B. (2016). The effect of language specific factors on early written composition: The role of spelling, oral language and text generation skills in a shallow orthography. *Reading and Writing, 29*, 501–527. https://doi.org/10.1007/s11145-015-9617-5

Babayigit, S., & Stainthorp, R. (2010). Component processes of early reading, spelling, and narrative writing skills in Turkish: A longitudinal study. *Reading and Writing, 23*(5), 539–568. https://doi.org/10.1007/S11145-009-9173-Y

Bahr, R. H., Silliman, E. R., Berninger, V. W., & Dow, M. (2012). Linguistic pattern analysis of misspellings of typically developing writers in grades 1-9. *Journal of Speech, Language, and Hearing Research, 55*, 1587–1599. https://doi.org/10.1044/1092-4388(2012/10-0335)

Berninger, V., & Abbott, S. (2020). *Revised PAL research based reading and writing lessons and reproducible instructional materials*. Pearson.

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing. An Interdisciplinary Journal, 4*, 257–280. https://doi.org/10.1007/BF01027151

Berninger, V., Vaughan, K., Abbott, R., Brooks, A., Abbott, S., Reed, E., et al. (1998). Early intervention for spelling problems: Teaching spelling units of varying size within a multiple connections framework. *Journal of Educational Psychology, 90*, 587–605. https://doi.org/10.1037/0022-0663.90.4.587

Berninger, V. W., Nagy, W., & Beers, S. (2011). Child writers' construction and reconstruction of single sentences and construction of multi-sentence texts: Contributions of syntax and transcription to translation. *Reading and Writing, 24*(2), 151–182. https://doi.org/10.1007/s11145-010-9262-y

Bosse, M.-L., Brissaud, C., & Le Levier, H. (2020). French pupils' lexical and grammatical spelling from sixth to ninth grade: A longitudinal study. *Language and Speech, 64*, 224–229. https://doi.org/10.1177/0023830920935558

Catach, N. (1986). *L'orthographe française*. Nathan.

Catach, N., Gruaz, C., & Duprez, D. (1980). *L'orthographe française: traité théorique et pratique avec des travaux d'application et leurs corrigés*. Nathan.

Daffern, T., Mackensie, N. M., & Hemmings, B. (2017). Predictors of writing success: How important are spelling, grammar and punctuation? *Australian Journal of Education, 61*(*1*), 75–87. https://doi.org/10.1177/0004944116685319

Direction de l'évaluation, de la prospective et de la performance [DEPP]. (2016). Les performances en orthographe des étudiants en fin d'école primaire (1987-2007-2015) [Students' Spelling Performances at the End of Elementary School (1987-2007-2015)]. *Note d'information*, 28.

Dubois, J. (1965). *Grammaire structurale du Français: Nom et pronom*. Larousse.
Fayol, M. (2014). Written language: Learning to read and to spell. In C. Fäcke (Ed.), *Manual of language acquisition* (pp. 162–176). Walter de Gruyter.
Fayol, M. (2016). From language to text: The development and learning of translation. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research, second edition* (pp. 130–143). The Guilford Press.
Gak, V. G. (1976). *L'orthographe du français, essai de description théorique et pratique*. Selaf.
Graham, S., Berninger, V. W., Abbott, R. D., Abbott, S. P., & Whitaker, D. (1997). Role of mechanics in composing of elementary school students. A new methodological approach. *Journal of Educational Psychology, 89*(1), 170–182. https://doi.org/10.1037/0022-0663.89.1.170
Jaffré, J.-P., & Fayol, M. (2005). Orthography and literacy in French. In R. M. Joshi & P. G. Aaron (Eds.), *Handbook of orthography and literacy*. L.E.A.
Joshi, R. M., & Aaron, P. G. (2005). *Handbook of orthography and literacy*. L.E.A.
Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first to fourth grade. *Journal of Educational Psychology, 80*(4), 437–447. https://doi.org/10.1037/0022-0663.80.4.437
Kellogg, R. T. (2001). Competition for working memory among writing processes. *American Journal of Psychology, 114*(2), 175–191. https://doi.org/10.2307/1423513
Kent, S., Wanzek, J., Petscher, Y., Al Otaiba, S., & Kim, Y. S. (2014). Writing fluency and quality in kindergarten and first grade: the role of attention, reading, transcription, and oral language. *Reading and Writing, 27*(7), 1163–1188. https://doi.org/10.1007/s11145-013-9480-1
Khomsi, A., Nanty, I., Parbeau-Guéno, A., & Pasquet, F. (2005). *ECL – Évaluation des compétences linguistiques écrites au collège*. Pearson ECPA.
Kim, Y.-S., Al Otaiba, S., Puranik, C., Folsom, J. S., Greulich, L., & Wagner, R. K. (2011). Componential skills of beginning writing: An exploratory study. *Learning and Individual Differences, 21*(5), 517–525. https://doi.org/10.1016/j.lindif.2011.06.004
Kim, Y.-S., Al Otaiba, S., Folsom, J. S., Gruelich, L., & Puranik, C. (2014). Evaluating the dimensionality of first grade writing composition. *Journal of Speech, Language, and Hearing Research, 57*, 199–211. https://doi.org/10.1044/1092-4388
Kim, Y. S., Al Otaiba, S., Wanzek, J., & Gatlin, B. (2015). Toward an understanding of dimensions, predictors, and the gender gap in written composition. *Journal of Educational Psychology, 107*(1), 79–95. https://doi.org/10.1037/a0037210
Largy, P., & Fayol, M. (2001). Oral cues improve subject-verb agreement in written French. *International Journal of Psychology, 36*, 121–132. https://doi.org/10.1080/00207590143000009
Lucci, V., & Millet, A. (1994). *L'orthographe de tous les jours*. Honoré Champion.
Maki, H. S., Voeten, M. J. M., Vauras, M. M. S., & Poskiparta, E. H. (2001). Predicting writing skill development with word recognition and preschool readiness skills. *Reading and Writing, 14*(7–8), 643–672. https://doi.org/10.1023/a:1012071514719
Manesse, D., & Cogis, D. (2007). *Orthographe: A qui la faute ?* E.S.F.
Martinet, C., Valdois, S., & Fayol, M. (2004). Lexical orthographic knowledge develops from the beginning of reading acquisition. *Cognition, 91*, B11–B22. https://doi.org/10.1016/j.cognition.2003.09.002
McCutchen, D. (2000). Knowledge, processing, and working memory: Implications for a theory of writing. *Educational Psychologist, 35*(1), 13–23.
Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities in French. *Child Development, 76*, 324–339. https://doi.org/10.1111/j.1467-8624.2005.00848_a.x
Rey-Debove, J. (1984). Le domaine de la morphologie lexicale. *Cahiers de Lexicologie, 45*, 3–19.
Richards, T., Aylward, E., Raskind, W., Abbott, R., Field, K., Parsons, A., Richards, A., Nagy, W., Eckert, M., Leonard, C., & Berninger, V. (2006). Converging evidence for triple word form theory in children with dyslexia. *Developmental Neuropsychology, 30*, 547–589. https://doi.org/10.1207/s15326942dn3001_3
Richards, T., Berninger, V., & Fayol, M. (2009). FMRI activation differences between 11- year-old good and poor spellers' access in working memory to temporary and long-term ortho-

graphic representations. *Journal of Neurolinguistics, 22*, 327–353. https://doi.org/10.1016/j.jneuroling.2008.11.002

Seymour, P. H., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94*, 143–174. https://doi.org/10.1348/000712603321661859

Treiman, R. (2018). Teaching and learning spelling. *Child Development Perspectives, 12*(4), 235–239. https://doi.org/10.1111/cdep.12292

Venezky, R. (1995). From orthography to psychology to reading. In V. Berninger (Ed.), *The varieties of orthographic knowledge. II. Relationships to phonology, reading, and writing* (pp. 23–46). Kluwer Academic Publishers.

Veronis, J. (1986). Etude quantitative sur le système graphique et phono-graphique du français. *European Bulletin of Cognitive Psychology, 6*, 501–553. http://www.jstor.org/stable/41557647

# *It Was a Dark and Stormy Sentence*: Teaching the Fine Art of Sentence Construction


**Bruce Saddler**

**Abstract** Learning to compose a variety of effective sentences is an important element within the writing process and one of the first tasks writers take on when learning to communicate in writing. Although creating a sentence may seem simple, for many writers it can be quite challenging. To improve sentence construction skills, writers need to develop an ability to place clear ideas into a variety of sentence formats using a style that is pleasing to the reader. One well-research method to teach sentence construction skills is sentence combining. In this chapter sentence combining is introduced and situated within the writing process. Additionally, the theoretical benefits and practical applications are reviewed along with recent empirical studies. Finally future directions for research are suggested.

**Keywords** Sentence combining · Sentence construction · Writing difficulties · Grammar · Syntax

## 1 Introduction

Once upon a time, when I was a young classroom teacher, I struggled with teaching various aspects of the writing process to my students. The processes of writing, including planning, drafting, and revision seemed far beyond their reach as they were challenged by seemingly basic tasks such as sentence construction. Many of my students wrote only very simple sentences, or sentences connected with a long series of "ands". My experiences taught me that writing is a very tough skill to teach (and to learn) because of all the sub-components involved. I remember often wondering where to begin and where to go with my instruction.

I looked to other teachers at my school for suggestions and then read through the bits of literature on the topic that I could find, without much satisfaction. One day I

B. Saddler (✉)
University at Albany – State University of New York, Albany, NY, USA
e-mail: bsaddler@albany.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_9

remember receiving an advertisement for a writing workshop called "I wrote 10,000 words today," which seemed a lofty goal when compared to the brief segments of text my young writers were creating for me. The title intrigued me and so I went. The presenter was Dr. Steve Graham, then on faculty at the University of Maryland College Park. I sat through his day long workshop taking copious notes about the power of harnessing self-regulation, direct instruction, and strategies to teach the components of writing.

At the end of the workshop I introduced myself to Dr. Graham and shared my desire to learn more about the mysteries of writing instruction. He mentioned an opportunity to come to Maryland as a Ph.D. student with him advising my program of study, and the chance seemed too good to let slip by. I applied and started my studies in the summer of that same year.

As anyone who has earned or is earning a doctoral level degree well knows, one of the main goals is to decide the topic for your dissertation. Since there were many areas of writing that interested me, I had difficulty deciding what to focus on. My initial thoughts centered on grammar instruction as that is what I had utilized in my own teaching. However, Steve, who at that point was my doctoral advisor, suggested I take a look at an older writing intervention called "sentence combining". By that point in my program I had learned two things about Steve: he liked the Peanuts comics by Charles Schultz, and you could trust his ideas about writing completely. I began to explore the literature spurred by Steve's suggestion, finding sentence combining to be a method to teach a variety of writing skills in a way that I had never thought about before. The more I read the numerous research studies, papers, reports, and books written about sentence combining the more I was convinced that teachers needed to incorporate this skill into their writing instruction.

I finished that dissertation. Steve was the chair of my committee and sentence combining my topic. Since graduating, I have continued my work with sentence combining at a research University. During my career I have conducted numerous studies about sentence combining's effects on young writers, received several grants to fund my research, written many journal articles and presented the benefits of sentence combining at conferences and workshops in the U.S. and abroad. But before I begin expanding on my contributions to sentence combining, sentence level interventions need to be situated within the larger arena of the writing process.

## 2 Writing Process

Now as anyone who has researched or taught writing knows, skill with written communication is an exceedingly intricate process to teach effectively and learn competently (Scardamalia & Bereiter, 1986). This has certainly been my experience as both a teacher and researcher of writing.

Yet writing must be taught and learnt during the school years as it is not only a critical skill for students during their school years, but also an important element for success in many careers after graduation. Any writer, from weekend warrior to

Pulitzer Prize winner, must manage a tasks including planning, organizing, revising, and editing of text that may operate recursively, and at times nearly simultaneously, while creating a composition. Of these skills, I believe that composing focused, meaningful, and interesting sentences that deliver well-defined ideas to a reader, can be exceptionally challenging.

Sentence construction is one of the first tasks writers take on when learning to communicate in writing. Young writers run into sentences early in their writing experiences even before entering primary classrooms. In fact, even very young writers start to string words together in an effort to transmit their ideas. These sentences, containing words or word-like representations and scribbled with crayons or painted with finger paints, represent the building blocks of coherent and effective writing, and these first sentences can set a writer on journey towards the ability to effectively construct sentences; a characteristic of expert writing (Beers & Nagy, 2009; Berninger et al., 2011).

Constructing a sentence may seem a simple task, yet consider the many skills and processes involved. In my own writing my sentence begins with a thought as I decide what I am trying to say (my message) and to whom (my audience) I am saying it to. (Right now I am wondering who will be reading this and what they think of my ideas so far!) I might begin by asking what my audience needs to hear on this subject? Then I must decide how to begin the sentence. More questions might arise: What word might work best? What word should come next? Is there a better word that I can use here? Words are tough and require much consideration, for as Mark Twain once wrote, "the difference between the almost right word and the right word is the difference between a lightning bug and lightning." As the sentence grows I must control the words and ideas to make certain there is a logical flow and rhythm. When the sentence is finished I must conclude with some form of punctuation. I must then check the just completed sentence to make certain it actually says what I intended to say and follows the grammatical rules of my language. If not, then I need to revise my work by perhaps adding, changing or taking away words. I might also have to adjust the punctuation or delete the entire sentence and start again. Then the process begins again as I plow into my next sentence. This new sentence must fit logically with the prior sentence and continue the idea being presented while offering new information.

At frequent points I must pause my forward thinking and reflect on what has been produced. While reading back through my existing text I may ask myself: Have I written too many short, choppy similar sentences? Are there fragments, run-ons, or ramblings? Do too many begin with the same word? Are too many of the same length or passively written? Finally, and perhaps most importantly, I must test all of my sentences as a complete composition to ensure they are accurately conveying the intended message I wanted and make corrections as needed just as I did for each individual sentence.

Even in this simplistic scenario of my composing, which is only a shallow representation of the deep cognitive machinations actually going on, sentence construction is clearly a complex process. To improve sentence construction skills, writers need to develop an ability to place clear ideas into a variety of sentence formats, for

example, by expressing their thoughts via mixtures of simple, compound, or complex sentences as appropriate, instead of a series of simple ones connected by coordinating conjunctions.

They must also develop a style of writing that is adapted to the audience and genre. *Style* in writing can be considered a writer's way with words (Nemans, 1995). For example, how a writer decides on the best syntactical arrangements in a given piece of writing relates directly to and reveals their particular style. Style is a writer's "fingerprint" representing their unique interpretation of message creation. Five different writers, if given a particular topic and a specific set of data about that topic, would likely craft five uniquely formed compositions, each with a particular style. That style sets them apart from other writers. This can easily be seen in professional writers whose style can run the gamut from Hemingway's plain and direct wording, to the intensely vivid, expansive prose such as Victor Hugo mastered. These writers each have a singular rhythm and pattern that is obviously effective.

Though it's imperative for any writer to create well-structured sentences, and to develop a signature style in their writing, doing so can be demanding. So much so, in fact, that many writers in the elementary grades may struggle with constructing sentences, and writers who struggle with writing, including writers with learning disabilities (LD), may have considerable difficulties that impact their sentences in several ways when compared to typically achieving writers: their sentences may be shorter, simpler, and less grammatically correct, with capitalization, punctuation, and spelling errors that reduce their overall quality (Graham et al., 2017). Often their stories are brief containing few sentences that may be filled with words of seven letters or less (Houck & Billingsley, 1989). Their sentences may feature word omissions or improper ordering, incorrect verb and pronoun usage or missing subjects and verbs, and a less varied, and many times less effective, vocabulary (Englert & Raphael, 1988; Morris & Crump, 1982). Additionally, the connector "and" is overly relied on when trying to link their thoughts together (Anderson, 1982). Finally, students with LD had lower motivation to write when compared to their typically developing peers (Graham et al., 2017). When added together, it's not unreasonable to believe these factors could make their compositions harder to understand, less enjoyable to read, and less likely to be rated well when scored.

Now for many writers, sentence construction (or syntax) skills so important in creating a variety of effective and engaging sentences, actually can improve with age alone (cf. Hunt, 1965; O'Donnell et al., 1967). However, the gap between the syntactical maturity of writers with disabilities and without appears to increase with age (Andolina, 1980). For many children with disabilities, controlling and manipulating sentences may indeed improve with age, but at a slower pace whereas typically developing writers show periods of rapid improvement during the primary and intermediate grades. Practically speaking, this means that the syntactic maturity of the writings of children with disabilities remains very simple and less effective as written language becomes more important to school success and more syntactically complex (Morris & Crump, 1982). Unfortunately, this performance gap is not closed by time, maturation, or typical school interventions (Christenson et al., 1989; Newcomer & Barenbaum, 1991).

There are several reasons researchers and teachers should improve sentence writing ability with empirically based interventions: First, sentence production difficulties may interfere with planning, content generation, and revising of text because the writer's attention is occupied with lower level skills depleting cognitive resources needed for higher level processes (Graham et al., 2017). Second, a writer lacking knowledge of effective sentence level skills may not be able to translate their ideas into text (Hayes & Flower, 1986), which could reduce the lucidity of the communication. Lastly, difficulties constructing well-designed and grammatically correct sentences may reduce the appreciation of the ideas within the composition by making the material more difficult for others to read.

Researchers have considered the importance of intervening on the sentence level. For example, Graham et al. (2012) suggests that less skilled writers need to develop competence in framing ideas within a variety of sentence formats, for instance, a compound or complex sentence rather than a series of simple ones.

This idea is not a new one, as since the mid 1960s researchers when researchers first suggested that schools prompt students to create more syntactically "mature' sentences. For example in a seminal sentence level writing study by Hunt (1965), schools were advised to facilitate student's ability to create mature writing patterns. Although the recognition of the importance of teaching sentence construction skills has been consistent for many years, how best to teach such skills has undergone many changes. Recently, researchers have suggested that teaching methods to accomplish this could involve direct, motivating and stimulating language experiences that accelerate syntactical pattern development during the school years (Troia, 2014) but this was not always the prevailing wisdom.

During the 1960s and 1970s grammar was the preferred sentence level instructional method in the United States. Such instruction focused direct teaching of parts of speech and sentence types alongside the diagramming of sentences to identify constituent elements. However, a seminal writing meta-analysis by Hillocks (1986) revealed that the widespread study of grammar did not contribute to the quality of a student's writing, or the use of proper mechanics. In addition, grammar taught in a formal manner and removed from actual writing had bored writers, decreasing their desire to write (Jean & Simard, 2011). Dissatisfaction with these outcomes stimulated researchers to advance sentence combining as an alternative method of improving sentence level writing ability.

## 3 Sentence Combining

Sentence combining practice involves explicitly teaching students how to manipulate or rewrite short, syntactically simple sentences into sentences that are more varied in terms of style, as revealed in the arrangement of words and word choice, and complexity, represented by length and syntactic structure (Saddler, 2009). For example, a series of simple sentences a young writer might produce such as: "The sky was blue. The sky was sunny. The sky was beautiful." could be combined in

differently, for example: "*The sunny blue sky was beautiful*". Or, "*The beautiful sky was sunny and blue*". Or "*The sky, which was sunny and blue, was beautiful*". The modest act of playing with the options available in their syntax can help writers think about the sound and substance of their language and how they may best represent their ideas.

Sentence combining is derived directly from Chomsky's transformational-generative grammar theory (1957). Chomsky believed traditional Latinate grammar inadequate for describing the intricacies found in the English language. So he proposed a language theory he believed adequately explained syntactical structure formation that contained transformational rules working alongside existing grammatical rules.

This new theory, transformational-generative grammar, purported that language can be modeled through mathematical-like formulas that govern it through a finite set of rules for sounds, word formation, and syntax. Through these rules, Chomsky attempted to explain how language users can create complex sentences from just a few basic (or kernel) sentence patterns; the kernel sentences being the spoken or written expressions of the essential thoughts of the writer. In the English language, kernel sentences are simple sentences (S) consisting of a subject (NP) and a predicate (VP). In mathematical terms, this arrangement could be represented as $S = NP + VP$. Chomsky believed that typical sentences in written texts consist of many basic kernels, each adding to the general gist of the text.

Sentence combining researchers embraced Chomsky's theories, basing their practice on the notion that the sentences we actually speak or write are derived from these kernel sentence units. In both transformational-generative grammar and sentence combining, the kernel states can be considered the "deep structure" or the underlying idea for the "surface structure" that is created or generated when kernels are "transformed" or changed. A writer analytically combines these minimal sentences to create a variety of possible sentences best fitting their goals. For example, the sentence "I quitted my seat, and walked on, although the darkness and storm increased every minute, and the thunder burst with a terrific crash over my head." (Shelley, 2003, p. 68) represents the surface structure of a deep structure that began in Mary Shelley's mind, perhaps in the form of "kernels" or basic sentences such as these:

I quitted my seat.
I walked on.
The darkness increased.
The storm increased.
The increase was by the minute.
The thunder burst.
It burst with a crash.
The crash was terrific.
The thunder was over my head.

These kernels could be transformed in many ways, for example:

- With the darkness and storm increasing every minute, I quitted my seat and walked on though the thunder burst with a terrific crash over my head.
- As the terrific crash of thunder burst over my head and the darkness and storm increased by the minute, I quitted my seat and walked on.

There is reason to believe that professional writers engage in this same practice as they think of different ways to say something. James Joyce, for example, would walk the streets while mentally rearranging the words of a single sentence until he found the right combination. If reorganizing the deep structure of text is at the heart of sentence combining practice and, as the theory goes, at the heart of the writing process itself, then SC practice may represent the same types of activities real writers engage in when composing.

Sentence combining practice allows writers to work through problems they may encounter while creating and manipulating text during writing. When writers, such as James Joyce, compose sentences and paragraphs, they make decisions about vocabulary and sentence structure, the arrangements of words within their sentences, adding, subtracting or rearranging as needed. It is unlikely they are only trying to create longer sentences; therefore, that should not be the end goal of SC practice either. Instead the end goal of SC should be to in many cases to say more in fewer words (Strong, 1986) and to say those words in the best way possible. To reach this goal, SC practice should prompt students to consider syntactic options in their writing through practice in consciously controlling and manipulating syntax. Through this practice students can learn that improvements in the clarity and variety of their sentences can help improve the overall quality of their compositions rather than the length alone. Through practicing various sentence combinations, students learn that they can choose or "judge", whether the sentences should be combined, and if they are to be combined, how best to combine them.

The reality is that what is usually needed in any writing is a variety of sentences. The writer could add shorter sentences for the impact brevity can provide. Longer sentences could be used for the conveyance of complex ideas that require the writer to create a more comprehensive explanation. SC practice helps students come to understand when sentences might be better left short, or might be more effective if made longer. In other words, SC practice can help writers understand the choices available to them at the sentence level (Saddler & Asaro-Saddler, 2010).

Although these may seem to be routine acts, they are not for many writers. Such writers may need to hear and read many sentence constructions that would not initially come to their mind while writing. Unfortunately though, hearing and reading are not enough; they also need to physically try out the sentence forms they are taking in by actively manipulating the syntax in writing (Moffett, 1968). According to Moffett, only a comparison of sentence alternatives—in the context of what the author is trying to accomplish will teach this type of judgment. Following this line of reason then it's only through comparing various combinations of the same sentences that students can sense the slight changes to meaning various sentence combinations have and the effect those sentence changes may make on the

overall meaning of the composition. There is no writing activity better suited than sentence combining to accomplish this.

Combining sentences in frequent sessions where the exercises are carefully modeled and openly discussed can expose writers to a variety of syntactic structures they can utilize while composing or revising to convey their ideas more effectively. In addition, such practice can also provide writers with a systematic method to explore language without the need to generate content/ideas, thus reducing some of the cognitive burden associated with the composing process. The exercises can provide a venue for playing with words and ideas while also providing focused, interesting language experiences. In comparison to grammar instruction, such practice represents applied use of syntax rather than knowledge about syntax and grammatical rules.

## 4 Theoretical Principles

The benefits of sentence combining practice are based on three theoretical principles. First, writers need instruction in formulating a clear understanding of the written sentence, the limits of the simple sentence, and the syntactic options that are available to them when crafting sentences (Neuleib & Fortune, 1985). SC practice can help writers learn and practice the language choices available to them (Saddler & Preschern, 2007). Although real writers may not typically write down strings of kernel sentences and then combine them as students will do while practicing SC, writers will draft a version, then adjust and revise that version, perhaps several times, before finally coming close to their final message. Students can use sentence combining practice to mimic this process of revision and adjustment by systematically trying out various sentence forms and exploring the options and syntactical alternatives available to them at the sentence level in their writing (Saddler & Preschern, 2007).

Sentence combining exercises, by presenting students with sets of simple sentences and requiring them to combine those sentences any way they wish as long as the product is grammatical, allows for the possibility of more than one "right" answer. This outcome, the awareness of the existence of acceptable alternatives, should be encouraged and can become itself a subject of instruction (Nutter & Safran, 1984). SC allows writers to practice the delicate art of adjusting the placement of information in their sentences to better transmit their message to the audience. Sentence combining practice also encourages writers to tighten and clarify their thoughts by de-combining lengthy sentences, or by rearranging, elaborating, or editing parts of sentences while also varying the sentence patterns appearing in paragraphs.

Second, once writers become more comfortable and fluent with the process of sentence formation and re-formation, the overall cognitive strain of writing may be decreased (Graham, 1982), and the student can free up cognitive space (i.e. working

memory) for higher level functions including navigating the processes of writing, considering the audience's needs, and the overall goals of the assignment.

We use working memory to store and process information, however working memory has a limited capacity, meaning any activity that makes demands on working memory may interfere with another activity. In writing, if an activity such as handwriting, spelling, punctuation, or sentence construction for example occupies too great an amount of working memory, the writer has little "cognitive capital" left to use on other functions. This may have a direct impact on the writers' ability to fluently produce ideas. However, as writers mentally and operationally manipulate syntactic structures through frequent SC practice, they may in essence make their syntactic skills more automatic, and in doing so freeing mental energy (Saddler et al., 2018; Stotski, 1975).

Third, gains in syntactic fluency lead to quality writing (Strong, 1986) and research suggests that SC helps increase syntactic maturity or fluency in writers (e.g., Hunt, 1965). The basic theory behind syntactic fluency is that skilled writers' produce quality writing and one possible evidence of writing skill is syntactic fluency. This theory is supported by evidence suggesting raters tend to score writing with a high degree of syntactic maturity as superior in quality (Phillips, 1996). In addition, syntactic complexity may also elicit more favorable responses from competent readers (de Beaugrande, 1985). Therefore, as the theory goes, if a writer makes gains in syntactic fluency those gains may contribute to quality writing (Strong, 1986).

## 5 Research Base

Sentence combining has a lengthy research base. In fact, over 90 studies conducted during the last 6 decades have demonstrated that sentence combining is an effective method for helping students produce more syntactically complex sentences (cf. Gale, 1968; Hunt, 1965; Mellon, 1969; O'Hare, 1973) and may improve the overall quality of compositions (cf. Combs, 1975; Perron, 1974).

Sentence combining researchers have focused on two essential questions: First, what effect does sentence combining have on a writers' syntactical maturity? Second, can sentence combining improve the overall quality of writing?

Research indicates sentence combining practice improves the syntactical maturity of children at nearly all grade levels. Hillocks (1986) reported that over 60% of the sentence combining studies performed between 1973 and 1985 produced significant results on measures of syntactical maturity while 30% reported improvements at a non-significant level and only 10% demonstrated no improvements or mixed results. As Kerek et al. (1980) stated, sentence combining "has been proven again and again to be an effective means of fostering growth in syntactic maturity" (p. 1067). However, research has not provided a definitive answer to the second question. Although researchers attained significant increases in writing quality after

sentence combining practice (e.g., Saddler & Asaro, 2008), others found non-significant or mixed results (Hillocks, 1986).

Three important publications provide additional support for SC. First, an extensive analysis of grammar and sentence combining literature in England conducted by the English Review Group (Andrews et al., 2006) concluded that the extensive 45 year history of sentence combining research suggests that sentence combining has been proved to work and should be considered as an important element in a repertoire of activities, especially for 7–14 year olds, where most of the research has been conducted. Furthermore, in Writing Next: A Report to the Carnegie Corporation (Graham & Perin, 2007), sentence combining had a consistently positive and moderate effect on writing and was listed among the practices recommended for inclusion within effective writing programs. These findings echoed the same conclusion of Hillocks (1986) when he stated "research shows sentence combining, on the average, to be more than twice as effective as free writing as a means of enhancing the quality of student writing. If we want our students to become conscious of writing style and actually to improve their own, I cannot recommend any pedagogy more enthusiastically than sentence-combining exercises. Research shows that sentence combining improves the overall style of the writing of most students who practice it" (Hillocks, 1986). Additionally, Graham et al. (2015) conducted a meta-analysis on effective writing practices and concluded that teaching sentence construction skills through sentence combining improved the writing of students in grades 4–7, with an effect size of 0.56. Together, these three studies suggest that SC instruction yields positive outcomes.

Unfortunately, most of the research of sentence combining as an instructional method is rather dated—mostly being from the 1960s to 1970s—and limited to the upper grade levels. Encouragingly, there have been important attempts to update and expand the literature base within the last 20 years. Three studies with younger writers provides further support for sentence combining practice while extending prior research in several important ways.

**Saddler and Graham (2005)** Saddler and Graham (2005) were the first to explore the effects of sentence combining practice using a peer assisted grouping arrangement versus traditional grammar instruction. In this study 44 students were selected from all fourth-grade students in 9 classrooms within three schools in the Washington D.C. Metropolitan Area. The students were identified as skilled or less skilled writers based on the Test of Written Language, 3rd edition (TOWL-3) results and teacher report. Students were paired and assigned a sentence combining or grammar instructional condition. Participants in each condition received 30 lessons, 25 min in duration, three times per week for 10 weeks.

Instruction was scaffolded via explanation and modeling in both groups with the SC group practicing four sentence combining procedures and the grammar group several parts of speech. The experimental group was first taught how to combine sentences using the conjunctions *and*, *but*, and *because*, then adjectives and adverbs, and finally adverbial and adjectival clauses. The instructor explained and modeled performing particular sentence combinations prior to the guided practice phase.

During guided practice, students were placed in dyads (one stronger writer, one weaker writer) and worked on writing their own solutions to the exercises. During the independent practice phase, the pair wrote a short story and revised pieces utilizing the sentence combining strategy to help generalize the sentence combining skill being learned. Using the peer assisted grouping arrangement, one student assumed the role of the coach while the other student applied the strategy during both the composition and revision process.

The grammar instruction component concentrated on skills associated with nouns, verbs, adjectives, adverbs, subjects, and predicates. The instructor modeled and explained how to appropriately apply a part of speech in a sentence. The students worked in pairs during guided practice completing exercises that required supplying of a missing part of speech in a sentence. The grammar instruction pairs also engaged in writing a short story and revising their pieces utilizing in a manner identical to the experimental group.

To assess treatment effects participants wrote and revised a story and completed sentence combining progress monitoring measures during the interventions. Analysis revealed a statistically significant effect of sentence combining instruction on all progress monitoring and standardized assessment measures. Participants in the sentence combining condition were twice as likely to combine two or more sentences into a semantically and syntactically correct single sentence than those students in the grammar instruction condition. Effects in the area of revision were more modest but moderate and still indicative of a higher level of achievement for the students in the sentence combining condition. Additionally, for the students in the sentence combining condition, post-test story quality improved to some extent in this study whereas in the grammar condition it was unchanged.

**Saddler et al. (2008b)** A replication single subject design study was conducted in 2008 by Saddler, Behforooz and Asaro. In this study students with and without documented learning disabilities were provided with sentence combining instruction with a peer assistance component. Six students were included in the study: three had an identified learning disability and three were typically achieving, with all described as having weak writing skills by their classroom teachers. Students were randomly assigned to instructional pairings.

This study extended Saddler and Graham (2005) in two ways. First, exercises were included to assist in generalization of the sentence combining skills from sentence creation to paragraph writing. Secondly, less skilled writers were paired with more skilled writers to assess the ability of the less skilled writer to learn the skills from a more able peer. Each student pair received 18 lessons, 25 min in duration, separated into three units of instruction, three times a week for 6 weeks.

Lessons contained similar procedures to Saddler and Graham's (2005) study, however two lessons in this study focused on generalization. In the third lesson, the students were asked to apply the sentence combining skills they had been practicing in isolated exercises to a paragraph revising transfer task. The fourth lesson provided the students with a topic and several random facts written as phrases. The students

were instructed to take those phrases and combine them into sentences to form a cohesive paragraph.

Treatment effects were measured using a sentence-combining task, story quality (based on a rubric), writing complexity, and frequency of use of taught sentence combining constructions in connected text. Results indicated that each student improved their ability to combine sentences at a very significant level of effectiveness. In addition, all of the students improved the quality of their stories and the complexity of their sentences.

**Saddler et al. (2008a)** The effect of a sentence combining and peer assisted grouping strategy was studied further in another single subject design by Saddler et al. (2008a). As in Saddler et al. (2008b), sentence-combining practice with a peer assistance component was employed with the addition of a parallel writing task, a Peer-Editor Checklist to improve the transfer of the sentence-combining skills to story writing, and the collection of anecdotal evidence regarding the efficacy of the intervention.

Four grade four students with learning disabilities and writing difficulties participated in paired dyads. Sessions lasting for 35 min occurred three times per week, with each pair of writers receiving 18 lessons. The sentence-combining curriculum was adapted from the Saddler and Graham (2005) study and involved rewriting short, kernel sentences using three units of six lessons each. Each unit contained a particular type of sentence construction including adjectives, phrase insertions and the use of connectors, "but" and "because" to combine sentence kernels. In addition, each unit also included a parallel writing piece that was peer-revised using a checklist.

Measures used to document instructional effects included sentence-combining ability, story quality, number and quality of revisions, and instances of taught sentence-combining constructions in connected text. Results indicated that all students improved their sentence-combining ability and the quality of their stories. The number of revisions attempted improved, however the overall story quality did not improve as a result of the attempted revisions. All writers included more of the taught sentence combining constructions in their compositions then were present at pretest.

Direct practice improved sentence-combining ability in this study as in previous sentence-combining studies, however, unlike previous studies, in this study the taught constructions appeared to a greater extent in the post test stories than at pretest. Additionally, the quality of the posttest stories improved for all of the writers and revising behavior increased, yet did not lead to improving quality of second drafts. Finally, in this study, unlike the previous two, anecdotal data collected about the students' perception of the strategy revealed that the overall steps of making sure sentences made sense was seen as beneficial and that the instruction in general was enjoyable. Student comments also suggested that the Peer-Editor Checklist increased their ability to identify the use of sentence-combining skills in each other's writing, and make effective revising suggestions as they progressed through the intervention.

The results of these studies add support for the use of sentence combining as a method to improve sentence construction ability. Furthermore, these studies demonstrate that sentence combining is effective with young writers at various ability levels, including writers with disabilities, and may favorably impact the quality of compositions and quantity of revisions. This finding is especially significant since early intervention using effective instruction may maximize the writing development of all children, reduce the number of students who develop writing problems from poor instruction, and lessen the difficulties and motivational challenges experienced by children with writing disabilities (Graham & Harris, 2004). These studies also reveal that peer grouping can be an effective instructional arrangement during sentence combining practice. Finally, in all of these studies, the intervention lesson structure was arranged following the principles of Self-Regulated Strategy Development (SRSD; Graham & Harris, 2005) including direct instruction, modeling, shared and independent practice and generalization.

Encouragingly, since the publication of these three studies, several other researchers have begun to test the potential of sentence combining; for example:

First, Limpo and Alves (2013) examined the effectiveness of two strategy-focused interventions on the opinion essay writing ability of fifth and sixth graders students. Two instructional groups received either planning or sentence-combining instruction over 12 weekly 90-min lessons. Lessons followed the Self-Regulated Strategy Development (SRSD) model. The intervention groups were compared with a practice control group that received their typical writing instruction. Results indicated that when compared to the control group, the experimental group's instruction improved planning, sentence-combining instruction, opinion essay quality and text length. A positive correlation between self-efficacy and writing quality was found for both intervention groups; and planning instructional effects generalized to summary writing.

Second, Lee and Lee (2016) examined effects of sentence manipulation activities on reading comprehension finding that sentence level practice enhanced syntactic knowledge and reading comprehension on English proficiency tests.

Third, Balthazar and Scott (2018) investigated the effects of a sentence treatment on the language performance of 10–14 year old children with specific language impairment. Treatment effects were measured by sentence probes administered at baseline, treatment, and posttreatment phases, and comparisons of pre-post performance on oral and written language tests and tasks. The treatment improved performance on the sentence probes for the majority of participants, with the largest effect sizes found for adverbial and relative clauses. Treatment gains were discovered on a comprehensive oral language test, but not on reading and writing measures. The researchers suggest that a focused intervention can produce improvements in complex sentence productions of older school children with language impairment.

Fourth, Goodrich et al. (2020) investigated whether sentence-combining instruction was effective for improving writing outcomes of Spanish-speaking language-minority (LM) students with poor sentence construction skills. Seven Spanish-speaking LM children in third to fifth grade participated in sentence-combining lessons to teach adjective placement. Results revealed a functional

relation between sentence-combining instruction and student performance on sentence-writing probes. In addition, improvement in number of correct writing sequences on the sentence-writing probe was found.

Fifth, Telesca et al. (2020) examined the impact of sentence combining with an explicit metalinguistic approach in comparison to typical science instruction on written expression and understanding of comparison/contrast in science for eighth-grade students who struggled with literacy. The researchers found that their treatment enhanced the experimental group's score in listing similarities and differences between two science concepts on a graphic organizer, however no significant differences between the two groups in their use of syntactic factors typical of academic text when responding to a science compare and contrast writing prompt were noted.

Finally, Walter et al. (2021) investigated the effectiveness of a sentence-combining intervention as compared with a spelling intervention and a business as usual control condition. Participants were 7–10 year old struggling writers, several of whom also performed poorly on measures of reading and oral language. Children receiving the sentence-combining intervention showed significant improvements in the sentence combining measure at post-test and delayed post-test compared to those receiving the spelling intervention and controls. Regression analyses revealed that participants in the sentence-combining intervention, with a low sentence combining score and low reading skills at baseline (or better baseline spelling skills), were more likely to show improvements at post-test. The researchers suggested that when devising interventions for struggling writers, specific profiles of skills should be considered. Specifically, sentence combining may be more appropriate for those students whose primary area of difficulty is reading, rather than poor spelling or oral language.

These recent studies add interesting and valuable contributions to the existing SC literature base by extending the reach of the research to language minority learners, metalinguistic approaches, content areas such as science, reading comprehension, and different writing genres including opinion essays. Encouragingly, several of the studies included elements of SRSD, which has been proven to be a potent method of arranging instruction for children with disabilities (Graham & Harris, 2018).

## 6 Research Supported Benefits

The sentence combining literature, when considered as a whole, indicates it has the potential to enhance writing abilities in several meaningful ways:

First, SC practice can raise a writer's awareness that syntactic options exist in their writing and promote their willingness to experiment with a wider range of those options thereby improving linguistic performance by introducing students to sentence options not within their familiar repertoire (Strong, 1986). Through systematic practice opportunities combining sentences into different arrangements, the syntactic patterns may become more familiar to the writer (de Beaugrande,

1985) and greater familiarity with syntactic options may boost a students' confidence in their ability to manipulate sentence syntax, making them more willing to vary, experiment, and innovate in their writing.

Second, by revising practice kernel sentence clusters and considering the effectiveness of the new constructions, SC exercises may help writers to begin to consider how different syntactic options may impact readers (Strong, 1976; Neuleib & Fortune, 1985). Hopefully, through this process of making judgments about their sentences and how the sentences will be received by an audience, they begin to consider how their writing sounds from the readers' perspective.

Third, discussing sentence combining exercises may also help students become confident about punctuation, as the exercises can reveal how punctuation organizes sentence elements in situ (Lindemann, 1995). The exercises could prompt discussions between writers regarding the types of punctuation needed in a particular sentence. For example, a comma may be required especially if a coordinating conjunction (for, and, nor, but, or, yet, so) is used as a connector in a new construction. However, a comma would not be used if two sentences are combined without a connector (i.e. The girl was happy; the boy was unhappy.). A comma could also be used after most introductory elements especially long ones (i.e. In the grey house with the wide front porch, the boy sat and hoped for time to pass quickly), but can be left out if the intro is short (i.e., In time the day became sunnier and happier for everyone.) or the writer chooses not to emphasize the beginning of the sentence.

Fourth, sentence combining practice may foster skill in revision. Revising has been called a process of transforming sentences (Elbow, 1985). SC can teach basic revision skills such as expansion or reordering of ideas, and tightening of language (Strong, 1985) and therefore could conceivably be most valuable in the revising stage of the writing process. Furthermore, Hillocks (1986) suggests that sentence combining might be most effective during the revision process as it offers writers practice with alternative syntactic structures they can select for their own compositions.

Improving revising skill is important for any writer—but especially so for young writers because revising can be so challenging—that many writers avoid the process or adopt a least effort strategy by first making changes to the elements in the sentence that are easiest to change (Hunt, 1983). Likely this will be only surface corrections such as word changes or punctuation adjustments followed by slight elaborations (i.e. adding a descriptive word) and deletions (c.f. Nodine et al., 1985). Their last choice, and the most cognitively challenging one, will be a partial or complete restructuring of their language. One reason for this may be that when revising, children tend to avoid tampering with what they call "basic sentence plans" meaning the ideas they initially set down in the form of sentences (Scardamalia & Bereiter, 1986; Hillocks, 1986). But even when writer's try to change the basic plan, often their revisions are not effective, because the original version of the text, present on the page, represents so strong a stimulus that it can literally block or inhibit the ability of the writer to think and create a new way to say the same idea (Scardamalia & Bereiter, 1986).

When a writer can break from the existing text, if the only alternatives they consider are those that come spontaneously to mind, revision may have little chance of success. A writer must deliberately bring alternatives to mind, or the original text will win for lack of competition (Scardamalia & Bereiter, 1986). Even very skilled writers, ones who can move beyond surface corrections and who display a readiness to take an audience into account in their syntactic choices, may not consider the relative value of various choices on their message. As a result, although they may value the rhetorical impact of a particular choice, they settle too quickly and too adamantly on a particular choice without trying other possibilities (Neuleib & Fortune, 1985).

For writers who make few revisions, or who settle on first choices without considering other syntactic options, gaining mindful access to syntactical alternatives becomes essential so they might transcend the original text on the page (Scardamalia & Bereiter, 1986). Scardamalia and Bereiter (1986), Hillocks (1986), and Hunt (1983) suggest that sentence combining practice might provide a hierarchically organized and systematic knowledge of syntactic structures that might enable writers to thoughtfully consider alternatives in sentence structures.

## 7 Future Directions of Sentence Combining Research

Although the studies mentioned here are important contributions in updating and expanding the research base supporting sentence combining, there is still much work left to do as there are many things we need to know about sentence combining. First, investigators have not examined the possibility of the cognitive load reduction hypothesis for sentence combining instruction. If this hypothesis is valid, a reduction in cognitive load should follow sentence combining instruction. For example, the amount of time students might take to react to a secondary-task (e.g., an audible signal) occurring while they construct sentences should be reduced. Likewise, as students improve their fluency with combining sentences, a proportional increase in time committed to other writing processes should occur (see Rijlaarsdam & van den Bergh, 1996). Secondly, we do not know yet how best to transfer sentence combining skills to connected writing. Therefore, future studies need to explore generalization as a primary goal. Thirdly, the meta-cognitive aspects of this instruction need to be explored since we don't yet know how a student approaches a sentence combining task cognitively, or how such instruction may help reorganize their existing schema regarding the act of constructing sentences. Fourthly, the interactions between pairs of writers as they approach sentence combining tasks should be recorded and analyzed to gain understandings of the cognitive impact acquiring sentence combining skills has and in how peer interactions impact that process. Fifth, future research should include participants with other disabilities, including emotional and behavioral and autism spectrum disorders. Sixth, interventions could include reading and vocabulary instruction, since both could impact sentence writing ability. Seventh, sentence combining's impact on revision needs to be explored to a greater

degree. We know very little about how SC impacts this area of the writing process. Finally, likewise, we need to expand knowledge of the utility of SC in content areas and with a variety of writing tasks across genres.

Clearly although much has been explored, there is much work left to do. Teaching the fine art of sentence construction is a complex and messy, yet essential, undertaking. In my own career, I can say with all honesty that the modest contributions I have made to the literature are a direct result of the prompting of Dr. Graham to explore sentence combining. I owe any success I have had to Steve and his suggestion to take a look at an old method of teaching writing called sentence combining. So I leave you with a set of kernels to practice combining:

The Doctor was wise.
The Doctor was helpful.
The Doctor gave advice.
The advice was good.
The advice was about sentence combining.
The Doctor's name was Graham.

## References

Anderson, P. L. (1982). A preliminary study of syntax in the written expression of learning disabled children. *Journal of Learning Disabilities, 15*, 359–362.

Andolina, C. (1980). Syntactic maturity and vocabulary richness of learning disabled children at four age levels. *Journal of Learning Disabilities, 13*, 27–32.

Andrews, R., Torgerson, C., Beverton, S., Freeman, A., Locke, T., Low, G., Robinson, A., & Zhu, D. (2006). The effect of grammar teaching on writing development. *British Educational Research Journal, 32*, 39–55.

Balthazar, C. H., & Scott, C. M. (2018). Targeting complex sentences in older school children with specific language impairment: Results from an early-phase treatment study. *Journal of Speech, Language, and Hearing Research, 61*(3), 713–728. https://doi.org/10.1044/2017_JSLHR-L-17-0105

Beers, S. F., & Nagy, W. E. (2009). Syntactic complexity as a predictor of adolescent writing quality: Which measures? Which genre? *Reading & Writing, 22*(2), 185–200.

Berninger, V., Nagy, W., & Beers, S. (2011). Child writers' construction and reconstruction of single sentences and construction of multi-sentence texts: Contributions of syntax and transcription to translation. *Reading & Writing, 24*(2), 151–182.

Chomsky, N. (1957). *Syntactic structures*. Mouton & Co.

Christenson, S. L., Thurlow, M. L., Ysseldyke, J. E., & McVicar, R. (1989). Written language instruction for students with mild handicaps: Is there enough quantity to endure quality? *Learning Disability Quarterly, 12*, 219–229.

Combs, W. E. (1975). Some further effects and implications of sentence combining exercises for the secondary language arts curriculum. (Doctoral dissertation, University of Minnesota, 1975). *Dissertation Abstracts International, 36*, A1266.

de Beaugrande, R. (1985). Sentence combining and discourse processing: In search of a general theory. In D. A. Daiker, A. Kerek, & M. Morenberg (Eds.), *Sentence combining: A rhetorical perspective* (pp. 61–75). Southern Illinois University Press.

Elbow, P. (1985). The challenge for sentence combining. In D. A. Daiker, A. Kerek, & M. Morenberg (Eds.), *Sentence combining: A rhetorical perspective* (pp. 232–245). Southern Illinois University Press.

Englert, C. S., & Raphael, T. E. (1988). Constructing well-formed prose: Process, structure, and metacognitive knowledge. *Exceptional Children, 55*, 513–520.

Gale, I. F. (1968). An experimental study of two fifth grade language arts programs (Doctoral dissertation, Ball State University, 1968). *Dissertation Abstracts International, 28*, A4156.

Goodrich, J. M., Hebert, M., Savaiano, M., & Andress, T. T. (2020). Effects of sentence-combining instruction for Spanish-speaking language-minority students. *Elementary School Journal, 120*(4), 715–749.

Graham, S. (1982). Composition research and practice: A unified approach. *Focus on Exceptional Children, 14*, 1–16.

Graham, S., & Harris, K. R. (2004). Writing instruction. In B. Wong (Ed.), *Learning about learning disabilities* (3rd ed., pp. 281–313). Academic.

Graham, S., & Harris, K. (2005). *Writing better*. Brookes.

Graham, S., & Harris, K. R. (2018). Evidence-based writing practices: A meta-analysis of existing meta-analyses. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Design principles for teaching effective writing: Theoretically and empirically grounded principles* (pp. 13–37). Brill.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools* (A report to Carnegie Corporation of New York). Alliance for Excellent Education.

Graham, S., Bollinger, A., Booth Olson, C., D'Aoust, C., MacArthur, C., McCutchen, D., & Olinghouse, N. (2012). *Teaching elementary school students to be effective writers: A practice guide* (NCEE 2012-4058). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications_reviews.aspx#pubsearch

Graham, S., Harris, K. R., & Santangelo, T. (2015). Research-based writing practices and the common core: Meta-analysis and meta-synthesis. *The Elementary School Journal, 115*, 498–522.

Graham, S., Collins, A. A., & Rigby-Wills, H. (2017). Writing characteristics of students with learning disabilities and typically achieving peers: A meta-analysis. *Exceptional Children, 83*, 199–218.

Hayes, J. R., & Flower, L. S. (1986). Writing research and the writer. *American Psychologist, 41*, 106–113.

Hillocks, G. (1986). *Research on written composition*. ERIC Clearinghouse on Reading and Communication Skills.

Houck, C. K., & Billingsley, B. S. (1989). Written expression of students with and without learning disabilities: Differences across the grades. *Journal of Learning Disabilities, 22*, 561–575.

Hunt, K. W. (1965). *Grammatical structures written at three grade levels* (Research Report No. 3). National Council of Teachers of English.

Hunt, K. W. (1983). Sentence combining and the teaching of writing. In M. Martlew (Ed.), *The psychology of written language: Developmental and educational perspectives*. Wiley.

Jean, G., & Simard, D. (2011). Grammar teaching and learning in L2: Necessary, but boring? *Foreign Language Annals, 44*, 467–494.

Kerek, A., Daiker, D. A., & Morenberg, M. (1980). Sentence combining and college composition. *Perceptual and Motor Skills, 51*, 1059–1067.

Lee, H., & Lee, J. H. (2016). Reassessment of sentence combining and decombining; effects of sentence manipulation activities on reading comprehension. *English Teaching: Practice & Critique, 16*(1), 55–71.

Limpo, T., & Alves, R. (2013). Teaching planning or sentence-combining strategies: Effective SRSD interventions at different levels of written composition. *Contemporary Educational Psychology, 38*, 328–341.

Lindemann, E. (1995). *A rhetoric for writing teachers*. Oxford University Press.

Mellon, J. C. (1969). *Transformational sentence combining: A method for enhancing the development of syntactic fluency in English composition*. National Council of Teachers of English.
Moffett, J. (1968). *Teaching the universe of discourse*. Houghton Mifflin.
Morris, N. T., & Crump, W. D. (1982). Syntactic and vocabulary development in the written language of learning disabled and non-learning disabled students at four age levels. *Learning Disabilities Quarterly, 5*, 163–172.
Nemans, B. S. (1995). *Teaching students to write*. Oxford University Press.
Neuleib, J., & Fortune, R. (1985). The use of sentence combining in an articulated curriculum. In D. A. Daiker, A. Kerek, & M. Morenberg (Eds.), *Sentence combining: A rhetorical perspective* (pp. 127–137). Southern Illinois University Press.
Newcomer, P. L., & Barenbaum, E. M. (1991). The written composing ability of children with learning disabilities: A review of the literature from 1980 to 1990. *Journal of Learning Disabilities, 24*, 578–593.
Nodine, B. F., Barenbaum, E., & Newcomer, P. (1985). Story composition by learning disabled, reading disabled, and normal children. *Learning Disability Quarterly, 8*, 167–179.
Nutter, N., & Safran, I. J. (1984). Improving writing with sentence combining exercises. *Academic Therapy, 19*, 449–455.
O'Donnell, R. C., Griffin, W. J., & Norris, R. C. (1967). *Syntax of kindergarten and elementary school children: A transformational analysis* (Research Rep. No. 8). National Council of Teachers of English.
O'Hare, F. (1973). *Sentence combining*. National Council of Teachers of English.
Perron, J. D. (1974). An exploratory approach to extending the syntactic development of fourth grade students through the use of sentence-combining methods (Doctoral dissertation, Indiana University, 1974). *Dissertation Abstracts International, 35*, A4316.
Phillips, S. E. (1996). *Sentence combining: A literature review*. Eric Document Reproduction Service No. ED 398 589.
Rijlaarsdam, G., & van den Bergh, H. (1996). The dynamics of composing – An agenda for research into an interactive compensatory model of writing: Many questions, some answers. In M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications*. Mahwah, NJ.
Saddler, B. (2009). Sentence combining. *Perspectives on Language and Literacy, 3*, 27–29.
Saddler, B., & Asaro, K. (2008). Beyond noun-verb: The use of sentence combining to improve sentence writing ability. *Insights on Learning Disabilities, 5*(2), 41–50.
Saddler, B., & Asaro-Saddler, K. (2010). Writing better sentences: The effects of sentence combining practice. *Preventing School Failure, 54*(3), 159–163.
Saddler, B., & Graham, S. (2005). The effects of peer-assisted sentence combining instruction on the writing of more and less skilled young writers. *Journal of Educational Psychology, 97*(1), 43–54.
Saddler, B., & Preschern, J. (2007). Improving sentence writing ability through sentence-combining practice. *Teaching Exceptional Children, 39*(3), 6–11.
Saddler, B., Asaro, K., & Behforooz, B. (2008a). The effects of peer assisted sentence combining practice on four writers with learning disabilities. *Learning Disabilities: A Contemporary Journal, 6*, 17–31.
Saddler, B., Behforooz, B., & Asaro, K. (2008b). The effects of sentence combining instruction on the writing of fourth grade students with learning disabilities. *Journal of Special Education, 42*, 79–90.
Saddler, B., Ellis-Robinson, T., & Asaro-Saddler, K. (2018). Using sentence combining instruction to enhance the writing skills of children with learning disabilities. *Learning Disabilities A Contemporary Journal, 16*(2), 191–201.
Scardamalia, M., & Bereiter, C. (1986). Research on written composition. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 778–803). Macmillan.
Shelley, M. (2003). *Frankenstein*. Barnes and Noble Classics, Inc.

Stotski, S. (1975). Sentence combining as a curricular activity: Its effect on written language development and reading comprehension. *Research in the Teaching of English, 9*, 30–71.

Strong, W. (1976). Close-up: Sentence combining. *English Journal, 24*, 56–65.

Strong. W. (1985). How sentence combining works. In D. A. Daiker, A. Kerek, & M. Morenberg (Eds.), *Sentence Combining: A Rhetorical Perspective* (pp. 334–352). Carbondale and Edwardsville, ILL: Southern Illinois University Press.

Strong, W. (1986). *Creative approaches to sentence combining*. ERIC Clearinghouse on Reading and Communication Skills and the National Council of Teachers of English.

Telesca, L., Ehren, B., Hahs-Vaughn, D., Zygouris-Coe, V., & Kong, A. (2020). The effect of metalinguistic sentence combining on eighth-grade students' understanding and written expression of comparison and contrast in science. *Journal of Speech, Language, and Hearing Research, 63*, 3068–3083.

Troia, G. (2014). *Evidence-based practices for writing instruction* (Document No. IC-5). Retrieved from University of Florida, Collaboration for Effective Educator, Development, Accountability, and Reform Center website: http://ceedar.education.ufl.edu/tools/innovation-configuration/

Walter, K., Dockrell, J., & Connelly, V. (2021). A sentence-combining intervention for struggling writers: Response to intervention. *Reading and Writing*. https://doi.org/10.1007/s11145-021-10135-8

# Part IV
# Writing Instructions for Students with Learning Disabilities

# Building Writing Skills for Summaries and Quick Writes


**Linda H. Mason and Jenna Basile**

**Abstract** Students with disabilities often lack the skills needed to complete short writing tasks such as summaries and quick writes. This is especially problematic given the number of time-limited short writing tasks required during classroom instruction and used for assessing learning. Much of the research on writing within time constraints has focused on production skills such as handwriting with young writers, writing sentences, or writing to curriculum-based measure (CBM) 3-min prompts. A smaller group of researchers have also addressed the impact of providing explicit instruction for supporting the writing process in short extended writing tasks such as summaries and quick writes, mirroring what is asked for in classroom instruction and assessment. Findings of 14 empirical studies indicate that when teachers model and provide structured strategy instruction in combination with procedures for self-regulation, specifically self-regulated strategy development (SRSD), students' writing improves across quantity and quality measures. Researchers have also established that including procedures for fluency building skills in writing instruction contributes to writing quality in timed writing tasks. Implications for research and for including SRSD instruction for short writing tasks such as summaries and quick writes to support students' with disabilities writing are discussed.

**Keywords** Summaries · Quick writing · Learning disabilities

With the focus on writing across the curriculum (e.g., Common Core State Standards Initiatives: NGA, 2010), students' with disabilities proficiency in expressing ideas in a variety of writing formats is critical. Many of these students, however, struggle to demonstrate proficiency when writing a restricted/brief written response such as a summary or quick write or when writing an extended format such as an essay or report. For many students with disabilities, writing within simple and complex formats is challenging due to a lack of the self-regulation and cognitive skills needed

L. H. Mason (✉) · J. Basile
College of Education and Human Development, George Mason University, Fairfax, VA, USA
e-mail: lmason20@gmu.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_10

for producing and completing a final written product (Harris et al., 2011; Graham et al., 2017). Regardless of the format, the lack of writing skills for expressing ideas and demonstrating knowledge negatively impacts students' ability to maximize learning. In content classes, for example, writing-to-learn formats such as summaries and quick writes are often used to allow students the opportunity to recall, clarify, and question information and to demonstrate their thinking about a topic by writing an opinion; and, to allow teachers the opportunity to assess student understanding of material read or taught (Frey & Fisher, 2012; Mason et al., 2012).

In the early 1980s, Brown and Day (1983) noted the difficulty in writing for tasks with restricted formats, tasks difficult for many across grade levels and age. These researchers established the effectiveness of explicit strategy instruction plus self-management training for improving the metacognitive skills needed to write an effective summary (e.g., Day, 1986). Given the findings of Brown and Day's research, it follows that this research-based approach for writing instruction, with individualized support, would be expected for students who have not yet developed the skills to complete short-extended writing tasks independently. In fact, programs of intervention research for elementary and adolescent writers have provided frameworks for effective instruction (e.g., Graham, 2019; Graham et al., 2012; Graham & Perin, 2007). Strategy instruction for written expression, for example, teaches students how to break writing tasks into manageable subtasks. Instruction that includes teaching and developing skills in self-regulation improves students' self-awareness and control (Harris et al., 2011). Best practice for writing includes scaffolding instruction with planned guided and independent practice to support students' independence over time and to foster generalization. In this chapter, we focus on and review intervention literature focused on teaching students with disabilities how to condense content for writing in structured short formats such as summaries and quick writes.

## 1 Summarization

Written summaries for content learning have four critical characteristics; summaries are short (i.e., less than one and one-half page), are written in the author's own words, tell important information, and state information to be learned (Friend, 2000). Summaries, a type of informative writing, are most often written to support learning after reading text (Graham & Hebert, 2011). In some cases, five rules have been taught to expedite summary writing from text read: delete trivial information, delete redundant information, substitute superordinate terms for a list of terms or actions, select a topic sentence, and invent a topic sentence (Brown & Day, 1983). Summaries can be used to explain one aspect of the text, or more broadly summarize text information. For example, after reading a chapter about the solar system, students may be asked to summarize what they read about a specific planet, or they may be asked to summarize what they read about all eight planets.

In Graham & Perin (2007) meta-analysis of group experimental studies, a mean weighted effect size of 0.82 was found for summarization instruction in grades 5

through 12. Instruction in the four review studies included (a) explicit strategy instruction in three of the studies and (b) the use of summarization models in one study. Graham and Perin noted, "teaching adolescents to summarize text had a strong impact on their ability to write more concise text" (p. 463). In 2008, Rogers and Graham reviewed single-subject experimental studies. Only one summary writing study was noted; 100% PND (percent of non-overlapping data) was noted for text production (Nelson et al., 1992).

## 2 Quick Writes

Quick writes are time limited short constructed responses to a question prompt related to a specific topic. Student learning is supported by presenting a non-threatening, informal, and brief writing activity for students (Frey & Fisher, 2012). Quick writes require students to think about and explain what they know before content instruction to activate prior knowledge, during content instruction to monitor learning, or after content instruction to evaluate what has been learned (Wood & Harmon, 2001). Quick writes can be implemented for a variety of purposes; for example, a student may write (a) an informative response describing vitamins in healthy snacks, (b) a narrative about a time in a health food store, or (c) an opinion or argumentative response to convince their classmates to avoid junk food. Quick writes benefit students' comprehension of content by encouraging students to make connections through the writing process (Mason et al., 2009).

The first author's interest in learning about and evaluating quick writes came from a need expressed by middle school science and social studies teachers, teachers working with students with disabilities in their inclusive content classrooms. Teachers noted what has often been observed in these classrooms (e.g., Mason et al., 2009); when asking students with disabilities to write a short, constructed response such as a quick write, these students would write little or nothing, and the writing often demonstrated little knowledge of the content. A review of quick write instruction indicated that although quick writes are included as a recommended activity in science and social studies teacher manuals, procedures for teaching quick writes had not been published (Mason et al., 2009). At the time of our chapter review, quick writes had not been evaluated through systematic literature review or meta-analytic procedures (Graham & Perin, 2007; Rogers & Graham, 2008).

## 3 Method

This systematic review of literature examines the effectiveness and treatment acceptability of empirical intervention research in summary and quick writing instruction, specifically for students with high-incidence disabilities (e.g., attention deficit hyperactivity disorder: ADHD, emotional behavioral disability: EBD, learning disability: LD, other health impaired: OHI, speech language impairment: SLI)

across instructional settings. Our purpose was to update the literature on this type of writing and to situate summary and quick write interventions for students with high-incidence disabilities within the 2012 standards for supporting writing across the curriculum (e.g., Common Core State Standards Initiatives). Therefore, studies were excluded if they had been included in Graham & Perin (2007) true and quasi-experimental meta-analysis or in Rogers and Graham's (2008) single-subject meta-analysis.

Summarizing and quick writing intervention studies for students with disabilities were located through a systematic search of prominent databases such as ApaPsychinfo, Academic Search Complete, Eric, Teacher Reference Center, Psychology and Behavioral Sciences Collection, and Education Research Complete. The following keywords were used throughout the search: *summary writing, quick writing, writing interventions,* and *special education*. The After eliminating duplicate studies and screening the literature for applicability to the inclusion criteria for the current review, a total of 56 studies were considered for further review. The database search and the descendent searches of existing literature and ancestorial searches (e.g., Benedek-Wood et al., 2014; Garwood et al., 2019; Hoover et al., 2012; Mason et al., 2009) found self-regulated strategy development (SRSD) instruction for summarizing and quick writing to be the only intervention implemented across all studies.

## *3.1 Inclusionary and Exclusionary Criteria*

Studies were included in this review if they utilized an experimental single-subject or group research design and were peer-reviewed and written in English. In addition, reviewed studies included (a) interventions with a focus on instruction for summary and quick writing, (b) students in grades 1–12, and (c) students previously diagnosed as having a high-incidence disability (e.g., ADHD, EBD, LD, OHI, SLI). Studies that included students without disabilities were reviewed only if the data was disaggregated for those with disability diagnoses. Studies focused on alternate forms of writing such as journal writing and extended writing tasks such as essays and reports and studies that included a technology-based intervention were eliminated. Dissertation and theses were eliminated from consideration for this chapter and review.

## *3.2 Coding and Data*

All studies were coded by the first and second author. Studies were coded for the study's experimental design, school setting, student demographics (e.g., gender, grade level, age, disability status), procedures (e.g., instructor, instructional time, number of phase probes in single-subject studies), measures such as quality (e.g.,

total number of points earned on a rubric) and number of strategy elements or information units and number of words written (number count, no ceiling), and treatment acceptability. Effect size information (ES) was used to report study measures results of each study's measures. Treatment acceptability was coded and reported descriptively.

Group experimental study ES results were coded for Cohen's *d*. In studies that did not have this statistic, the chapter authors calculated mean differences with Cohen's $d = M^1 - M^2/SD_{pooled}$. An ES was considered to be small (.20), medium (.50), or large (.80) as suggested Cohen (1988). Single-subject studies were coded for the percentage of non-overlapping data (PND) found in visual graphs of data. In studies that did not report PND, the chapter authors calculated the PND percentage (PND % = nonoverlapping data points/all data points) if the measure was reported in graph format. PND was calculated using the proportion of post-intervention data points less than the highest data probe in baseline. For single-subject design studies, 90% were considered a large effect; 70–90% considered a medium effect; and 50–70% considered a small effect (Scruggs et al., 1987).

## 4 Results

Fourteen studies met the inclusion and exclusion criteria—four evaluated summarization instruction (Asaro-Saddler et al., 2018; Ennis, 2016; Rogevich & Perin, 2008; Saddler et al., 2017) and 10 evaluated quick writing instruction (Benedek-Wood et al., 2014; Ciullo et al., 2019; Mong Cramer & Mason, 2014; Garwood et al., 2019; Hoover et al., 2012; Mason et al., 2011, 2013a, b—two studies; Mason et al., 2010). Eleven studies used single-subject design methodology; three studies used a quasi-experimental design (Asaro-Saddler et al., 2018; Mason et al., 2013b; Rogevich & Perin, 2008). Students' grade levels ranged from grade 4 to grade 12 (no studies were found involving students early than grade 4). Studies were implemented in inclusive classrooms, in learning support and alternative classrooms, in alternative day schools, and in residential treatment facilities.

Components and results across studies are illustrated by research design in two tables. Table 1 describes the single-subject studies, including probes across phases, and provides results as PND. Table 2 describes quasi-experimental studies and provides effect size results for Cohen's *d.* Setting (e.g., type of school), student demographics (e.g., gender, grade level, age, disability), instruction (time for and number of sessions, instructor), and results (e.g., quality, elements or knowledge/idea units, number of words written) are captured in both tables.

All reviewed studies included explicit strategy instruction with self-regulation, specifically self-regulated strategy development (SRSD) instruction. All authors reported procedures for evaluating treatment fidelity across lessons; this was noted to be strong across the studies. Given that all studies implemented SRSD instruction, this instructional approach will be described next, followed by an overview of each study.

**Table 1** Single-Subject Design Studies

| Study | Student description | Setting | Instruction | Percentage of non-overlapping date (PND) |
|---|---|---|---|---|
| *Persuasive quick writing* | | | | |
| Garwood et al. (2019) | *n* = 11 female<br>13–17 yrs old<br>*n* = 8 EBD<br>*n* = 2 ADHD<br>*n* = 1 no disability | Residential treatment facility<br>Mixed grade science class | Six to seven 30-min sessions<br>Whole class instruction provided by the school's science teacher | Post-instruction[a]<br>96% (0–8 pt quality)<br>98% (# elements)<br>87% (#words)<br>Maintenance[a]<br>100% (0–8 pt quality)<br>100% (# elements)<br>73% (# words) |
| Hoover et al. (2012) | *n* = 4 female<br>16–19 yrs old<br>*n* = 4 LD | Inclusive high school<br>*n* = 2 11th grade *n* = 2 12th grade | Five 30-min sessions<br>Five 10-min sessions<br>One-to-one instruction provided by first author (students' special education teacher) | Post-instruction[a]<br>60% (# elements)<br>25% (# words)<br>Maintenance[a]<br>71% (# elements)<br>14% (# words) |
| Mason et al. (2013a, b) | *n* = 3 male<br>15–17 yrs old<br>*n* = 3 EBD | Inclusive high school<br>*n* = 1 9th grade<br>*n* = 2 11th grade | Five 30-min sessions<br>One to three 10-min sessions<br>One-to-one instruction provided by the third author (students' special education teacher) | Post-instruction<br>79% (0–7 pt quality)<br>68% (#elements)<br>68% (#words)<br>Maintenance<br>83% (0–7 pt quality)<br>50% (# elements)<br>66% (# words) |

| | | | | |
|---|---|---|---|---|
| Mason et al. (2011) | Study one | Study one | Study one | Study one |
| | *n* = 2 female<br>*n* = 4 male<br>12–13 yrs old<br>*n* = 4 LD<br>*n* = 1 LD/ADHD<br>*n* = 1 ADHD | Inclusive middle school<br>*n* = 5 6th grade<br>*n* = 1 7th grade | Five to six 45-min sessions<br>Student paired instruction provided by the third GRA author | Post-instruction<br>56% (0–10 pt quality)<br>94% (# of elements)<br>Maintenance<br>75% (0–10 pt quality)<br>100% (# elements) |
| | Study two | Study two | Study two | Study two |
| | *n* = 6 female<br>*n* = 4 male<br>12–13 yrs old<br>*n* = 9 LD<br>*n* = 1 ADHD | Inclusive middle school<br>*n* = 10 sixth grade | Five to six 45-min sessions<br>Small group instruction provided by the students' special education teacher | Post-instruction<br>62% (0–10 pt quality)<br>77% (# elements)<br>Maintenance<br>50% (0–10 pt quality)<br>67% (# elements) |
| Mason et al. (2010) | *n* = 1 female<br>*n* = 5 male<br>12–14 yrs old<br>*n* = 3 EBD<br>*n* = 1 EBD/autism<br>*n* = 1EBD/ADHD/LD | Alternative middle school<br>6th grade | Five 30-min sessions<br>Three 10-min sessions<br>One-to-one instruction provided by the third GRA author | Post-instruction<br>84% (0–7 pt quality)<br>20% (# elements)<br>0% (# words)<br>Maintenance<br>60% (0–7 pt. quality)<br>20% (# elements)<br>0% (# words) |
| *Persuasive quick writing extended* | | | | |
| Mong Cramer and Mason (2014) | *n* = 1 female<br>*n* = 7 male<br>Age n/a<br>*n* = 8 EBD | Alternative middle school<br>6th grade | Five 45-min planning and composing sessions<br>Five 45-min revision sessions<br>Student paired instruction provided to by the first GRA author | Post-instruction[a]<br>96% (0–7 pt quality)<br>71% (# elements)<br>50% (# words)<br>Maintenance<br>n/a |

(continued)

**Table 1** (continued)

| Study | Student description | Setting | Instruction | Percentage of non-overlapping date (PND) |
|---|---|---|---|---|
| Ciullo et al. (2019) | *n* = 4 female<br>*n* = 4 male<br>Age n/a<br>*n* = 5 LD<br>*n* = 2 ADHD<br>*n* = 1 no disability | Inclusive elementary school<br>*n* = *6* 5th<br>*n* = *2* 4th | 50-min sessions—3 for reading, 5 for writing, 1 for timed-practice<br>Student paired instruction provided by an interventionist, a retired special education teacher | Post-instruction<br>94% (# elements)<br>Maintenance<br>n/a |
| *Informative quick writing* | | | | |
| Benedek-Wood et al. (2014) | *n* = 41 female,<br>*n* = 37 male<br>10–12 yrs old<br>*n* = 7 LD<br>*n* = 2 ADHD<br>*n* = 1 EBD | Inclusive middle school<br>Four 6th grade science classrooms | Four 30-min sessions<br>Four 10-min session<br>Whole class instruction provided by the science teacher (third author) | Post-test (class average)<br>100% (0–5 pt quality)<br>71% (# knowledge units)<br>100% (# words)<br>Maintenance<br>n/a |
| *Summary writing from informational social studies text* | | | | |
| Ennis (2016) | *n* = 2 female<br>*n* = 1 male<br>Age n/a<br>*n* = 2 EBD | Residential treatment facility<br>9th, 10th, and 11th grade | Six 40–50 min sessions<br>One-to-one instruction provided by the students' social study teacher | Post-instruction[a]<br>100% (0–14 pt quality)<br>100% (# elements)<br>100% (# words)<br>Maintenance<br>n/a |
| Saddler et al. (2017) | *n* = 1 female<br>*n* = 5 male<br>10–11 yrs old<br>*n* = 6 EBD | Elementary self-contained classroom<br>5th to 6th grade mixed classroom | Six 30–40 min sessions<br>Paired instruction provided by the fourth GRA author | Post-instruction[a]<br>94% (0–17 pt. quality)<br>Maintenance<br>n/a |

*Note. yrs* years, *#* number written, *EBD* emotional behavioral disability, *ADHD* attention hyperactivity disorder, *LD* learning disability, *pt* point, *n/a* data not available in published study, *GRA* graduate research assistant

[a]Group PND calculated by chapter authors

**Table 2** Group-experimental studies

| Study | Student description | Treatment groups | Setting | Instructional delivery | Posttest effect size results |
|---|---|---|---|---|---|
| *Persuasive quick writing* | | | | | |
| Mason et al. (2013b) | *n* = 134 female<br>*n* = 145 male<br>Mean age = 13.28<br>*n* = 55 IEP | *n* = 33 treatment<br>*n* = 51 eligible comparison<br>*n* = 195 whole class comparison | Inclusive middle school<br>8th grade | 5 30-min sessions + <5 10-min sessions<br>One-to-one instruction provided by 7 GRAs | Treatment to eligible comparison<br>*d* = 1.11 (0–7 pt quality)<br>*d* = 0.81 (# elements)<br>*d* = 0.58 (# words)<br>Treatment to whole class comparison<br>*d* = 0.88 (0–7 pt quality)<br>*d* = 0.39 (# elements)<br>*d* = −0.33 (# words) |
| *Summary writing from informational science text* | | | | | |
| Rogevich and Perin (2008) | *n* = 63 male<br>13–16 yrs old<br>*n* = 32 BD (32)<br>*n* = 31 BD/ADHD | *n* = 15 BD treatment<br>*n* = 17 BD comparison<br>*n* = 15 BD/ADHD treatment<br>*n* = 16 BD/ADHD comparison | Residential treatment facility<br>Mixed grade level classes | Five 45-min sessions<br>Small group (3–4) instruction provided by the first author, a clinician at the school | 0–20 pt quality scale for the # information units written<br>BD treatment to BD comparison[a]<br>*d* = 3.91 (science text)<br>*d* = 2.98 (social studies text)<br>*d* = 5.08 (multi-source text)<br>*d* = 2.82 (maintenance)<br>BD/ADHD treatment to BD/ADHD comparison[a]<br>*d* = 2.61 (science text)<br>*d* = 1.48 (social studies text)<br>*d* = 1.65 (multi-source text)<br>*d* = 0.89 (maintenance) |
| *Summary writing from informational social studies text* | | | | | |
| Asaro-Saddler et al. (2018) | *n* = 13 female,<br>*n* = 17 male<br>16–19 yrs old<br>*n* = 25 LD<br>*n* = 3 OHI<br>*n* = 1 SLI<br>*n* = 1ASD | *n* = 16 treatment<br>*n* = 14 control | Resource classroom<br>Grade n/a | 22 lessons (time n/a)<br>Whole resource class instruction provided by the school's special education teacher | Treatment to control[a]<br>*d* = 0.97 (0–14 pt quality scale)<br>*d* = 0.80 (reading comprehension)<br>*d* = 2.06 (# words) |

*Note. yrs* years, *#* number written, *BD* behavioral disability, *ADHD* attention hyperactivity disorder, *LD* learning disability, *OHI* other health impaired, *SLI* speech language impaired, *ASD* autism spectrum disorder, *pt* point, *d* = Cohen's *d*, *n/a* data available in published study, *GRA* graduate research assistant

[a]*d* calculated by chapter authors

### *4.1 SRSD Instruction*

To facilitate students' strategy acquisition, six instructional stages are implemented throughout SRSD instruction: develop background knowledge, discuss it, model it, memorize it, guided practice, and independent practice (Graham & Harris, 1996, 2003; Mason et al., 2012). In SRSD, strategy use and self-regulation of the writing process is gradually shifted from teacher-directed instruction to the student through scaffolding. Instruction is criterion-based rather than time-based; students demonstrate mastery of a particular stage or procedure before moving on to the next phase of instruction. Students' independent strategy use is supported over time and context with booster sessions as needed.

In SRSD instruction, goal setting, self-monitoring, self-instruction, and self-reinforcement are explicitly taught and supported to meet students' individual needs. These self-regulation procedures are targeted to address the students' environment, writing task, and skill level. Specific, proximal, and appropriately challenging goals, to meet the genre-specific task, are established. Students self-monitor their goals by assessing whether or not the goal has been achieved and then self-recording the result. Students are taught to use six basic personalized self-instructions to support their writing: (1) problem definition, (2) focus of attention and planning, (3) strategy use, (4) self-evaluation and error correcting, (5) coping and self-control, and (6) self-reinforcement. Students self-reinforce by selecting a reinforcer or by self-rewarding.

The review of this collection of SRSD instructional studies begins with summary writing, followed by a review of the quick write studies. In order to gain a sense of the development of procedures in methodology and for the development of instruction in lines of research, studies are reviewed in literature timeline (i.e., date the studies occurred).

### *4.2 SRSD for Summary Writing*

SRSD for writing summaries from text was explored in (a) one quasi-experimental study (Rogevich & Perin, 2008) and one single-subject study (Ennis, 2016) for TWA—Think before reading, think After reading, think While reading, (b) one single-subject study (Saddler et al., 2017) for WIN—Write a topic sentence, Identify the important information, and Number, and (c) one-quasi-experimental (Asaro-Saddler et al., 2018) for WINDOW—Write the main idea, Identify important information, Number important information, Develop sentences, Organized sentences using transition words, Write an ending sentence.

In the four studies reviewed, the reading comprehension strategy (i.e., TWA, WIN, or WINDOW) was taught prior to students applying the strategy to writing a summary. A key element in each of these studies is the focus on improving students

with disabilities skills in identifying important information to learn and to write about. Applying Brown and Day's (1983) rules, students are taught to delete trivial and redundant information and use superordinate terms where possible (e.g., "predator" for alligator, hawk, snake, and wolf). Each of the studies included procedures for self-regulation, such as goal sheets and self-monitoring checklists, to support both reading and summary strategy application.

### TWA and Summary Writing Studies

Rogevich and Perin (2008) determined the effects of TWA (Mason, 2004; Mason et al., 2006) plus Written *S*ummarization (TWA-WS) for ideas units (i.e., most important ideas) written after reading a science text. The quasi-experimental design included 63 males in a residential treatment facility. Students were equally assigned to intervention (behavior disorder (BD)/Intervention and BD + ADHD/Intervention) and comparison (BD/Practice with no treatment and BD + ADHD/Practice with no treatment). Five 45-min instructional sessions were provided by a school's clinician to groups of three to four students. During pre- and post-instruction assessment, students were given a science passage to read, followed by 15 min to write a summary. Each summary could earn up to 20 quality points based on the number of idea units written. Post hoc comparisons showed a significantly higher performance, with *large* effects, for the intervention groups when compared comparison groups (*d* ranged from 2.61 to 3.91). Results indicated that skills generalized to social studies text with *medium* to *large* effects (*d* ranged from 1.48 to 2.98), generalized to multi-source text with *large* effects (*d* ranged from 1.65 to 5.08); and, maintained 3 weeks after the intervention (*d* ranged from 0.89 to 2.82). Rogevich and Perin noted the limitations for instructional time, only five sessions were implemented. All students were positive about the intervention, one stating, "This [intervention] has really helped me figure out what's important in a passage. They should teach us how to do this in school, but they don't, they just expect us to already know how to do this" (p. 147).

In a multiple-probe, multiple-baseline design across students study, Ennis (2016) explored the effects of TWA + PLANS (Pick goals, List goals, And, make Notes, Sequence notes) on social studies summary writing with three high school students with EBD (two female, one male) in a residential treatment facility. Instruction was implemented by the students' social studies teacher, 2–3 days per week in 40–50 min sessions. Results, calculated by the authors of this chapter, indicated 100% PND *large* effects from baseline (3–4 probes) to post-intervention (3–5 probes) for the 0–14 point quality scale, for summary elements, and for the total written words. Maintenance data and treatment acceptability data was not reported. Ennis noted the positive effects were supported by the use of individualized instruction and in a setting where positive behavioral interventions and support (PBIS) was implemented with fidelity.

**WIN and WINDOW Studies**

In a 2017 a multiple-probe, multiple-baseline design across students study, Saddler et al. explored the effects of WIN with FRI (Facts, Readings, and Ideas) with six 5th and 6th grade students with EBD (one female, five male) in a self-contained classroom. Summary writing was taught in six lessons, 30–40 min per day by a graduate research assistant (GRA). The quality of students' 20-min writing probes were assessed using a 0–17 point rubric. All students were noted to improve quality from baseline (three to five probes) to post-instruction (three probes). Chapter authors calculated *large* effects with 95% PND for writing quality. Maintenance data and treatment acceptability data was not reported. The study's positive findings suggest SRSD for the WIN strategy may increase the summary writing quality and skills of students with EBD.

In 2018, Asaro-Saddler et al. implemented a pre-/post-test quasi-experimental design to explore the effects of WINDOW for reading comprehension and summary writing skills of students in a resource classroom. WINDOW reinforced self-regulation by asking students to complete a WINDOW Self-Check Guide to evaluate their written summary. Students included 30 students (13 female, 17 male) with IEPs (25 with LD, 3 with OHI, 1 with SLI, 1 with autism spectrum disorder) assigned to the control ($n = 14$) or treatment groups ($n = 16$). Instruction was implemented in 22 lessons, taught twice per week by two of the school's trained special education teachers. Over the four-month duration of the intervention, the control group received their usual language arts instruction. The study effect size results, calculated by chapter authors, showed that all students in the treatment group, when compared to the control group, showed higher scores with *large* effects for quality on a 0–14 point scale ($d = 0.97$) and for the number of words written ($d = 2.06$). Reading comprehension, measured by students' responses to six questions after reading, indicated a *moderate* effect ($d = 0.80$). Maintenance and treatment acceptability data was not reported. Results indicate that the WINDOW strategy may increase the summary writing and reading comprehension skills of high school students with disabilities.

## *4.3 SRSD for Quick Writes*

SRSD for quick writing has been evaluated in nine multiple-baseline single-subject design studies with middle and high school students with disabilities (Benedek-Wood et al., 2014; Ciullo et al., 2019; Mong Cramer & Mason, 2014; Garwood et al., 2019; Hoover et al., 2012; Mason et al., 2011, 2013a—two studies; Mason et al., 2010) and one quasi-experimental study (Mason et al., 2013b). SRSD for quick writing was provided to students in one-to-one sessions, in small-group sessions, and to a whole class in elementary, middle, and high school settings.

Ten minutes is generally given for quick writing in content classrooms (Frey & Fisher, 2012). Therefore, in each of the following reviewed studies, with the

exception of Ciullo et al. (2019) where students were given 13-min to write, SRSD instruction was implemented to teach students how to write a quick write in a 10-min time frame. In each study, writing instruction ranged from five to seven 30-min lessons for strategy acquisition, plus one to five lessons for independent practice in writing within the 10-min (or 13-min) time frame.

Quick write instruction follows procedures outlined in SRSD instruction with additional lessons for writing within a specified time. Prior to instruction, the teacher asked the students to complete a quick write to establish performance prior to instruction.

- Lesson One. In the first SRSD lesson, the necessary prerequisite skills for strategy mastery are developed, the selected genre-specific strategy is discussed and explained, and students begin to memorize each strategy step. The teacher emphasizes how the strategy steps and other procedures to be learned will support writing improvement.
- Lesson Two: In the second lesson, the teacher cognitively models quick writing using the selected strategy and all instructional materials (e.g., mnemonic sheets, graphic organizers, graphing charts). The 10-min time limit is extended for modeling, so the teacher can explicitly model each step and each self-regulation procedure. Following modeling, students establish personal goals and self-instructions, and then evaluate their prior performance from the previously collected writing sample.
- Lesson Three: The teacher provides collaborative group and/or peer practice in the third lesson. This guided practice lesson is repeated as needed until students demonstrate independence in quick writing without time constraints. In each lesson, students self-monitor their writing performance and record the number of strategy steps completed on a graphing chart.
- Lesson Four: Once students have mastered an untimed quick write, the teacher cognitively models a second time to show students how to use the quick write strategy in the time limit. Self-instructions to regulate writing in a specified time frame (e.g., 10 min) are critical in this modeling lesson.
- Lesson Five: Guided practice is repeated in multiple lessons until students demonstrate independence in writing within the specified time frame.

### Persuasive Quick Writing

In each of the persuasive quick writing studies, the POW (Pick my idea, Organize my notes, Write and say more) + TREE (Topic sentence, Reasons: three or more, Explain each reason, Ending sentence) strategies were used (a) to support three writing processes for brainstorming ideas, planning, and composing and (b) to support the elements of good persuasion by using TREE elements for planning and writing. Students self-regulated by setting goals for quick writing, by writing a good topic and ending sentence, and by counting the number of reasons and explanations written. Students were encouraged to include at least one rebuttal and refutation for

the position they supported. Rebuttal and refutations did not influence scores for the number of elements and number of words written, both on a continuous scale. However, this element did influence the quality of writing which was evaluated on a scale of 0 to an upper limit of 7 or 10. All single-subject studies for persuasive quick writing implemented a multiple-probe, multiple-baseline design across students or across classrooms.

**Single-Subject Studies** In the first of two quick writing studies, Mason et al. (2011) evaluated the persuasive/opinion quick writing performance of six students (2 female, 4 male) with disabilities (4 with LD, 1 with LD/ADHD, 1 with ADHD), assigned to pairs and instructed by a GRA in an inclusive middle school. The second study included 10 students (6 female, 4 male) with disabilities (9 with LD, 1 with ADHD), placed in three small groups and instructed by a trained special education teacher. In both studies, the POW + TREE strategy was taught in six to seven 45-min lessons. During baseline (3–7 probes), post-instruction (3–5 probes), and 2-week maintenance (1–3 probes), students' responses were evaluated for quality on a 1–10 point scale, the number of TREE elements, and the number of words written. Quality from baseline to post-instruction increased with *small* effects in both studies (56% PND; 62% PND respectively). Quality at maintenance indicated a *medium* effect in study one (75% PND) and a *small* effect in study two (50% PND). An overall improvement from baseline to post-instruction for the number of written elements indicated a *large* effect (94% PND) in study one and a *medium* effect (77% PND) in study two; at maintenance a *large* effect was noted in study one (100% PND) and a *small* effect in study two (67% PND). Number of words written was not graphed. All students in study one reported that POW + TREE instruction helped them improve their writing. In study two, students' responses were positive overall; one student, however, noted preferring expressive writing, and another noted that they preferred the four-square strategy because "you do not have to think" (p. 216).

In a third study, Mason et al. (2010) examined the effects of POW + TREE for five middle school students with EBD (1 female, 5 male) in an alternative school. Following baseline (5–9 probes), five 30-min and three 10-min sessions were provided in a one-to-one setting outside of the classroom by a GRA. Students writing quality, scored on a 0–7 point scale, demonstrated a *medium* effect of 84% PND at post-instruction (5 probes) and *small* effect with 60% PND at maintenance (1 probe). No PND effect was found for the number of elements or words written. Treatment acceptability was assessed by asking the students to write a persuasive quick write to the following prompt: "Should students your age be taught how to write using POW+TREE?" (p. 146). All students wrote at least three reasons to support POW + TREE instruction, noting that the strategy helped with organization, thinking things through, writing more, and generating ideas.

Mason et al. (2013a) examined the effects of POW + TREE with three male students with EBD in an inclusive high school. Assessment and instruction were provided the students' special education teacher, a doctoral student at the time of the study, prior to the start of the school day. Five to seven quick write probes were

collected a baseline followed by five one-to-one instructional sessions. Following instruction, six to seven post-instruction and two maintenance probes were collected. Quality, on a 0–7 point scale, indicated a *medium* effect after the intervention (79% PND) and at maintenance (83% PND). The number of response elements written indicated *small* effects of 68% PND after instruction and 50% PND at maintenance. The number of words written indicated a *small* effect for all three students, 68% PND immediately after instruction, and 66% PND at maintenance. Students reported the intervention had positive effects: "It has helped me become a better writer. It is easier for me to write. Before I would just think of something and write about it, but now I have a strategy" (p.172).

Hoover et al. (2012) examined the effectiveness of POW + TREE with four female students with LD in an inclusive high school. Instruction was provided before school by the first author, a doctoral student and the students' special education teacher at the time of the study. Prior to instruction, five to eight baseline probes were collected. Each student then completed five 30-min instructional sessions and five 10-min sessions. Results of post-instruction (5–6 probes) and maintenance (1–2 probes), calculated by chapter authors, indicated *small* and *medium* effects on number of elements written after instruction (60% PND) and at maintenance (71% PND) and no effect for the number of words written after instruction (25% PND) or at maintenance (14% PND). Quality was not reported. Results of the social validity measure indicated that all students viewed the intervention positively. One student noted "The lessons were good and easy to learn because the mnemonics are easy to learn" (p. 32).

**Quasi-Experimental Study** Using a comparison group pre-/post-test quasi-experimental study, Mason et al. (2013b) examined the writing performance of 279 eighth-grade students enrolled in four urban charter schools. The WJ-Fluency III subtest was used to screen all students in eighth grade for eligibility. Groups included 33 students in treatment ($n$ = 12 with IEP; 11 female, 22 male); 51 in an eligible comparison—students who met criteria but did not consent for instruction ($n$ = 18 with IEP; 22 female, 29 male); and 195 in an non-eligible comparison group ($n$ = 25 with IEP; 101 female, 94 male). All students were assessed by their teacher in whole-class assessment prior to and after instruction. Seven GRAs ants delivered POW + TREE instruction to the treatment group in five 30-min and five 10-min one-to-one sessions. Results indicated the treatment group, compared to the eligible comparison group, had *large* effects at post-test for quality on a 0–7 point scale ($d$ = 1.10) and number of elements ($d$ = 0.81) and *medium* effects for number of words ($d$ = 0.58). Results comparing the treatment group to the non-eligible comparison group indicated *large* effects for quality ($d$ = 0.88). *Small* effects were found for the number of elements written ($d$ = 0.39). Results also indicated that the whole class comparison group wrote more words than treatment at post-test ($d$ = 0.33). All 33 treatment students were positive about instruction, noting "It improved how I write an essay by using reason and explanation and I don't mix up the words anymore when I write" and that it helped others "become a great writer" (p. 244).

**Persuasive Quick Writing Extended** SRSD for POW + TREE quick writing was extended in two studies by examining the adding revision instruction to the lessons (Mong Cramer & Mason, 2014) and by examining the effect of instruction when combined with a reading strategy (Ciullo et al., 2019). Mong Cramer and Mason implemented a multiple-probe, multiple-baseline design across students study with an alternating treatment design to evaluate the effects of POW+ TREE quick writing paired with a peer revision strategy, LEAF (Listen as the author reads, Explain what you like best, Ask evaluation questions, Finalize your comments). Eight students in an alternative program for adolescents with EBD (1 female, 7 male) were taught in pairs by the first author, a doctoral student at the time of the study. The experimental design included four phases: (a) five baseline prompts, (b) five 45-min POW + TREE sessions, (c) four 45-min LEAF sessions, and (d) alternating assessments with three peer and three independent revisions. During assessment sessions, students were given 10 min for planning, class time for working with their peer, and 10 min to write a final quick write. Results, calculated by chapter authors, following POW + TREE+LEAF instruction, indicated *large* effects for quality on a 0–7 point scale (96% PND), *medium* effects for number of elements (71% PND), and a *small* effect for number of words (50% PND). Maintenance data was not reported. Results showed that students with EBD can provide meaningful peer feedback and that they recognized instruction and feedback from a peer were helpful, for example, "Yeah, two heads are better than one" (Cramer, 2011, p. 180).

Ciullo et al. (2019) examined the effects of SRSD to teach four female and four male fourth-grade students (5 with LD, 2 with ADHD, 1 no disability) to paraphrase social studies text and then to compose a persuasive quick write about the text read. A retired special education teacher provided 50-min instructional sessions—three for reading the text, five for quick writing, and one for timed-practice. Students were first taught the reading comprehension paraphrasing strategy TRAP (Think about you know and what to know, Read a paragraph and Ask yourself—"What are the main ideas and details", Put the main ideas and details into your own words) in three consecutive sessions. Next, instruction for POW + TREE was provided. Given the age of the students in this study, quick writing time was extended from 10 to 13 min. All students' performance in writing persuasive parts/elements improved from baseline (3–8 probes) to post-instruction (5–7 probes) with *large* effects for the number of TREE elements written (94% PND). Quality, number of words written, and maintenance was not reported. All students felt instruction was beneficial. One student noted, "TRAP and TREE are easy. Other kids should use TRAP and TREE and then they can write more like I do" (p.19).

### Informative Quick Writing

In a staggered multiple-probe, multiple-baseline design across classrooms study in four middle school science classrooms, Benedek-Wood et al. (2014) examined the effects of SRSD for informative quick writing. One sixth-grade general education

science teacher delivered the intervention to 78 students across four classes (41 female, 37 male). Of the 78 students, 10 students had been diagnosed with a disability including seven with LD, two with ADHD, and one with EBD. The POW+ TIDE (Topic sentence, Important Details, Ending sentence) strategy was introduced to students after completing a minimum of five baseline probes. Through the staggered A-B design across classes, students were introduced to the strategy imbedded within their typical science instruction. After five POW+ TIDE sessions, students were provided with four fluency practice lessons including 10 min of segmented time for teacher feedback. One week after the final lesson, two to five post-instruction probes were collected from each class. Whole class results at post-instruction indicated 100% PND (*large* effect) for quality on a 0–5 point scale, 71% PND (*medium* effect) for knowledge units, and 100% PND (l*arge* effect) for number of words. Maintenance data was not reported. Students' responses ($n = 78$) on a five-point scale social validity questionnaire ranged from 3.2 to 4.2 (1 = Strongly disagree; 5 = Strongly agree). Scores reflected that students felt instruction was helpful but they did not think they would use the strategies in other classes.

Garwood et al. (2019) examined the effects of a multiple-probe, multiple-baseline design across classes for persuasive quick writing in science in a residential treatment facility. The 11 female secondary students (8 with EBD, 2 with ADHD, 1 no disability) had experienced sexual abuse and other forms of trauma. The students' science teacher provided instruction for POW + TREE in six 30- to 35-min lessons at the beginning of the science class. In addition to baseline (5–7 probes) and post-instruction probes (5 probes), a maintenance probe was given the week following instruction. PND, calculated by the chapter authors, indicated *large* effects for quality on a 0–8 point scale (96% PND post-instruction, 100% PND at maintenance) and number of strategy parts (98% PND post-instruction, 100% PND at maintenance). The number of words indicate *medium* effects (87% PND post-instruction, 73% PND at maintenance). All students commented positively about the intervention: (a) What have you learned? "How to communicate my opinion and present my arguments" and (b) How do you think this will help other students? "It will help them on tests and in jobs" and "It helps to make writing less stressful" (p. 12). The science teacher also provided feedback noting the intervention was "easy to learn," "a great benefit to the students," and "something these kids needed" (p. 9).

## 5 Discussion

Effects of SRSD instruction for improving students' with high-incidence disabilities written summarization and quick writing, reviewed in this chapter, extends research Brown and Day (1983) completed with older students without disability in the 1980s (e.g., Day, 1986) and the research reviewed by Graham & Perin (2007) and Rogers and Graham (2008). The studies we reviewed were situated across settings, including elementary, middle, and high school inclusive classrooms, special education classrooms, and class settings in residential treatment facilities.

Instruction in eight of the 14 studies was provided by the school personnel. The large PND effect sizes noted in quasi-experimental summary writing studies (e.g., Asaro-Saddler et al., 2018; Rogevich & Perin, 2008) and single-subject quick write studies (e.g., Benedek-Wood et al., 2014; Ennis, 2016; Garwood et al., 2019) provides evidence that SRSD instruction can be used effectively by classroom teachers. Instructional time appears to have an impact on results. For example, in studies where time was restricted to five to six sessions for strategy acquisition (e.g., Hoover et al., 2012; Mason et al., 2011, 2013a, b) effects were smaller than studies with more instructional sessions such as in Asaro-Saddler's et al. 22 session intervention.

Of the 14 studies, all except Ennis (2016) and Asaro-Saddler et al. (2018), which included complex reading plus summary writing, provide evidence that students with disabilities can be taught to write a brief summary or quick write within a 10- to 20-min time period. The Ennis and Asaro-Saddler studies highlight the need for instruction that explicitly defines writing task demands, addresses methods and tools to meet those demands, individualizes based on the needs of students with disabilities, and provides the time needed for writing when tasks are complex.

Out of the seven studies implemented by school personnel, only Garwood et al. (2019) measured teacher perceptions. In this study, the science teacher responded positively to all given survey questions except time for implementation. This response reflects what has been noted by teachers in survey research and has been called for in recommendations for administrators and policy-makers (Graham, 2019). As noted by Graham, advocacy to promote the importance of writing, and to include the time needed for writing instruction in the curriculum, is critical if we hope to improve students' outcomes in this critical academic skill.

Importantly, students participating in the one summary writing study (Rogevich & Perin, 2008) and all 10 quick writing studies, where treatment accessibility data was collected, reported that instruction had been beneficial and helped them become better writers. SRSD instruction, and the strategies taught, were viewed favorably. Some students, however, noted that often the logistical aspects of a study such as coming to school early (Hoover et al., 2012; Mason et al., 2013a, b) or missing homework time (Mason et al., 2013a, b) was problematic. Some students noted a less demanding writing task such as narrative writing was preferred (Mason et al., 2011). Students in the reviewed studies learned how to write, but some appeared to view this as not important outside the context of the study, for example, outside of the science classroom (Benedek-Wood et al., 2014). The few negative comments highlight the need for helping students understand the importance of learning how to write well.

In conclusion, the magnitude of the effects in the reviewed instruction, although varied from *small* to *large* effects across settings, instruction, tasks, and disabilities, indicates additional large-scale replication and efficacy study is warranted for SRSD instruction for summarization and quick writing, especially when combining this instruction with a reading comprehension strategy. In addition, researchers, administrators, and teachers need to do more to support a student's "journey as a writer" (Graham, 2019. p. 296).

## References[1]


*Asaro-Saddler, K., Muir-Knox, H., & Meredith, H. (2018). The effects of a summary writing strategy on the literacy skills of adolescents with disabilities. *Exceptionality, 26*(2), 106–118. https://doi.org/10.1080/09362835.2017.1283626

*Benedek-Wood, E., Mason, L. H., Wood, P. H., Hoffman, K. E., & McGuire, A. (2014). An experimental examination of quick writing in the middle school science classroom. *Learning Disabilities: A Contemporary Journal, 12*(1), 69–92. http://web.b.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=0&sid=e16389a7-bae2-4de5-9259-42d1058d42da%40pdc-v-sessmgr05

Brown, A. L., & Day, J. D. (1983). Macrorules for summarizing texts: The development of expertise. *Journal of Verbal Learning and Verbal Behavior, 22*(1), 1–14. https://doi.org/10.1016/S0022-5371(83)80002-4

*Ciullo, S., Mason, L. H., & Judd, L. (2019). Persuasive quick writing about text: Intervention for students with learning disabilities. *Behavior Modification, 45*(1),122–146. https://doi.org/10.1177/0145445519882894

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Inc.

Cramer, M. M. (2011). *The effects of strategy instruction for writing and revising persuasive quick writes on middle school students with emotional behavioral disorders* (Order No. 3483694). The Pennsylvania State University. https://search-proquest-com.mutex.gmu.edu/docview/902633738?pq-origsite=summon

Day, J. D. (1986). Teaching summarization skills: Influences of student ability level and strategy difficulty. *Cognition and Instruction, 3*(3), 193–210. https://doi.org/10.1207/s1532690xci0303_3

*Ennis, R. P. (2016). Using self-regulated strategy development to help high school students with EBD summarize informational text in social studies. *Education and Treatment of Children, 39*(4), 545–568. https://doi.org/10.1353/etc.2016.0024

Frey, N., & Fisher, D. (2012). *Improving adolescent literacy: Content area strategies at work*. Pearson Prentice Hall. https://www.pearson.com/us/higher-education/program/Fisher-Improving-Adolescent-Literacy-Content-Area-Strategies-at-Work-4th-Edition/PGM240947.html

Friend, R. (2000). Teaching summarization as a content area reading strategy. *Journal of Adolescent & Adult Literacy, 44*(4), 320–329. http://www.jstor.org/stable/40015345

*Garwood, J. D., Werts, M. G., Mason, L. H., Harris, B., Austin, M. B., Ciullo, S., Magner, K. M., Koppenhaver, D. A., & Shin, M. (2019). Improving persuasive science writing for secondary students with emotional and behavioral disorders educated in residential treatment facilities. *Behavioral Disorders, 44*(4), 227–240. https://doi.org/10.1177/0198742918809341

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education, 43*(1), 277–303. https://doi.org/10.3102/0091732X18821125

Graham, S., & Harris, K. R. (1996). Self-regulation and strategy instruction for students who find writing and learning challenging. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 347–360). Lawrence Erlbaum Associates, Inc. https://www.taylorfrancis.com/books/9780203811122

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323–344). The Guilford Press. https://www.guilford.com/books/Handbook-of-Learning-Disabilities/Swanson-Harris-Graham/9781462518685

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566


[1] *Reviewed Articles

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Graham, S., Bollinger, A., Booth Olson, C., D'Aoust, C., MacArthur, C., McCutchen, D., & Olinghouse, N. (2012). *Teaching elementary school students to be effective writers: A practice guide* (NCEE 2012-4058). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/WWC_Elem_Writing_PG_Dec182018.pdf

Graham, S., Collins, A. A., & Rigby-Wills, H. (2017). Writing characteristics of students with learning disabilities and typically achieving peers: A meta-analysis. *Exceptional Children, 83*(2), 199–218. https://doi.org/10.1177/0014402916664070

Harris, K. R., Graham, S., MacArthur, C., Reid, R., & Mason, L. H. (2011). Self-regulated learning processes and children's writing. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 187–201). Routledge. https://psycnet.apa.org/record/2011-12365-001

*Hoover, T., Kubina, R., & Mason, L. H. (2012). Effects of quick writing instruction for high school students with learning disabilities. *Exceptionality, 20*(1), 20–38. https://doi.org/10.1080/09362835.2012.640903

Mason, L. H. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. *Journal of Educational Psychology, 96*(2), 283–296. https://doi.org/10.1037/0022-0663.96.2.283

Mason, L. H., Hickey Snyder, K., Sukhram, D. P., & Kedem, Y. (2006). Self-regulated strategy development for expository reading comprehension and informative writing: Effects for nine 4th-grade students who struggle with learning. *Exceptional Children, 73*(1), 69–89. https://doi.org/10.1177/001440290607300104

Mason, L. H., Benedek-Wood, E., & Valasa, L. (2009). Quick writing for students who struggle with writing. *Journal of Adolescent and Adult Literacy, 53*(4), 313–322. https://doi.org/10.1598/JAAL.53.4.4

*Mason, L. H., Kubina, R., Valasa, L. L., & Cramer, A. (2010). Evaluating effective writing instruction of adolescent students in an emotional and/or behavioral support setting. *Behavioral Disorders, 35*(2), 140–156. http://www.jstor.org/stable/43153813

*Mason, L. H., Kubina, R., & Taft, R. (2011). Developing quick writing skills of middle school students with disabilities. *Journal of Special Education, 44*(4), 205–220. https://doi.org/10.1177/0022466909350780

Mason, L. H., Reid, R., & Hagaman, J. (2012). *Building comprehension in adolescents: Powerful strategies for improving reading and writing in content areas*. Brooks Publishing Co. Inc. https://products.brookespublishing.com/Building-Comprehension-in-Adolescents-P261.aspx

*Mason, L. H., Kubina, R., & Hoover, T. (2013a). Effects of quick writing instruction for high school students with emotional and behavioral disabilities. *Journal of Emotional and Behavioral Disorders, 21*(3), 163–175. https://doi.org/10.1177/1063426611410429

*Mason, L. H., Kubina, R. M., Kostewicz, D., Mong Cramer, A., & Datchuk, S. (2013b). Improving quick writing performance of middle school struggling learners. *Contemporary Educational Psychology, 38*(3), 236–246. https://doi.org/10.1016/j.cedpsych.2013.04.002

*Mong Cramer, A. M., & Mason, L. H. (2014). The effects of strategy instruction for writing and revising persuasive quick writes for middle school students with emotional and behavioral disorders. *Behavior Disorders, 40*(1), 37–51. https://doi.org/10.17988/0198-7429-40.1.37

National Governors Association (NGA) Center for Best Practices & Council of Chief State School Officers (CCSSO). (2010). *Common core state standards: English language arts*. NGA & CCSO. http://www.corestandards.org/ELA-Literacy/

Nelson, J. R., Smith, D. J., & Dodd, J. M. (1992). The effects of teaching a summary skills strategy to students identified as learning disabled on their comprehension of science text. *Education and Treatment of Children, 15*(3), 228–243. https://www.jstor.org/stable/42899268?seq=1#metadata_info_tab_contents

Rogers, L. A., & Graham, S. (2008). A meta-analysis of single-subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879–906. https://doi.org/10.1037/0022-0663.100.4.879

Rogevich, M. E., & Perin, D. (2008). Effects on science summarization of a reading comprehension intervention for adolescents with behavior and attention disorders. *Exceptional Children, 74*(2), 135–154. https://doi.org/10.1177/001440290807400201

*Saddler, B., Asaro-Saddler, K., Moeyaert, M., & Ellis-Robinson T. (2017). Effects of a summarizing strategy on written summaries of children with emotional and behavioral disorders. *Remedial and Special Education: RASE, 38*(2), 87–97. https://doi.org/10.1177/0741932516669051

Scruggs, T. E., Mastropieri, M. A., & Casto, G. (1987). The quantitative synthesis of single subject research: Methodology and validation. *Remedial and Special Education, 8*(2), 24–33. https://doi.org/10.1177/074193258700800206

Wood, K. D., & Harmon, J. M. (2001). *Strategies for integrating reading and writing in middle and high school classrooms*. National Middle School Association.

# Can Argumentative Writing Improve Math Knowledge for Elementary Students with a Mathematics Learning Disability?: A Single-Case Classroom Intervention Investigation

**Sharlene A. Kiuhara, Joel R. Levin, Malynda Tolbert, Megan Erickson, and Kenny Kruse**

**Abstract** Students with high incidence disabilities continue to perform considerably lower than their same-aged peers without disabilities in the areas of written expression and mathematical reasoning. This is especially concerning for students who come from diverse cultural, linguistic, and socio-economic backgrounds. We examined the effectiveness of a writing-to-learn mathematics intervention designed for students with a mathematics disability. The intervention incorporated the six-stages of Self-Regulated Strategy Development (SRSD) that targeted students' understanding of fractions as numbers and their argumentative writing and mathematical reasoning. A single-case multiple-baseline design was implemented with seven special education teachers who were randomly assigned to the staggered tiers of the design. Following 2 days of professional development and training, the teachers initiated the intervention in their classrooms. Visual and statistical analyses of the data revealed selected positive baseline-to-intervention phase changes in students' performance during implementation of SRSD. Implications and future directions of the research are discussed.

**Keywords** Argument writing · Writing-to-learn mathematics · Self-regulation strategy development

S. A. Kiuhara (✉) · M. Tolbert · M. Erickson · K. Kruse
University of Utah, Salt Lake City, UT, USA
e-mail: s.kiuhara@utah.edu

J. R. Levin
University of Arizona, Tucson, AZ, USA



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_11

# 1 Introduction

A recent meta-analysis conducted by Graham et al. (2020) has demonstrated that writing as a learning activity increases content-area learning for students in Grades 1–12, including mathematics. When learning mathematics, Resnick (1987) re-characterized the "3-Rs" as reasoning, writing, and arithmetic, as essential skills for children to carry out a series of steps for solving a mathematical problem flexibly and accurately. The National Governors Association and the Council of Chief State School Officers also recognized the importance of developing students' reasoning and language when learning mathematics and included eight practices in the Common Core State Standards for Mathematics (CCSS-M) to develop students' mathematical expertise (CCSS-M, 2010). Such activities involve constructing arguments, communicating their reasoning using clear definitions and explanations to justify their answers, and critiquing their peers' reasoning. Thus, students use language to become active agents in constructing new knowledge (Boscolo & Mason, 2001; Newell, 2006) and are likely to show larger gains in their problem-solving performance (Ball & Bass, 2003; National Mathematics Advisory Panel [NMAP], 2008).

In this chapter, we report a mathematics-intervention study that extends two feasibility studies presented by Hacker et al. (2019) and Kiuhara et al. (2020) in which the planning and constructing of written arguments facilitated the fraction learning and quality of mathematical reasoning of 5th and 6th-grade students with a mathematics learning disability (MLD). Although their previous research targeted students with and at-risk for a mathematics learning disability (MLD), the students in the current study were classified as having a MLD and were receiving specialized instruction in mathematics.

## *1.1 What Are Barriers for Students with MLD When Learning Fractions?*

Developing foundational knowledge about fractions is an essential building block for developing algebraic reasoning and predicting success for learning secondary mathematics (Bailey et al., 2015; Siegler et al., 2012). Proficiency with fractions requires students to make a conceptual shift from understanding whole numbers to two quantities that convey a single numerical value (e.g., $\frac{2}{3}$). This shift views the magnitude of the numerator and denominator as a unit rather than as separate numbers (Fuchs et al., 2013; Jordan et al., 2013). Although many 4th-graders fail to perform at or above the 25th percentile in mathematics, students with disabilities continue to score *below basic proficiency levels* compared to their same-aged peers without disabilities (National Center for Educational Statistics, 2015). Research has shown that students with MLD also experience difficulty using mathematical notation and accurately solving fraction problems involving multiple steps (Bryant &

Bryant, 2008; Geary, 2011). Students with MLD continue to struggle learning fractions through middle school (Mazzocco et al., 2013), which places them even further behind their same-grade peers without disabilities.

## *1.2 A Place for Writing Instruction in Math Class?*

Researchers have found that writing about one's learning promotes deeper engagement and active reasoning about new ideas (Bangert-Drowns et al., 2004; Graham et al., 2020; Hubner et al., 2006). Children use language to make sense and meaning of mathematical content, and language ability, in turn, supports and reinforces the conceptual knowledge needed for learning mathematics (Desoete, 2015; Vukovic & Lesaux, 2013). However, research has also indicated that students with MLD (a) may exhibit comorbidity with language development and learning mathematics (Korhonen et al., 2012); (b) often struggle with reasoning and communicating their ideas (NAEP, 2011); (c) often have difficulty with working memory, processing speed, and self-regulation when approaching and completing a task (Geary, 2011; Jitendra & Star, 2011; Jordan et al., 2013); and (d) have limited background and vocabulary knowledge for explaining or justifying their solutions and the solutions of their peers (Gersten et al., 2009; Krowka & Fuchs, 2017; Woodward et al., 2012). For students who may have MLD and speak English as a second language, using the process of writing to develop targeted academic vocabulary is essential (Cuenca-Carlino et al., 2018).

Writing provides students with a permanent record of their thinking (Hacker & Dunlosky, 2003). Therefore, writing about math learning may be helpful to students with MLD because writing allows students to reformulate and make sense of complex mathematical concepts. Engaging students in activities for constructing logical arguments and evaluating the logic or reasoning of peers are considered mathematical practices that develop mathematical expertise (CCSS-M, 2010; NMAP, 2008). However, some may argue that combining writing and mathematics may create further barriers to learning mathematics, especially for children with MLD.

## *1.3 Writing Strategies Instruction and Writing-to-Learn Mathematics*

Because we are interested in the benefits of using argumentative writing as a learning tool to develop mathematical knowledge and reasoning for students with MLD, we drew upon the extensive Self-Regulated Strategy Development (SRSD) evidence base that sequences explicit writing and self-regulation strategies for students with and without mild to moderate support needs (see Graham et al., 2012; Gillespie & Graham, 2014; Hebert & Powell, 2016; Harris & Graham, 2009). Briefly

described here, SRSD consists of six stages of instruction (Develop Background Knowledge, Discuss It, Model It, Memorize It, Support It, and Independent Performance) to help students independently manage their learning and writing process. The six stages focus on developing students' background knowledge for understanding the purpose, value, and characteristics of the strategies. Students learn to self-regulate and self-monitor their learning process, memorize the learning strategies, and give and receive feedback on their writing. Together with the teacher or other students, students actively engage in discussions, learn the terminology and vocabulary needed to articulate their mathematical understandings, and carry out the rhetorical structures of a specific writing genre, such as argumentation. During instruction, the stages are taught recursively, so the teacher can repeat stages and differentiate instruction to meet individual students' learning needs. The teacher models how to use the learning and self-regulation strategies and facilitates students' learning until each student can use the strategies independently and support their peers in their learning and writing process (Kiuhara et al., 2020).

Two strategies called FACT + $R^2C^2$ (herein referred to here as FACT) were designed to help students with and at-risk for MLD reason through a mathematics problem while constructing a written argument using the six stages of SRSD. FACT represents the following steps: F = *Figure* out plan (What is my task? Do I understand the problem? What do I need to know? What tools do I need?); A = *Act* on it (What reasons, evidence and support will I use? What words will I choose? How will I interpret my results?); C = *Compare* my reasoning with a peer's (What is similar or different? What are my reasons? Does it make sense? Can we make improvements?); and T = *Tie* it up in an argument which prompts students to go through the steps of RRCC. $R^2C^2$ represents the following steps: R = Did I *restate* the task?; R = Did I provide *reasons*, evidence, and support?; C = Did I provide a *counterclaim* that addresses an answer different from my own?; and C = Did I wrap it up with a *concluding* statement?

FACT was empirically tested in studies by Hacker et al. (2019) and Kiuhara et al. (2020). Kiuhara and colleagues initially tested the effects of FACT using a pre-posttest cluster-randomized controlled trial in which 10 teachers were randomly assigned to the FACT or business as usual conditions by teacher type (i.e., special educator or general educator) and grade (i.e., 4th, 5th, or 6th grades). The teachers in the FACT condition received 2 days of professional development (PD) before implementing the FACT lessons with their students. Treatment fidelity observations were conducted across 33% of the class sessions and was high (96% across all teachers, range 89–100%). The outcomes from pretest to posttest favored students in the FACT condition on a fraction test (Hedges' $g = 0.60$), quality of mathematical reasoning ($g = 1.82$); the number of argumentative elements ($g = 3.20$), and total words written ($g = 1.92$). We found that students with MLD in the FACT condition demonstrated greater gains in fraction scores from pretest to posttest than students without MLD ($n = 12$, $g = 1.04$). These findings showed promise for implementing a writing-to-learn math intervention in which students with MLD constructed arguments and critiqued the reasoning of their peers.

The study by Hacker et al. (2019) tested the effects of FACT on students with MLD using a single-case multiple baseline design (MBD) with an associated randomization test (Levin et al., 2018) to make informed decisions for further development of the lessons (Levin, 1992). Five special education teachers from different schools implemented the intervention with their 5th- and 6th-grade students ($n = 34$). The teachers received 2 days of PD before they were randomized to begin instruction at staggered points in time. A pre- posttest fraction measure indicated an average increase from the beginning to the end of the study ($d = 0.70$). However, Parker et al.'s (2014) rescaled Nonoverlap of All Pairs (NAP) effect-size indices (Gafurov & Levin, 2022) at the classroom-level produced mixed results for fraction accuracy, .32 (range, .13 to .69); mathematical reasoning, .61 (.08 to .97); number of rhetorical elements, .45 (.15 to .81); and total words written, .14 (.54 to .78). Teachers were observed every third lesson. Although overall treatment fidelity across the teachers was high (87%), some teachers were showing stronger student gains than others and some teachers showed higher treatment fidelity than others, ranging from 66% to 99%. Thus, further examination was needed to increase treatment fidelity and to identify the skills and knowledge needed for special education teachers to implement content-rich language and mathematics instruction.

### *1.4 Purpose of the Present Study*

The purpose of the present study was to build from the efficacy studies presented earlier by Hacker et al. (2019) with a similar single-case randomized design MBD and associated statistical analysis. Doing so here provided us with data to identify areas for additional refinement to the FACT lessons and the Writing-to-Learn PD protocol. The primary research question that guided our study was: To what extent do students with MLD who receive the FACT intervention demonstrate gains in fraction knowledge? We also wanted to understand the effects of the intervention on their quality of mathematical reasoning, argumentative elements, and total words written. We predicted that using argument writing as a tool during learning would encourage students to be more precise in communicating their mathematical reasoning (Graham et al., 2020; Resnick, 1987).

## 2 Method

### *2.1 Setting and Participants*

This study took place in a large and ethnically diverse school district located in the intermountain region of the United States. After receiving institutional review board approval to conduct this study, we contacted key district personnel who identified

licensed special education teachers. Eight special education teachers from eight different elementary schools consented to participate, and three sequential forms of teacher randomization took place prior to the start of the study. First, eight teachers were randomly assigned to four dyads. Then, each dyad was randomly assigned to predetermined dates during which the pair of teachers participated in 2 days of PD before implementing the FACT intervention. Finally, each teacher from the same dyad was randomly assigned to one of two adjacent dates within the eight intervention tiers of the MB design.

Absenteeism presented an unforeseen challenge for one teacher because of a family emergency, which considerably delayed the second of 2 days of her initial PD phase of the study (discussed below), along with her assigned intervention start date and continuing intervention-day absences. Therefore, that teacher and her students ($n = 4$) were not considered to be part of the formal study's data analysis and so the following results and discussion are based on the remaining seven teachers' classes.

Teachers provided specialized mathematics instruction in small groups to 5th and 6th-grade students with MLD for 45 min per day, four times per week. Five schools qualified for Title 1 services. The teachers had taught on average for 13.38 years (range = 5–27 years). Four teachers had a master's degree, two had a bachelor's degree, and one had a Juris doctorate. All teachers were female.

The 5th and 6th-grade students from each teacher's specialized mathematics class were invited to participate in the study if they (a) had at least one mathematics learning goal on their Individualized Education Plan (IEP), (b) were receiving specialized instruction in mathematics, and (c) were able to write a complete sentence on a standardized writing test.

The student participants ($n = 27$) included 15 (56%) 5th-graders, 12 (44%) 6th-graders, 14 (52%) male, and 10 (37%) English learners (ELs). Fifteen students (56%) were Latinx, 8 (30%) were White, 2 (7%) were Black, and 2 (7%) were multiracial. The students were administered two screening measures: (a) the mathematics subtest of the *Wide Range Achievement Test, 4th Ed.* (*WRAT-4*) (Wilkinson & Robertson, 2006) and (b) a writing subtest from *Wechsler Individual Achievement Test (WIAT), 3rd Ed.* (Wechsler, 2009), in which students wrote a short essay response to an expository prompt. Students' writing was scored for word count, theme development, and text organization. All students scored below the 14th percentile on the mathematics subtest ($M = 6.92$; range = 3.17–13.36). The average students' percentile ranking on the written expression subtest was 6.90 (range = 3.17–13.67).

## 2.2 Single-Case Intervention Design

A nonconcurrent MBD, with random assignment of classrooms to the staggered multiple-baseline levels (or "tiers") – a single-case design with a high degree of scientific credibility (Kratochwill & Levin, 2010; Levin, 1992; Levin et al., 2018; Slocum et al., 2022) – was adopted to evaluate the effectiveness of the intervention.

A MBD was implemented because the skills and knowledge that students may have acquired through participation were unlikely to be reversed, and this single-case intervention design allowed each teacher's class to serve as its own control (e.g., Horner & Odom, 2014; Kiuhara et al., 2017). The nonconcurrency of the design was necessitated by the irregularities of school-schedule timing, teacher absences, weather conditions, etc., for the commencement of each teacher's intervention on pre-specified calendar dates. Hence, the to-be-reported stagger of the design is represented by the number of pre- and post-intervention "sessions" rather than by actual chronological dates (Slocum et al., 2022). The purpose of implementing randomization in the design and associated statistical analysis was: (a) to improve the internal validity of the study, thereby providing a small-scale proxy to a randomized controlled trial (Kratochwill & Levin, 2010); and (b) to allow for a formal statistical assessment of the intervention's effectiveness based on a well-controlled Type 1 error probability while furnishing adequate statistical power for uncovering the intervention effects of interest (Levin et al., 2018).

The PD was provided by the third and fourth authors prior to the teachers' intervention start date. The fractions and argumentative writing components were guided by Hacker et al.'s (2019) and Kiuhara et al.'s (2020) PD protocols, which included PD components needed for teachers to "buy in" and effectively implement the intervention (e.g., multiple opportunities for active learning through modeling and practice and incorporating the same materials during PD that students have used previously) (Harris et al., 2012, 2014). We added discussion and activities that centered around developing and extending teachers' conceptual and procedural understandings of fractions, common misconceptions students exhibit when learning fractions, and using questioning to better engage students in learning mathematics (Borko et al., 2015; Polly et al., 2014; Jayanthi et al., 2017). The teachers attended the PD in pairs to encourage collective participation in a safe learning environment and allowed PD facilitators to provide feedback and establish rapport for lending support to teachers once the teachers began instruction during the intervention phase.

### *2.3 Intervention*

The FACT intervention consisted of five lessons using the six stages of SRSD described earlier (i.e., Develop Background Knowledge, Discuss It, Memorize It, Model It, Support It, and Independence), which embedded explicit instruction for improving students' writing knowledge and performance, self-efficacy, and strategic behavior with their writing and learning processes (Harris & Graham, 2009). We situated the language content of the FACT lessons around four mathematical practices: (a) construct arguments and critique the reasoning of peers, (b) make sense of problems and persevere in solving them, (c) use appropriate tools strategically, and (d) attend to precision by speaking and writing with precise mathematics vocabulary, describing relationships clearly, and calculating problems accurately (CCSS-M, 2010). The fraction content included understanding equivalence, comparing

fractions that refer to the same whole, composing and decomposing fractions, adding and subtracting with like and unlike denominators, and using equivalent fractions to solve problems with unlike denominators. The writing content included writing arguments to support claims with reasons and evidence. The fraction content was taught using a sequence of multiple representations (e.g., fraction blocks, number lines, area models, and numerical and mathematical notation) (Hughes et al., 2014; Witzel et al., 2003).

## *2.4 Measures*

### Distal Fraction Measure

The *easyCBM* Math Number and Operations assessment (Tindal & Alonzo, 2012) was modified to include 27 fraction items (0–27 points possible). The test items consisted of multiple-choice questions that focused on magnitude, equivalence, comparing two fractions and adding and subtracting fractions with like and unlike denominators. Two equivalent forms of the test were counterbalanced and administered to students before and at the end of the intervention phase and at the end of the intervention and post-intervention. Two independent scorers scored 100% on all assessments. Interrater reliability (IRR) between two scorers was 100%.

### Progress Monitoring Fractions

Twenty equivalent fraction probes were developed and administered to students each week during the study. Each fraction probe consisted of 14 questions (1 point each) and was divided into three sections: (a) placing fractions on a number line; (b) comparing the magnitude of two fractions; and (c) computational accuracy for adding and subtracting two fractions. Two independent scorers scored 100% on all assessments with an IRR of 100%.

### Progress Monitoring Writing

Twenty equivalent writing probes were administered weekly to measure students' ability to construct an argumentative paragraph in which students justified their solution to a fraction problem during untimed conditions. The students' written responses were scored for quality of mathematical reasoning, argumentative writing elements, and total words written following the procedures used by Hacker et al. (2019) and Kiuhara et al. (2020). Two independent scorers scored 100% of all assessments.

**Quality of Mathematical Reasoning**

Students' papers were scored holistically for the quality of mathematical reasoning following the procedures outlined in Kiuhara et al. (2020). The scoring index was from 0 to 12, with higher scores indicating higher reasoning quality and computational accuracy. For example, a score of 0 indicated that the student wrote no response or showed no understanding of the problem (e.g., I don't know). A score of 5 or 6 indicated a student solved the problem correctly but had gaps in reasoning to support their answer or provided little support in a counterclaim. The responses that scored 11 or 12 included a clear and focused understanding of the problem, an accurate and fully supported position, a supported counterclaim, and controlled writing with sequencing and strong transitions. Two independent scorers scored 100% of all assessments. Disagreements of more than + or –3 points between the two scorers were resolved by discussion. IRR was 90%.

**Argumentative Elements**

Students' papers were scored for six argumentative elements (0–36 points) following the procedures from Kiuhara et al. (2020). The papers were scored for the following: (a) included a statement that represented the mathematics task (e.g., My task is to compare the fractions $\frac{1}{4}$ and $\frac{2}{3}$.); (b) stated a claim or answer to the mathematics problem (e.g., I think that $\frac{2}{3}$ is greater than $\frac{1}{4}$.); (c) provided reasons and elaborations to support the claim (e.g., I used a number line, and $\frac{2}{3}$ is closer to one whole.); (d) provided a counterclaim or an incorrect solution to the problem (e.g., Others may argue that $\frac{1}{4}$ is greater $\frac{2}{3}$.); (e) provided reasons and elaborations to support the counterclaim (e.g., My peer might think $\frac{1}{4}$ is greater because 4 is greater than 3); and (f) provided a concluding statement (e.g., However, the number line shows that $\frac{2}{3}$ is greater than $\frac{1}{4}$ because it is closer the 1). Two independent scorers scored 100% of all assessments. Disagreements of more than + or –3 points between the two scorers were resolved by discussion. IRR was 96%.

**Total Words Written**

Students' writing was scored for total words written following the procedures outlined by Kiuhara et al. (2020). The third author typed verbatim the students' writing probes into a word processing program to eliminate bias for handwriting, spelling, and grammar errors (Kiuhara et al., 2012). A second scorer checked the typed probes for accuracy and resolved any differences with the first scorer. The word-processing program calculated the total number of words written.

### *2.5 Treatment Fidelity*

Each teacher was observed for a minimum of 33% of instructional sessions during the intervention phase of the study. Based on our findings from the MBD study reported by Hacker et al. (2019), we established that a teacher required further coaching and support if their treatment fidelity during the observed session did not reach a 90% criterion. Across all teachers, the average percent of instructional components that were in agreement between two independent raters was 92% with a range of 83–97%. On average, the teachers completed the intervention in 28 instructional days (range = 16–42 days).

### *2.6 Approach to Analysis*

As was already noted, a randomized single-case MBD and associated randomization tests (Levin et al., 2018) were adopted for the present study. Following those procedures, the eight participating teachers were randomly assigned to the design's staggered tiers, with a planned two outcome-observation stagger between tiers. The resulting design was implemented over a 16-week period and assured a minimum of 5 baseline (A-phase) observations and 4 intervention (B-phase) outcome observations for each teacher/classroom (but see the first paragraph of the immediately following section). All randomization tests were based on the average outcome performance of each teacher's class, which ranged in size from two to seven students, and those tests were directional (*viz.*, positing that the mean of the intervention phase would exceed the mean of the baseline phase) based on a Type I error probability of .05. In addition, because previous related research (e.g., Kiuhara et al., 2020) suggested that any expected intervention effects would not emerge in an immediate fashion, a two-outcome observation delay was built into the randomization-test analyses of all measures of mean between-phase change, based on the "data-shifting" procedure of Levin et al. (2017, p. 24) as operationalized in Gafurov and Levin's (2022) freely available *ExPRT* single-case randomization-test package.

## 3 Summary of Results

The to-be-summarized randomization-test analyses were conducted on five weekly outcome measures: Fractions Test, Mathematical Reasoning, Argumentative Elements, and Total Words Written. In addition, more fine-grained analyses were conducted on just the intervention's Lesson 1 and Lesson 2 content, which aligned with the first three stages of SRSD instruction (i.e., Develop Background Knowledge, Discuss It, Memorize It). As an overview, the analyses yielded mixed results, which in turn lead to inconclusive interpretations This is attributable in part to the present

MBD's actual between-tier stagger of only one outcome assessment per week (equivalent to four instructional lessons, depending on the teaching pace of the teacher) rather than the planned-for two staggers (eight instructional lessons). That was an unfortunate consequence of the aforementioned intervention-scheduling constraints for the 7 participating teachers. As a result, less clear differentiation among the present study's staggered tiers serves to reduce the typically high scientific credibility of a randomized MBD (Levin et al., 2018, Figs. 2 and 3, pp. 298–300). An illustrative graph depicting the results for one of the outcome measures, Mathematical Reasoning, is presented in Fig. 1, with the two-observation delayed phases labeled Phase A* and Phase B*.

*Primary analyses*. Wampold and Worsham's (1986) MBD randomization-test procedure, the most appropriate and statistically powerful MBD randomization test available (Levin et al., 2018), was applied to the four primary outcome measures but did not produce any statistically significant A- to B-phase improvements (*p*-values ranging from .94 to .16). That said, there was a large *general* between-phase statistical performance increase on three of the four measures, according to Busk and Serlin's (1992) "no-assumptions" average *d*s and Parker et al.'s (2014) average NAPs. After rescaling the NAPs to range from 0.0 to 1.00, they represent the average (across-classrooms) proportion of observations in the seven classrooms' baseline and intervention phases that do not overlap (for Mathematical Reasoning: $d = 6.86$, NAP = .74; for Argumentative Elements; $d = 6.14$, NAP = .78; and for Total Words Written: $d = 2.45$, NAP = .55). At the same time, a "visual analysis" (Kratochwill et al., 2021) of the outcome data revealed that four of the seven classes' observed improvements were not synchronous with the introduction of the intervention (see the Mathematical Reading outcomes for Teachers 1, 3, 4, and 6 in Fig. 1). As such, those improvements could have been attributable, at least to some extent, to the students' year-long growth spurred by teachers' provision of regular mathematics instruction during the baseline phase, which varied across the staggered intervention start dates. It is worth noting that the apparent precipitous decline and recovery of the intervention-phase mean for Teacher 5's class is based on the performance of only two students.

Two other sets of statistical analyses shed additional light on the efficacy of the FACT intervention. In one, analyses were conducted to examine whether the just-noted findings were consistent with another aspect of MBD logic, namely, that as each tier in the design exhibits an A- to B-phase improvement, the lower-level tiers do not – consistent with what has been referred to as a "vertical analysis" (see Kratochwill et al., 2021). An illustration of a differentiated vertical stair-step pattern can be appreciated through an examination of the Mathematical Reading measure outcomes presented in Fig. 1. A statistically more powerful modification of a stepwise between-tiers comparison MBD randomization-test procedure of Levin et al. (2018) that was originally proposed by Revusky (1967) is sensitive to assessing that pattern and yielded statistically significant results ($p$s < .023) on all measures except the fractions test ($p = .17$). Thus, it can be concluded that students' observed

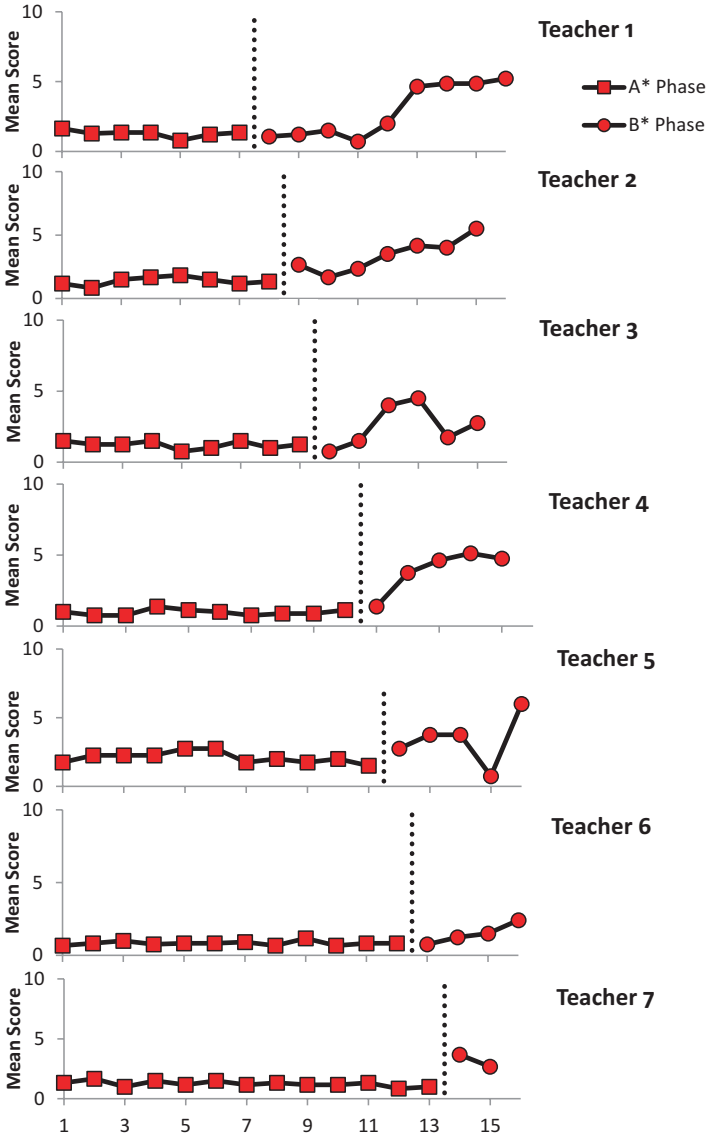



**Fig. 1** Quality of mathematical reasoning

mathematics performance increases occurred, at least to some extent, in accord with the hoped-for staggered tier improvements of successful MBD intervention outcomes.

Second, based on the aforementioned visual analyses of the data, it was noted that on two of the outcome measures (Mathematical Reasoning and Argumentative Elements) students exhibited a low stable level of performance during their baseline phase, followed by a steady increase in performance during their intervention phase.

In formal between-phase regression line "slope-change" analyses (Levin et al., 2021), no two-observation delay was included because Class 7 did not provide enough intervention-phase outcomes for a meaningful slope to be calculated. An illustration of the present study's baseline-to-intervention-phase slope increases may be seen for the Mathematical Reasoning measure presented in Fig. 1, where the actual intervention start point for each class occurred two observations earlier than is indicated there. From a statistical perspective, Wampold-Worsham MB randomization tests documented a significant increase in the slopes of students' performance between the baseline and intervention phases, from an across-students average slope of −.01 to .52, $p$ = .026, and from −.02 to .70, $p$ = .012, for Mathematical Reasoning and Argumentative Elements, respectively.

## *3.1 Fine-Grained Analyses: Or with a Tip of the Hat to an Esteemed Mentor, Colleague, and Friend: "Fine Grahamed" Analyses*

The preceding analyses were conducted on the comprehensive 20 parallel forms assessments that included all the FACT intervention content. During the earlier parts of most students' intervention phase, those assessments contained content that had not yet been covered by the classroom teachers. Thus, students' performance on such content for those assessments would not be expected to be responsive to any intervention effects. A seemingly more appropriate analysis was therefore conducted on just the lesson content that would have been covered for all students on all assessments. As with the mixed results already reported for the complete test content: (a) the Wampold-Worsham MBD test indicated that there was no statistically significant A- to B-phase improvement in students' performance ($p$ = .47); but (b) the Levin et al.'s (2018) Modified Revusky MBD test again produced a statistically significant appropriately differentiated vertical pattern of improvement ($p$ = .037).

**Pretest and Posttest Measures of Fraction Operations**

A correlated-samples $t$-test was conducted on the distal 27-item fraction operations measure that was administered to the 27 participating students at the beginning and the end of the study. Students exhibited a statistically significant increase on that measure ($p < .001$), amounting to a 3-item gain, $d = 1.30$. For this twice-administered measure, the observed improvement over the course of the school year does not reflect the true impact of the FACT + $R^2C^2$ intervention *per se* because it is entangled with regular school mathematics instructional content by teachers during the baseline phases and student growth.

## 4 Discussion

In our single-case classroom intervention investigation, we aimed to understand whether writing-to-learn mathematics through argument writing would increase students' fraction knowledge, their quality of mathematical reasoning, and the number of argumentative elements and total words written when provided with explicit strategies instruction using the six stages of SRSD. Based on previous research showing the benefits of writing-to-learn content (Bangert-Drowns et al., 2004; Graham et al., 2020; Hubner et al., 2006), as well as benefits of explicit writing strategies instruction (Gillespie & Graham, 2014; Graham et al., 2012), we expected that learning to construct written arguments and to critique the reasoning of peers would: (a) help students with MLD address their misconceptions when solving fraction problems and (b) allow those students to develop a deeper understanding of fractions (Hacker et al., 2019; Kiuhara et al., 2020). Our analyses uncovered selected classroom-level effects on all four outcome measures, similar to our findings in the MBD study presented by Hacker et al. (2019). That is, some teachers and certain outcome measures were associated with more student gains than others. This could be attributed to several factors.

First, the FACT intervention using SRSD combines multiple learning components (i.e., fraction learning and argumentative writing) with behavioral components (i.e., self-regulating learning behaviors and self-monitoring affect and motivation), which students acquire. These multiple components may require extended class sessions for teachers to ensure a student with MLD can independently use the strategies. Kiuhara et al. (2020) found that on average and with high treatment fidelity across teachers, the general education teachers in the FACT condition completed the lessons in 23 45-min sessions compared to the special education teachers. The latter completed the lessons in 30 45-min sessions. This indicated that students with MLD may require extended time reaching the Independent Stage of SRSD, as we found in the current study. For the seven teachers in this study, the instructional pacing to complete the lessons ranged from 16 days to 42 days. The teachers in the intervention phase for the longest amount of time demonstrated high treatment fidelity and did not require additional coaching and support; however, they had a large number of ELs with MLD in their classrooms (range 33–57%).

Based on their research using SRSD to teach ELs the process of argumentative writing, Cuenca-Carlino et al. (2018) suggest incorporating culture and specific language needs into SRSD. The quality of teacher-student and student-student interaction during learning influences language development (Gersten et al., 2009; Klingner & Soltero-Gonzales, 2009). For example, although SRSD addresses components of student agency (i.e., students are taught strategies to engage in discourse affecting their learning behaviors), it may also be necessary for teachers to understand how (a) to expand students' ability to draw on their own languages and cultural experiences for reflecting on their original or novel approaches when solving problems (Klinger and Soltero-Gonzales) and (b) to facilitate ways for ELs to think aloud using their first language or with their peers when solving math problems (Garcia &

Sylvan, 2011). Conversely, the teacher who completed the FACT lessons in the shortest amount of time (i.e., 16 days) reported in her teaching log that she did not engage students in any supplemental learning activities and simply followed the manualized protocol for implementing the lessons almost verbatim. Although her treatment fidelity was high (91%), we learned that PD should include space for teachers to discuss the recursive approach to teaching SRSD and decision-making processes for determining when teachers should reteach, modify, or reorder the stages of SRSD (Cuenca-Carlino et al., 2018). Students benefit only when they are provided multiple opportunities to connect mathematical language and communicate their learning (Bangert-Drowns et al., 2004; Klein, 1999).

Second, one of our concerns from the previous study reported by Hacker et al. (2019) was to increase teacher "buy in" to implement a novel instructional intervention such as FACT. One way we addressed this was here to devote time during PD to discuss with teachers their experiences they had teaching fractions to students with MLD and some of the common misconceptions that students have when learning fractions (Borko et al., 2015). We also included a protocol for using questioning as a way for teachers to assess students' understanding and to provide opportunities for students to articulate their understandings (Borko et al., 2015; Polly et al., 2014; Jayanthi et al., 2017). We found that treatment fidelity was high across all seven teachers. It also proved beneficial for us to have established a minimum criterion level for knowing when to provide teachers with additional coaching during the implementation phase. For example, one teacher who did not meet the criterion of treatment fidelity (*viz*., greater than 85%) on more than one observed lesson had transferred the learning activities from the teacher's manual to a SmartBoard that the teacher used daily. This resulted in some of the FACT content not having been easily transferred and therefore was skipped by the teacher during instruction. The third author met with the teacher regularly to address the fidelity components that were missed during the observation and answered any questions the teacher had about the lesson.

Third, from a methodological perspective, our MBD with random assignment of teachers to tiers and associated randomization statistical tests represents a rigorous single-case intervention design (Kratochwill & Levin, 2010; Levin et al., 2019). Although logistical constraints and exigencies rendered it not possible to implement the intervention optimally here, future scientifically credible and statistically powerful single-case research that focuses on teaching SRSD should strive to include a greater stagger of the tiers' intervention start points, along with statistical randomization-test models that require a start point for each tier member that is randomly sampled from two or more acceptable potential intervention start points (see, for example, Levin & Ferron, 2021).

Finally, we did not account for the type of instruction the students received during the baseline phase, which ranged from 19 to 42 days, depending on the teacher's staggered intervention start date and the possible overlap with the district's pacing calendar for teaching or reviewing fractions. Taking all these various design and operational challenges into consideration (Kratochwill et al., 2021) leads directly to a general guiding principle that has emerged from our present and earlier FACT/

SRSD studies (see Hacker et al., 2019; Kiuhara et al., 2020). As with large-scale randomized controlled trials studies, in single-case classroom-based intervention investigations one's mantra should be: "Plan for the best but always expect the unexpected."

In sum, our previous FACT intervention studies found that engaging students with MLD in activities that focused on communicating with precise mathematical language and constructing arguments increased students' quality of mathematical reasoning, argumentative elements, and total words written (Hacker et al., 2019; Kiuhara et al., 2020). Although certainly not conclusive, the collective visual and statistical outcomes from this study are encouraging, in that they are suggestive of positive effects associated with using writing-to-learn mathematics and SRSD. They should provide continued motivation for classroom-based interventionists to extend this line of research in teaching mathematics and writing, especially for a wide range of ELs with MLD.

**Acknowledgements** We wish to thank the teachers and students who participated in this study, Drs. Noelle Converse and Leah Vorhees for their support, and the Utah State Board of Education Special Education Services for providing the funding for this project.

## References

Bailey, D. H., Zhou, X., Zhang, Y., Cui, J., Fuchs, L., Jordan, N. C., Gersten, R., & Siegler, R. S. (2015). Development of fraction concepts and procedures in U.S. and Chinese children. *Journal of Experimental Psychology, 129*, 68–83.

Ball, D. L., & Bass, H. (2003). Making mathematics reasonable in school. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 27–44). National Council of Teachers of Mathematics.

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*, 29–58.

Borko, H., Jacobs, J., Koellner, K., & Swackhamer, L. E. (2015). *Mathematics professional development: Improving teaching using the problem-solving cycle and leadership preparation models*. National Council of Teachers of Mathematics.

Boscolo, P., & Mason, L. (2001). Writing to learn, writing to transfer. In G. Rijlaarsdam (Series Ed.), *Studies in writing: Volume 7. Writing as a learning tool: Integrating theory and practice* (pp. 83–104). Kluwer Academic Publishers.

Bryant, B. R., & Bryant, D. P. (2008). Mathematics and learning disabilities. *Learning Disability Quarterly, 31*, 3–11.

Busk, P. L., & Serlin, R. C. (1992). Meta-analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single case research design and analysis: New directions for psychology and education* (pp. 187–212). Lawrence Erlbaum Associates, Inc.

*Common Core State Standards for Mathematics.* (2010). Retrieved January 1, 2014, from www.corestandards.org

Cuenca-Carlino, Y., Gozur, M., Jozwik, S., & Krissinger, E. (2018). The impact of self-regulated strategy development on the writing performance of English learners. *Reading and Writing Quarterly, 34*, 248–262. https://doi.org/10.1080/10573569.2017.1407977

Desoete, A. (2015). Language and math. In P. Aunio, R. Mononen, & A. Laine, Mathematical learning difficulties—Snapshots of current European research. *LUMAT, 5*, 647–674.

Fuchs, L. S., Schumacher, R. F., Long, J., Namkung, J., Hamlett, C., Cirino, P. T., Jordan, N. C., Siegler, R., Gersten, R., & Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology, 105*, 683–700.

Gafurov, B. S., & Levin, J. R. (2022). *ExPRT* (*Excel Package of Randomization Tests): Statistical analyses of single-case intervention data*. Current Version 4.3 (November 2022) is retrievable from the *ExPRT* website at http://ex-prt.weebly.com

Garcia, O., & Sylvan, C. E. (2011). Pedagogies and practices in multilingual classrooms: Singularities in pluralities. *Modern Language Journal, 95*, 385–400.

Geary, D. C. (2011). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental & Behavioral Pediatrics, 32*, 250–263.

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research, 79*(3), 1202–1242. https://doi.org/10.3102/003454309334431

Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children, 80*, 454–473.

Graham, S., McKeown, D., Kiuhara, S. A., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879–896.

Graham, S., Kiuhara, S. A., & MacKay, M. (2020). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research, 90*(2), 179–226. https://doi.org/10.3102/0034654320914744

Hacker, D. J., & Dunlosky, J. (2003). Not all metacognition is created equal. *Problem- based learning for the information age*. *New Directions for Teaching and Learning, 95*, 73–80.

Hacker, D. J., Kiuhara, S. A., & Levin, J. R. (2019). A metacognitive intervention for teaching fractions to students with or at-risk for learning disabilities in mathematics. Special issue on metacognition in mathematics education. *ZDM: The International Journal on Mathematics Education, 51*(4), 601–612. https://doi.org/10.1007/s11858-019-01040-0

Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *Teaching and Learning Writing, 6*, 113–135.

Harris, K. R., Lane, K. L., Graham, S., Driscoll, S. A., Sandmel, K., Brindle, M., & Schatschneider, C. (2012). Practice-based professional development for self-regulated strategies development in writing: A randomized controlled study. *Journal of Teacher Education, 63*(2), 103–119. https://doi.org/10.1177/002248711429005

Harris, K. R., Graham, S., & Adkins, M. (2014). Practice-based professional development and self-regulated strategy development for Tier 2, at-risk writers in second grade. *Contemporary Educational Psychology, 40*(1), 5–16.

Hebert, M. A., & Powell, S. R. (2016). Examining fourth-grade mathematics writing: Features of organization, mathematics vocabulary, and mathematical representations. *Reading and Writing, 29*, 1511–1537. https://doi.org/10.1007/s11145-016-9649-5

Horner, R. H., & Odom, S. L. (2014). Constructing single-case research designs: Logic and options. In T. R. Kratchowill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 27–51). American Psychological Association.

Hubner, S., Nuckles, M., & Renkl, A. (2006). Prompting cognitive and metacognitive processing in writing-to-learn enhances learning outcomes. In *Proceedings of the 28th annual conference of the cognitive science society* (pp. 357–362). Erlbaum.

Hughes, E. M., Witzel, B. S., Riccomini, P. J., Fries, K. M., & Kanyongo, G. Y. (2014). A meta-analysis of algebra interventions for learners with disabilities and struggling learners. *Journal of the International Association of Special Education, 15*(1), 36–47.

Jayanthi, M., Gersten, R., Taylor, M. J., Smolkowski, K., & Dimino, J. (2017). *Impact of the developing mathematical ideas professional development program on grade 4 students' and teachers' understanding of fractions* (REL 2017-256). Regional Educational Laboratory Southeast.

Jitendra, A. K., & Star, J. R. (2011). Meeting the needs of students with learning disabilities in inclusive mathematics classrooms: The role of schema-based instruction on mathematical problem solving. *Theory Into Practice, 50*, 12–19.

Jordan, N. C., Hansen, N., Fuchs, L. S., Siegler, R. S., Gersten, R., & Micklos, D. (2013). Developmental predictors of fraction concepts and procedures. *Journal of Experimental Child Psychology, 116*, 45–58.

Kiuhara, S. A., O'Neill, R., Hawken, L. S., & Graham, S. (2012). The effectiveness of teaching 10th grade students with a disability STOP, AIMS, and DARE for planning/drafting persuasive text. *Exceptional Children, 78*(3), 335–355.

Kiuhara, S. A., Kratochwill, T. R., & Pullen, P. C. (2017). Designing robust experimental single-case design research. In J. M. Kauffman, D. P. Hallahan, & P. C. Pullen (Eds.), *Handbook of special education* (2nd ed., pp. 116–136). Routledge.

Kiuhara, S. A., Gillespie Rouse, A., Dai, T., Witzel, B., Morphy, P., & Unker, B. (2020). Constructing written arguments to develop fraction knowledge. *Journal of Educational Psychology, 112*(3), 584–607. https://doi.org/10.1037/edu0000391

Klein, P. D. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Educational Psychology Review, 11*, 203–270.

Klingner, J. K., & Soltero-Gonzalez, L. (2009). Culturally and linguistically responsive literacy instruction for English language learners with learning disabilities. *Multiple Voices for Ethnically Diverse Exceptional Learners, 12*, 4–10.

Korhonen, J., Linnanmaki, K., & Aunio, P. (2012). Language and mathematical performance: A comparison of lower secondary school students with different level of mathematical skills. *Scandinavian Journal of Educational Research, 56*(3), 333–344. https://doi.org/10.1080/00313831.2011.599423

Kratochwill, T. R., & Levin, J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15*, 124–144. https://doi.org/10.1037/a0017736

Kratochwill, T. R., Horner, R. H., Levin, J. R., Machalicek, W., Ferron, J., & Johnson, A. (2021). Single-case design standards: An update and proposed upgrades. *Journal of School Psychology, 89*, 91–205.

Krowka, S. K., & Fuchs, L. S. (2017). Cognitive profiles associated with responsiveness to fraction intervention. *Learning Disabilities Research & Practice, 32*, 216–230.

Levin, J. R. (1992). Single-case research design and analysis: Comments and concerns. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case research design and analysis: New directions for psychology and education* (pp. 213–224). Erlbaum.

Levin, J. R., & Ferron, J. M. (2021). Different randomized multiple-baseline models for different situations: A practical guide for single-case intervention researchers. *Journal of School Psychology, 86*, 169–177.

Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2017). Additional comparisons of randomization-test procedures for single-case multiple-baseline designs: Alternative effect types. *Journal of School Psychology, 63*, 13–34.

Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2018). Comparison of randomization-test procedures for single-case multiple-baseline designs. *Developmental Neurorehabilitation, 21*, 290–311.

Levin, J. R., Kratochwill, T. R., & Ferron, J. M. (2019). Randomization procedures in single-case intervention research contexts: (Some of) "the rest of the story". *Journal of the Experimental Analysis of Behavior, 112*, 334–348.

Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2021). Investigation of single-case multiple-baseline randomization tests of trend and variability. *Educational Psychology Review, 33*, 713–737.

Mazzocco, M. M. M., Myers, G. F., Lewis, K. E., Hanich, L. B., & Murphy, M. M. (2013). Limited knowledge of fraction representations differentiates middle school student with mathematics learning disability (dyscalculia) versus low math achievement. *Journal of Experimental Child Psychology, 115*, 371–387.

National Center for Education Statistics, Institute of Education Sciences, National Assessment of Educational Progress for Mathematics. (2015). *The nation's report card: Mathematics 2015*. Retrieved from https://nces.ed.gov/nationsreportcard/mathematics/

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. U.S. Department of Education. Retrieved from www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

Newell, G. E. (2006). Writing to learn: How alternative theories of school writing account for student performance. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 235–247). Guilford Press.

Parker, R. I., Vannest, K. J., & Davis, J. L. (2014). Non-overlap analysis for single-case research. In T. R. Kratochwill & J. R. Levin (Eds.), *Single-case intervention research: Methodological and statistical advances* (pp. 127–151). American Psychological Association.

Polly, D., Neale, H., & Puglee, D. K. (2014). How does ongoing task-focused mathematics professional development influence elementary school teachers' knowledge, beliefs and enacted pedagogies? *Early Childhood Educational Journal, 42*, 1–10.

Resnick, L. B. (1987). *Education and learning to think*. National Academy Press.

Revusky, S. H. (1967). Some statistical treatments compatible with individual organism methodology. *Journal of the Experimental Analysis of Behavior, 10*, 319–330.

Siegler, R. S., Duncan, G. J., Davis-Kean, P. E., Duckworth, K., Claessens, A., Engel, M., Susperreguy, M. I., & Chen, M. (2012). Early predictors of high school mathematics achievement. *Psychological Science, 23*, 691–697.

Slocum, T. A., Pinkelman, S. E., Joslyn, P. R., & Nicols, B. (2022). Threats to internal validity in multiple-baseline design variations. *Perspectives on Behavior Science*. https://doi.org/10.1007/s40614-00326-1

Tindal, G., & Alonzo, J. (2012). *easyCBM*. Houghton Mifflin Harcourt.

Vukovic, R. K., & Lesaux, N. K. (2013). The language of mathematics: Investigating the ways language counts for children's mathematical development. *Journal of Experimental Child Psychology, 115*, 227–244.

Wampold, B., & Worsham, N. (1986). Randomization tests for multiple-baseline designs. *Behavioral Assessment, 8*, 135–143.

Wechsler, D. (2009). *Wechsler individual achievement test* (3rd ed.). Pearson.

Wilkinson, G. S., & Robertson, G. J. (2006). *Wide range achievement test* (4th ed.). PAR.

Witzel, B. S., Mercer, C. D., & Miller, M. D. (2003). Teaching algebra to students with learning difficulties: An investigation of an explicit instruction model. *Learning Disabilities Research & Practice, 18*(2), 121–131.

Woodward, J., Beckmann, S., Driscoll, M., Franke, M., Herzig, P., Jitendra, A., Koedinger, K. R., & Ogbuehi, P. (2012). *Improving mathematical problem solving in grades 4 through 8: A practice guide* (NCEE 2012-4055). National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications_reviews.aspx#pubsearch/

# Handwriting Difficulties: Different Manifestations and Underlying Functions


**Naomi Weintraub**

**Abstract** Handwriting difficulties (HD) are manifested in poor legibility, speed or both, which seem to represent different aspects of handwriting performance; product and process, respectively. As such, they may stem from different deficits in underlying functions. This study examined the distribution of HD manifestations among higher-education students, and the relationship between language, working memory as well as perceptual- and grapho-motor skills and handwriting legibility and speed. The sample included 110 higher-education students (18- to 33-years old), of whom 66 had specific learning disabilities (SLD) and HD and 44 were typically developing. Students were administered a handwriting test and various tests measuring the underlying functions. Results showed three distinct groups of students with HD. Most (62.1%) had poor legibility but not poor speed. Others had only slow handwriting (24.2%) and a minority had both (13.6%). Working memory was found to predict both poor legibility and speed. Yet, whereas having slow handwriting was also predicted by poor spelling and gender, having poor legibility was related to visual-spatial motor organization. These findings indicate that higher-education students with HD are a heterogeneous group. Therefore, for research and educational purposes distinctions should be made with reference to the different HD groups, because each appear to be associated with different underlying deficits.

**Keywords** Dysgraphia · Legibility · Handwriting speed · Working-memory · Higher-education students

N. Weintraub (✉)
School of Occupational Therapy, Hebrew University, Jerusalem, Israel
e-mail: naomi.weintraub@mail.huji.ac.il



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_12

# 1 Introduction

Handwriting difficulties (dysgraphia) refers to an impaired ability to produce handwritten words or letters (handwriting disorder; Berninger et al., 2015b). The prevalence of writing disorders is estimated to be between 7% and 15% of school-age students (Döhla & Heim, 2016). Although this prevalence is not much different (5–17%; Shaywitz & Shaywitz, 2005) from that of reading and spelling difficulties (dyslexia), it is one of the lesser researched manifestations of specific leaning disability (SLD; Katusic et al., 2009). Furthermore, similar to other neuro-developmental disorders, handwriting difficulties have been found to continue into adulthood (Cousins & Smyth, 2003; Rosenberg-Adler & Weintraub, 2020).

Handwriting difficulties may have both functional and emotional effects on individuals. They may limit students' higher-level writing performance such as their ability to express knowledge or ideas (McCloskey & Rapp, 2017) or note-taking (Peverly et al., 2013). Furthermore, handwriting difficulties may often cause frustration and affect students' self-esteem as well as motivation (Engel-Yeger et al., 2009). Consequently, they often minimize or avoid performing writing assignments, which may tax their grades (Feng et al., 2019). Students' grades may also be negatively affected due to legibility bias (Graham et al., 2011; Greifeneder et al., 2011). Finally, handwriting difficulties may influence a person's performance in daily activities (Tal-Saban & Weintraub, 2019) including career choice or advancement (McCloskey & Rapp, 2017).

## *1.1 Measures of Handwriting Performance*

The two main measures of handwriting are speed and legibility. Handwriting speed represents the process or production aspect of handwriting, and is often measured as the number of words or characters written per minute (Rosenblum et al., 2003). Legibility reflects the quality of the handwritten product, and is often defined as the degree to which the text is *perceived* as readable (Graham, 1986; Greifeneder et al., 2011). This perception is influenced by letter formation and the uniformity of spatial components (e.g., letter size, spacing within and between words, letter slant; Graham et al., 2006; Weintraub et al., 2007) and the texts' neatness (e.g., erasures). Another definition of handwriting legibility is the *actual* readability of the text (Graham, 1986; Rosenberg-Adler & Weintraub, 2020). Readability of the text may be influenced by letter formation and spelling (Graham et al., 2011).

The fact that the two handwriting measures represent different aspects of handwriting performance suggests that students with handwriting difficulties may comprise a heterogeneous group, where some may have either poor legibility or slow handwriting and others may have both. Furthermore, these difficulties may result

from deficits in different functions or skills. To date, most studies comparing handwriting performance of students with and without handwriting difficulties, did not differentiate between students with poor legibility versus those with poor speed (e.g., Duiser et al., 2013), or related to either one or the other. The few studies that did compare the underlying measures related to either legibility versus speed (e.g., Klein et al., 2011; Salameh-Matar et al., 2018) found that different underlying functions were associated with handwriting speed versus legibility. Both studies reported that visual-motor integration had a unique contribution in predicting legibility. By contrast, speed was predicted by the alphabet task (Salameh-Matar et al., 2018) or fine-motor coordination (Klein et al., 2011). Yet these studies focused on elementary-school students.

## *1.2 Conceptual Framework*

With the purpose of better understanding the possible underlying functions related to handwriting legibility and speed, we based this study in the psychomotor model by Van Galen (1991) and its later suggested revisions (i.e., Graham et al., 2006; Kandel & Spinelli, 2010). This model postulates that handwriting production is hierarchically controlled, and is based on parallel processing, where higher-level output, forms the foundation for the subsequent lower levels' processing. The first stages are psycholinguistic processes, including generating ideas, translating them into words, and selecting a word to be written.

Next, an orthographic representation of the selected word is retrieved from long-term memory, and phoneme-to-grapheme translation occurs, resulting in an abstract letter sequence of the word that is stored in the orthographic or graphemic buffer (Kandel & Spinelli, 2010; Van Galen, 1991). This is followed by grapheme to allograph conversion (i.e., the specific letter shape such as upper or lower case), activating the corresponding internal motor programs required for writing the selected allographs (e.g., letter form, size, rotation, direction). The next stage is visual spatial arrangement parameter setting (Graham et al., 2006), in which the correct placement of the letters on the page (e.g., in relation to the line, the margins or a previously written letter) is determined. Van Galen's (1991) model further asserts that in the final stage, the internal motor programs are converted into grapho-motor plans, specifying the muscles required, as well as the force, speed and slant necessary for producing the letters. One stage that was not clearly defined in this model is the grapho-motor production stage (McCloskey & Rapp, 2017), namely the movement of the hand and fingers for producing the strokes related to the different letters (Palmis et al., 2017). Nevertheless, the handwriting model of Van Galen (1991) and its revisions, delineate the need for integrating language, cognitive and perceptual-motor functions while handwriting.

## *1.3 Underlying Mechanisms of Handwriting*

Over the years, various studies examined the relationship between underlying functions and handwriting performance.

### Language Functions

Berninger (2000) described the different types of languages; language by ear, by mouth, by eye, and language by hand, referring to handwriting (or keyboarding). The four language systems share underlying functions (Berninger, 2000; Döhla & Heim, 2016). For example, a positive association was reported between reading and handwriting speed (Berninger et al., 1992, 2002; Bosga-Stork et al., 2016). Other researchers reported handwriting difficulties among students with dyslexia (Berninger et al., 2008; Graham et al., 2020; Sumner et al., 2013). Similarly, several researchers found that students' spelling was associated with handwriting speed (Abbott et al., 2010; Bosga-Stork et al., 2016; Limpo & Alves, 2013; Wicki et al., 2014), and legibility (Wicki et al., 2014) or contributed to the predictions of students' handwriting status (Rodríguez & Villarroel, 2017).

By contrast, other studies have shown that handwriting, spelling and reading develop autonomously and require unique language skills (Berninger et al., 2002; Bosga-Stork et al., 2016). For example, students may have reading or spelling difficulties but not handwriting difficulties (Berninger et al., 2015b; Döhla & Heim, 2016). Nevertheless, most of these studies were performed among elementary school students. Moreover, poor handwriters were usually a mixed group of students with both poor legibility and slow speed. Therefore, more evidence is required for examining the association between reading and spelling skills and each of the handwriting measures among higher-education students.

### Working Memory

Based on the conceptual model of handwriting production described above, it appears that working memory plays an important role in handwriting production, and may affect the handwritten product (e.g., letter formation, spatial organization). Various studies have shown an association between working memory and handwriting production (e.g., Adi-Japha et al., 2007; Olive & Kellogg, 2002). For example, in a study among students in 4th–9th grades with dysgraphia, Berninger et al. (2015a) found that the students had impaired working memory in the orthographic word storage and processing, as well as in focused/selective attention. Other studies have shown that handwriting performance was associated with attention deficits (Capodieci et al., 2018; Graham et al., 2016). In a study among higher-education students, Peverly et al. (2013) found a significant correlation between attention span and handwriting speed.

Another measure which taps working memory in the context of handwriting is the Alphabet task, which requires writing the alphabet letters from memory in their correct sequence and form. This task involves retrieving the alphabet letters (allographs) from long-term memory, storing them in their correct sequence while matching them with the motor programs, and then producing them in the correct order and form (Abbott & Berninger, 1993). Many studies have shown a relationship between the Alphabet task and handwriting speed (Berninger et al., 1997; Graham et al., 2000; Medwell & Wray, 2014). Yet, the association between the Alphabet task and legibility is scarce. In one study, Weintraub and Graham (2000) did not find that the Alphabet task had a significant contribution in predicting students' status as poor or good legibility.

As can be seen, in the different studies working memory was operationalized and measured in different ways. In some studies, the focus was attention, whereas in other studies working memory was directly related to the handwriting process. Furthermore, only few studies examined the association of working memory and handwriting legibility, and most related to school-age students.

### Perceptual- and Grapho-Motor Functions

As stated above, handwriting production entails two stages that were not described in the model proposed by Van Galen (1991); the spatial arrangement parameter setting (Graham et al., 2006) and the grapho-motor execution stage, in which the actual handwriting occurs. The former, requires writing the letters in their correct form, size and location in relation to the line and the other letters (Graham et al., 2006). Most studies have shown that students' scores on visual-motor integration tests were significantly correlated with handwriting legibility (Duiser et al., 2013; Weintraub & Graham, 2000), yet others did not find such associations (Prunty et al., 2016).

The grapho-motor execution stage of handwriting requires the use of the fingers to manipulate a writing tool (pencil or pen) while executing the motor program in a sequential, accurate and fluent manner. This act requires fine motor sequencing and control (McCloskey & Rapp, 2017; Smits-Engelsman et al., 2001). In fact, studies reported that motor control and finger dexterity were related to handwriting legibility (Duiser et al., 2013; Weintraub et al., 2010) or speed (Berninger & Rutberg, 1992; Salameh-Matar et al., 2018). However, the different studies measured motor control in various methods such as finger dexterity (e.g., Berninger & Rutberg, 1992; Weintraub & Graham, 2000) or specific grapho-motor tasks (e.g., Kaiser et al., 2009; Salameh-Matar et al., 2018). Therefore, the relationship between motor control and handwriting performance needs to be further explored.

The handwriting model and its revisions, as well as the studies mentioned above support the assertion that problems in any of the handwriting production stages due to deficits in language, cognitive, or perceptual-motor-functions may affect handwriting legibility, speed or both. Yet, the contribution of the different underlying functions in predicting handwriting legibility versus speed has not been sufficiently explored.

### *1.4 Study Purpose*

The overall purpose of this study was to enhance the understanding of handwriting difficulties among higher-education students. This study had several specific objectives: (a) to describe the distribution of handwriting difficulty manifestations (legibility versus speed) among these students; (b) to examine the relationship between basic underlying functions: language, working memory and perceptual-and graphomotor functions and handwriting legibility and speed; and (c) to examine the contribution of these underlying functions in predicting students' handwriting status (poor legibility versus control and slow handwriting versus control). I expected to find students with poor legibility, whose handwriting speed was intact, and other students, whose handwriting speed was slow, but legible, while a third group encountering difficulties in both measures. I also hypothesized that different functions were related to legibility versus speed. Finally, I expected to find different predictors of handwriting status in relation to poor legibility as opposed to slow handwriting.

## 2 Method

### *2.1 Participants*

The sample included 110 higher-education students between the ages of 18–33, of whom 66 had SLD with handwriting difficulties (henceforth, HD) and 44 were typically developing (TD). Most (74.3%) were first year students from various faculties (e.g., humanities, social sciences, law, medicine and natural sciences). Students were included in the study if they met the general inclusion criteria: (a) did not have a physical condition involving their upper limb, neck, or back (e.g., inflammation, or fractures) in the past 6 months; and (b) were not taking medications that could affect their handwriting performance (except for medication for attention deficit hyperactivity disorder; ADHD). Students were included in the HD group if they: (a) had been diagnosed in childhood or adolescence with SLD and HD and were currently eligible for test accommodations (based on self-report); and (b) were currently identified as having handwriting difficulties based on a standardized test.

Students were included in the TD group if: (a) they were not diagnosed with a SLD, ADHD or Developmental Coordination Disorder, and were not eligible (in the past or currently) for test accommodations (based on self-report); and (b) their handwriting performance was within the normal range, based on a standardized test. The sample was elicited from a database that was accumulated from three past studies that focused on handwriting and keyboarding of higher-education students, all using the same inclusion criteria as described above and recruitment methods (see Procedure section). In all studies, participants were recruited through a convenience sampling method.

## 2.2 Measures

### Background Questionnaire

This self-report Background Questionnaire was developed for use in our lab studies with higher-education students for the purpose of obtaining demographic information (e.g., gender, age), developmental, medical, and educational background information as well as routines and practices in the use of computers (e.g., typing hours per day).

### Handwriting Performance of Post-secondary Students Evaluation (HaPPS; Weintraub et al., 2012)

The HaPPS is a standardized test of handwriting performance of higher-education students. It consists of three tasks: A 10-min copying task; a 3-min writing to dictation task; and a 5-min expository writing composition task. Each task is scored for speed (i.e., number of characters) and legibility (i.e., percent of unreadable words; higher scores represent poorer legibility). The cutoff point for illegible handwriting is $Z > (2)$ in at least two of the handwriting tasks. The cutoff point for slow handwriting is $Z < (-1.5)$ in the copying and dictation tasks. The HaPPS was found to have a medium-high and statistically significant inter-rater reliability for legibility, $.61 < r < .87$, $p < .01$; and for speed, $.76 < r < .90$, $p < .01$, and to discriminate between students with and without handwriting difficulties both in speed, $F_{(4,112)} = 16.58$, $p = .001$; and in legibility, $F_{(4,110)} = 11.80$, $p = .001$ (Yusilles, 2012). In this study all tasks were used for forming the handwriting status groups, and the copying task was used for examining the relationship between underlying functions and handwriting performance.

### MATAL: Diagnosis of Learning Disabilities and Attention Disorders (Ben-Simon & Inbar-Weiss, 2012)

The MATAL is a computerized battery of standard tests for the diagnosis of learning disorders between the ages of 16–30. It consists of 20 cognitive tests. In this study we used two basic language tests: Oral text reading (scored for speed) and Spelling (while writing to dictation, scored for error rate). Both scores were transformed to *Z*-scores. Administration time for this test is approximately 30-min. This test was used to measure the initial stages of handwriting production, as described by Van Galen (1991).

### Working Memory

In this study we used two working memory tests. The first is based on the *Working Memory scale* of the Behaviour Rating Inventory of Executive Function-Adult version (BRIEF-A; Roth et al., 2005), which is a standardized self-report questionnaire that measures adults' executive functions in their daily activities.

The working memory scale consists of seven items that capture the capacity to actively hold information in mind for the purpose of completing a task or generating a response. The items themselves focus on the persons' attention and concentration capacity.

The second measure was the *Alphabet task* (Berninger & Rutberg, 1992). While this task is not direct measure of working memory, it evaluate students' efficiency in retrieving the alphabet letters from memory in their correct form and sequence, and storing them in an orthographic buffer prior to correctly producing them (Brooks et al., 2011). Thus, it requires the integration of orthographic retrieval and motor production. In this study the task was scored as the time it took the students to write the Hebrew alphabet letters in their correct form and sequence. This task taps various stages of the model by Van Galen (1991), including grapheme-allograph conversion, allograph-motor program association and letter production.

**Rey Osterrieth Complex Figure (ROCF; Rey, 1941 in Shin et al., 2006)**

The ROCF measures visuo-spatial-motor organization and memory. In the current study we focused on the Copy subtest (i.e., copying a complex geometric figure), using the Meyers and Meyers (1995) scoring system which taps organization and grapho-motor control. Raw scores are converted into age-related standard scores. This scoring system has been found to have high inter-rater reliability ($r$ = .93). The ROCF taps two main stages of the handwriting production, the visual spatial arrangement parameter setting and the handwriting execution stage.

**Finger Succession (Berninger & Rutberg, 1992)**

This test measures kinesthetic ability by examining the speed of performing a sequential fine-motor task, while the hands are out of peripheral vision. In the current study we focused on the Dominant Hand subtest. Scoring is based on the time (seconds) it takes the student to complete five correctly performed cycles (i.e., touching the thumb with each of the fingers from the fifth finger-"Pinky" to the second finger-"index"). Administration time for this test is approximately 2-min. The test taps the stage of handwriting execution.

## *2.3 Procedure*

Upon receiving ethical approval from the University's review board committee, notices were posted on the University's bulletin boards on campuses and on the network, inviting students to participate in the studies. Students who expressed an

interest in the study were asked to sign an informed-consent form, which included a specific clause for adding the data to the laboratory's database and filled out the Background Questionnaire. If they met the general inclusion criteria, they were individually administered the test battery that was administered in random order. Students who took medication for ADHD were requested to perform the test battery after taking their medication. The tests were administered by experienced clinicians who specialized in learning disabilities and had over 10 years of experience. All participants could choose between receiving payment according to the University's set price for volunteering to participate in experiments or receiving credit points if their program had such a requirement (i.e., participating in an experiment as part of their curriculum).

### *2.4 Data Analysis*

Statistical analysis was performed using the Statistical significance was set at $p < .05$. Descriptive statistics were used to describe the central tendency and variability of measures. Kruskal-Wallis, followed by Mann-Whitney tests were used to compare the groups' underlying functions due to the small group size. Pearson correlations coefficients were used to examine the relationship between the underlying functions and handwriting speed and legibility, respectively. Finally, we conducted logistic regression to determine the contribution of the different underlying functions as well as students' gender to the prediction of students' handwriting status (i.e., good and poor legibility or good and poor speed). This approach uses chi-square goodness-of-fit tests to determine which set of factors is most effective in predicting outcome status.

## 3 Results

### *3.1 Sample Description and Group Distribution*

First, we examined the distribution of the different manifestations of handwriting difficulties among the students with HD ($n = 66$). Results showed 41 (62.1%) students had only poor *legibility* (legibility group; $M_{age} = 24.32$, $SD = 2.86$), of whom 22 (53.7%) were males; 16 (24.2%) had only *slow* handwriting (Speed group; $M_{age} = 23.69$, $SD = 2.82$), of whom 13 (81.3%) were males; and 9 (13.6%) students had both poor legibility and slow handwriting (Both group; $M_{age} = 23.78$, $SD = 2.28$), of whom 8 (88.9%) were males. The mean age of the TD group ($n = 44$) was 23.22 ($SD = 2.19$), of whom 16 (36.4%) were males. The groups did not significantly differ ($p > .05$) in age as well as in the number of years they studied at the university. By contrast, the three handwriting difficulties groups significantly differed in gender distribution ($\chi^2 = 6.44$, $p = .04$). The two group that included students with slow handwriting had over 80% males as opposed to 50% in the legibility group.

Descriptive data of the four groups' handwriting performance are presented in Table 1. As can be seen, the students with only slow handwriting were similar to the TD group in their legibility scores. Yet, although within the normal range, the students in the only legibility group had lower speed scores compared to the TD group. Due to the small number of participants in the Both group, and our focus on students with only legibility or only speed difficulties, subsequent analysis related to these two groups as well as the control group.

Next we compared the underlying functions among the three groups. As can be seen in the Table 2 there was a significant main effect for reading, spelling and the two working memory measures, where the two handwriting groups scored significantly below the control. Yet, the two HD groups did not significantly differ between themselves ($p > .05$).

## *3.2 Handwriting Speed*

The correlations between the underlying functions and handwriting speed, separately for students with handwriting difficulties and the TD students, are described in Table 3. As can be seen, among the TD students, only the Alphabet task was significantly correlated with handwriting speed. In contrast, among the students with HD, spelling and the Alphabet task were the only functions that were significantly correlated with handwriting speed. Table 4 presents the contribution of the different underlying functions and gender in determining the likelihood that students would have good or poor handwriting speed. In this analysis both reading and finger succession were not entered into the model because they did not have a significant correlation with either speed or legibility. In addition, we did not include the visual-spatial-motor organization test (ROCF) because, unfortunately, in the TD group, only 18 students were administered this test.

Results indicated that the goodness of fit of the four predictor variables in this logistic regression model, compared to a constant model, was statistically significant, $\chi^2(4) = 16.58$, $p = .002$, explained 45.8% (Nagelkerke $R^2$) of the variance in handwriting speed, and correctly classified 90.6% of the students as having or not having speed difficulties. Each of the two working memory variables (the Alphabet

**Table 1** Handwriting legibility and speed (in copying) of the four study groups

| | Legibility group (n = 41) | | Speed group (n = 16) | | Both[a] group (n = 9) | | TD[b] (n = 44) | |
|---|---|---|---|---|---|---|---|---|
| Handwriting measures | M | SD | M | SD | M | SD | M | SD |
| Legibility[c] | 8.21 | 5.54 | 1.60 | 1.06 | 6.67 | 3.13 | 1.26 | .95 |
| Speed[d] | 116.36 | 18.61 | 82.40 | 12.73 | 81.60 | 14.67 | 129.06 | 16.13 |

Note
[a]Both: Difficulty in legibility and speed
[b]*TD* typically developing students
[c]Percent of unreadable words (higher scores–lower legibility)
[d]Number of characters per minute

**Table 2** Comparison of study groups in underlying functions

| | Legibility group | | | | Speed group | | | | TD[a] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n | M | Md | IQR[b] | n | M | Md | IQR[b] | n | M | Md | IQR[b] | $H^{c}_{(df=2)}$ |
| *Language* | | | | | | | | | | | | | |
| Reading | 38 | −.7 | −.9 | 1.6 | 13 | −.8 | −.8 | 2.4 | 44 | .3 | .1 | 1.3 | 13.3*** |
| Spelling | 38 | −.3 | .1 | 1.5 | 13 | −2.4 | .8 | 2.6 | 44 | .6 | .8 | .7 | 10.8** |
| *WM*[d] | | | | | | | | | | | | | |
| BRIEF-A[e] | 41 | 62.7 | 64.0 | 18.0 | 16 | 64.9 | 63.5 | 19.0 | 44 | 49.1 | 49.5 | 11.0 | 31.4*** |
| Alphabet | 24 | 1.7 | 1.7 | .8 | 11 | 1.5 | 1.5 | .5 | 36 | 2.2 | 2.2 | .6 | 20.6*** |
| *PGM*[f] | | | | | | | | | | | | | |
| ROCF[g] | 30 | −4.0 | −2.8 | 3.5 | 8 | −7.1 | −2.0 | 2.8 | 18 | −2.6 | −1.6 | 4.8 | 2.8 |
| Finger succession | 37 | 8.0 | 6.5 | 3.4 | 16 | 5.7 | 6.2 | 2.9 | 40 | 5.8 | 5.9 | 2.7 | 5.5 |

Note
**$p < .01$; ***$p < .01$
[a]*TD* typically developing
[b]*IQR* interquartile range
[c]*H* Kruskal-Wallis H-test
[d]*WM* working memory
[e]*BRIEF-A* Behaviour Rating Inventory of Executive Function-Adult
[f]*PGM* perceptual-grapho-motor
[g]*ROCF* Rey Osterrieth complex figure

**Table 3** Correlations between underlying functions and handwriting measures

| | Students with HD[a] | | | TD[b] students | | |
|---|---|---|---|---|---|---|
| | | Handwriting | | | Handwriting | |
| | n | Legibility | Speed | n | Legibility | Speed |
| *Language* | | | | | | |
| Reading text | 59 | .20 | .13 | 44 | −.24 | .09 |
| Spelling | 59 | .07 | .23* | 44 | −.06 | .09 |
| *Working memory* | | | | | | |
| BRIEF-A[c] | 66 | .35** | −.15 | 44 | .06 | −.15 |
| Alphabet task | 39 | .22 | .40** | 36 | −.13 | .38* |
| *Perceptual-grapho-motor* | | | | | | |
| ROCF[d] | 43 | −.28* | −.01 | 18 | −.25 | .07 |
| Finger succession | 59 | −.01 | .05 | 40 | −.05 | .26 |

Note
*p < .05, **p < .01
[a]*HD* handwriting difficulties
[b]*TD* typically developing
[c]*BRIEF-A* Behaviour Rating Inventory of Executive Function-Adult version
[d]*ROCF* Rey Osterrieth complex figure

task and BRIEF-A) had a unique significant contribution. For each point of increase in the Alphabet task, there was a .09 increased likelihood of having good handwriting speed and for each point of increase in working memory, there was a 1.14 increased likelihood of having good handwriting speed. The unique contribution of gender approached significance ($p = .075$).

**Table 4** Logistic regression analyses for predicting handwriting status for poor legibility and for slow handwriting

| | Legibility | | | | Speed | | | |
|---|---|---|---|---|---|---|---|---|
| Variables in model | B | SE | Odds ratio | 95% CI for exp β | B | SE | Odds ratio | 95% CI for exp β |
| Spelling | −.59** | .33 | .55 | .29–1.05 | .68 | .56 | 1.97 | .66–5.86 |
| BRIEF-A[1] | .06* | .03 | 1.07 | 1.01–1.12 | .13* | .06 | 1.14 | 1.02–1.27 |
| Alphabet task | −.76 | .63 | .47 | .14–1.62 | −2.37* | 1.22 | .09 | .01–1.02 |
| Gender | .80 | .68 | 2.23 | .59–8.44 | 3.04*** | 1.71 | 20.97 | .73–597.94 |

Note
*$p < .05$, **$p = .069$, ***$p = .075$
[a]*BRIEF-A* Behaviour Rating Inventory of executive Function-Adult version

### *3.3 Handwriting Legibility*

First, we examined the relationship between the underlying functions and handwriting legibility, separately, for students with handwriting difficulties and TD students. As can be seen in Table 3, none of the functions were found to be related to handwriting legibility among the TD students. Among the students with HD, only working memory and the ROCF were found to have a significant correlation with legibility.

Table 4 presents the contribution of the different underlying functions and gender in determining the likelihood that students would have good or poor handwriting legibility. Similar to speed, four variables were included in the model. Results indicated that the goodness of fit of the four predictor variables, compared to a constant model, was statistically significant, $\chi^2(4) = 17.83$, $p = .001$, explained 33.5% (Nagelkerke $R^2$) of the variance in handwriting legibility, and correctly classified 75.4% of the students as having good or poor legibility. In this model, only the BRIEF-A working memory had a unique significant contribution. For each point of increase in this variable, there was a 1.06 increased likelihood of having good handwriting legibility. In addition, the contribution of spelling approached significance ($p = .07$).

## 4 Discussion

Handwriting research has spanned for many decades and the knowledge and understanding of handwriting production has greatly increased. Yet, there are still gaps in knowledge and areas that have not been sufficiently explored. One such area is the underlying factors related to the different measures of handwriting performance, legibility versus speed, and specifically among higher-education students. This topic was the focus of this study.

## 4.1 Manifestations of Handwriting Difficulties

First, we examined the distribution of handwriting difficulties among higher-education students. Our results confirmed that, similar to other groups of students with SLD, students with handwriting difficulties are a heterogeneous group (Deuel, 1995). Most of the students (62.2%) in our sample had only poor legibility, fewer had only slow handwriting (24.2%), and even less had both poor legibility and speed. It is not clear if this distribution is representative, because we could not find previous data on this issue, either among school-age or among higher-education studies and thus, further studies, in this area are clearly needed.

## 4.2 Handwriting Speed

Handwriting speed represents the pace of producing a handwritten text. This pace reflects the fluency of the different stages of the handwriting process. Therefore, it is reasonable to assume that both language and the perceptual-motor functions, as well as working memory, would be related to handwriting speed. In fact, various studies reported a low to medium correlation between reading and handwriting speed (Berninger et al., 2010; Bosga-Stork et al., 2016; Medwell & Wray, 2014; Salameh-Matar et al., 2018). This is not surprising, given that both require similar abilities such as phoneme-grapheme conversion (Döhla & Heim, 2016). Moreover, a handwriting copying task entails reading the text. By contrast, our results showed that reading and handwriting speed were not significantly correlated. A possible explanation of our finding is that the association between reading and handwriting speed depends on the stage of development (Abbott et al., 2010; Berninger et al., 2006; Bosga-Stork et al., 2016; Fitzgerald & Shanahan, 2000). Most of the studies focused on elementary school students, where both reading and handwriting were still developing. Perhaps, at a level of higher-education, reading no longer plays a significant role in handwriting speed.

Additionally, the results showed a low significant correlation between handwriting speed and spelling, but the latter did not have a unique significant contribution in predicting students' handwriting speed status. This finding contradicts the results of Connelly et al. (2006) who reported that college students with dyslexia did not write significantly slower than an age-related control group. However, in a later study among 9-year-old children, Sumner et al. (2013) showed that the slower handwriting among students with dyslexia was not due to slow execution, but rather because of frequent pausing. They suggested that the pausing was related to poor spelling ability. These findings coincide with our findings showing the students' spelling ability was significantly related to handwriting speed.

Spelling and handwriting, are both transcription skills that draw on orthographic processes (Abbott et al., 2010; Rodríguez & Villarroel, 2017). As described in the handwriting model (Van Galen, 1991), at the initial stages of handwriting production, after determining the word to be written, the orthographic representations of the word are retrieved from long-term memory system (Kandel et al., 2017; Van Galen, 1991), suggesting that spelling abilities play a role in the initial stages of handwriting. This premise was supported by several studies that reported a low-medium correlation between spelling skills and handwriting speed, both among elementary school students (Abbott et al., 2010; Bosga-Stork et al., 2016) and higher-education students (MacKay et al., 2019).

Working memory is an additional underlying function which plays a role in the different stages of handwriting production (Olive & Kellogg, 2002; Van Galen, 1991). Our results showed that both self-reported working memory and the Alphabet task, had a significant unique contribution in predicting students' handwriting speed status. Working memory, and specifically attention, have been found to be related to handwriting speed (Capodieci et al., 2018; Peverly et al., 2013; Tucha & Lange, 2001). Other studies have reported the important role that the ability to retrieve and write the alphabet letters from memory in their correct order and form played in handwriting speed (Abbott & Berninger, 1993; Alves et al., 2016; Salameh-Matar et al., 2018) as well as in composition quality and quantity (Alstad et al., 2015; Graham et al., 2000). Yet most of these studies focused on school-age students. It is interesting to note that even at the level of higher-education, the Alphabet task appears to be related to the speed of handwriting production.

The Alphabet task requires integrating orthographic codes (allographs) with their motor programs and execute the programs as quickly as possible. As such, it also taps motor functions, including motor speed and accuracy. Yet, our results indicated that this test was a better predictor of handwriting speed compared to a general fine-motor test, i.e., finger succession. Although finger succession was found to be related to handwriting speed in earlier ages (Berninger & Rutberg, 1992), it appears that among young adults with slow handwriting it does not play a significant role.

Finally, our results showed that the contribution of gender in predicting students' handwriting speed status approached significance. This finding is not surprising, and is concurrent with previous studies showing the females tended to write faster than males (Berninger & Fuller, 1992; Graham & Weintraub, 1996) even among young adults (Barnett et al., 2011). This was also seen in the fact that above 80% of the students with poor speed were males. A review of gender differences in speed processing (Roivainen, 2011) may shed light on this finding. The review revealed the females seem to process language-related speed tasks involving the alphabet letters and rapid naming faster, whereas males were faster on reaction time tests and finger tapping. The authors suggested that this may be partly based on the frequency of engagement of females in language related activities at school and at home. With respect to the current study, perhaps at the higher-education level, language-related processing plays a more important role in handwriting production than fine-motor, as was seen in the contribution of the Alphabet task versus finger succession.

### *4.3 Legibility*

As mentioned above, legibility reflects the neatness and readability of the handwritten product (Graham, 1986; Greifeneder et al., 2011). It is believed to be affected by various factors including, letter formation, visual-spatial organization (Graham et al., 2006; Weintraub et al., 2007), as well as spelling (Graham et al., 2011). Our results showed that although there was no significant correlation between spelling and handwriting legibility, the contribution of spelling in predicting students' handwriting legibility status approached significance. To our knowledge, this is one of the first studies that examined the contribution of spelling to handwriting legibility among higher-education students. Yet, our finding supports those of Wicki et al. (2014), who reported a low significant correlation between spelling (orthographic) skills and handwriting legibility among school-age students. These findings suggest that spelling errors may not only affect readers' perception of students' writing composition quality (Graham et al., 2011), but also its legibility, especially if operationally defined as readability.

In examining working memory, our results showed that self-report working memory, which focuses on attention and concentration had a unique and significant contribution in predicting students' legibility status. This result supports previous studies showing the association between attention and handwriting legibility (Adi-Japha et al., 2007; Tucha & Lange, 2001). On the other hand, we did not find a significant correlation between the Alphabet task and handwriting legibility. This finding is concurrent with results of an earlier study among fifth grade students (Weintraub & Graham, 2000). Perhaps this is because the scoring of the Alphabet task focuses mostly on the handwriting process (i.e., speed of writing letters) rather than on the product (legibility; i.e., the concise production of the letters' motor programs).

Finally, similar to other studies (Kaiser et al., 2009; Weintraub & Graham, 2000) our results indicated that visual-spatial grapho-motor abilities (ROCF) were significantly related to handwriting legibility. This finding supports the additional stage suggested by of Graham et al. (2006) to Van Galen's (1991) handwriting model, the visual-spatial arrangement parameter setting. Interestingly, we did not find a significant relationship between finger succession (which measures fine-motor sequencing and kinesthetic awareness necessary for handwriting production), and handwriting legibility (which is influenced by concise letter formation; Graham et al., 2006; Weintraub et al., 2007). As stated above, this task may be too easy for higher-education students, showing small variability. Hence, perhaps more ecological tests, examining fine-motor functions, are more suitable for this age-group (Tal-Saban & Weintraub, 2019).

There is a common premise that the handwriting legibility of males is lower than that of females. The premise is based on various studies' findings (Barnett et al., 2018; Berninger et al., 2008; Weintraub & Graham, 2000). In most of these studies, the legibility rating reflected the readers' perception of legibility, which is mostly

affected by neatness and uniformity of spatial components (spacing, slant or letter size). However, a recent study (Margalit-Budznak, 2017) among middle school students, has shown that the percentage of boys identified as having poor legibility, based on the perceived legibility, was almost double compared to the percentage of boys identified when legibility was operationally defined as readability, as was the case in the current study. Therefore, perhaps gender does not play such an important role in predicting students' legibility status when legibility is defined as readability, which is less affected by the neatness of the handwritten product, but rather reflects the readers' ability to decipher the text, while relying on the text's context.

## *4.4 Study Limitations and Future Studies*

As far as we know, this is one of the first studies examining the distribution of handwriting difficulties (poor legibility or slow handwriting), and comparing the underlying functions related to each of these groups among higher-education students. One of the limitations of this study is that this sample was drawn from a data-base of several studies. Although the methodology of the studies was identical, including students' sampling and inclusion criteria, and most of the test battery, we did not administer all tests in all studies. Therefore, in some instances (e.g., ROCF), there was missing data which limited our analyses. Moreover, this study is based on students who volunteered for the study, and so, it is subject to referral bias. For example, perhaps being more visible, students with poor legibility were more aware of their difficulties compared to students with slow handwriting, and therefore they were more prone to volunteer to participate. Thus, future studies should be replicated with a larger and more heterogeneous sample. Furthermore, future studies should examine working memory or attention aspects not only as a self-report, but with more canonical tasks, to better understand how these functions predict handwriting speed or legibility.

## *4.5 Summary and Implications*

The results of this study support the premise that students with handwriting difficulties are a heterogeneous group. Some students may have more difficulties with the handwriting process (speed), whereas others may have problems with their handwriting product (legibility), or both. This finding suggests that both for research and educational purposes, it is important to more specifically define poor handwriting or students with handwriting difficulties. Such a distinction will assist in better understanding the different handwriting difficulties and will increase external validity of the studies.

Furthermore, the results of this study support our hypothesis that these types of difficulties may be associated with different underlying functions. From a theoretical perspective, these findings suggest that handwriting speed and legibility not only reflect different aspects of handwriting production (i.e., process and product), but also draw on different underlying processes, where working memory is required at all stages and aspects. These results further support the premise that difficulties at any stage of the handwriting production (Graham et al., 2006; Van Galen, 1991) may result in different types of handwriting problems. Therefore, in assessing handwriting performance it is important to establish at what stage do the students encounter difficulties, and what are the possible deficits in the underlying functions. This distinction may better support and enhance the efficacy of handwriting instruction or clinical intervention programs.

## References

Abbott, R. D., & Berninger, V. W. (1993). Structural equation modeling of relationships among developmental skills and writing skills in primary- and intermediate-grade writers. *Journal of Educational Psychology, 85*(3), 478–508. https://doi.org/10.1037/0022-0663.85.3.478

Abbott, R. D., Berninger, V. W., & Fayol, M. (2010). Longitudinal relationships of levels of language in writing and between writing and reading in grades 1 to 7. *Journal of Educational Psychology, 102*(2), 281. https://doi.org/10.1037/a0019318

Adi-Japha, E., Landau, Y. E., Frenkel, L., Teicher, M., Gross-Tsur, V., & Shalev, R. S. (2007). ADHD and dysgraphia: Underlying mechanisms. *Cortex, 43*(6), 700–709. https://doi.org/10.1016/s0010-9452(08)70499-4

Alstad, Z., Sanders, E., Abbott, R. D., Barnett, A. L., Henderson, S. E., Connelly, V., & Berninger, V. W. (2015). Modes of alphabet letter production during middle childhood and adolescence: Interrelationships with each other and other writing skills. *Journal of Writing Research, 6*(3), 199–231. https://doi.org/10.17239/jowr-2015.06.03.1

Alves, R. A., Limpo, T., Fidalgo, R., Carvalhais, L., Pereira, L. Á., & Castro, S. L. (2016). The impact of promoting transcription on early text production: Effects on bursts and pauses, levels of written language, and writing performance. *Journal of Educational Psychology, 108*(5), 665–679. https://doi.org/10.1037/edu0000089

Barnett, A. L., Henderson, S. E., Scheib, B., & Schulz, J. (2011). Handwriting difficulties and their assessment in young adults with DCD: Extension of the DASH for 17- to 25-year-olds. *Journal of Adult Development, 18*(3), 114–121. https://doi.org/10.1007/s10804-011-9121-3

Barnett, A. L., Prunty, M., & Rosenblum, S. (2018). Development of the Handwriting Legibility Scale (HLS): A preliminary examination of reliability and validity. *Research in Developmental Disabilities, 72*, 240–247. https://doi.org/10.1016/j.ridd.2017.11.013

Ben-Simon, A., & Inbar-Weiss, N. (2012). *MATAL test battery for the diagnosis of learning disabilities: User guide*. National Institute for Testing and Evaluation and the Council for Higher-Education (Hebrew).

Berninger, V. W. (2000). Development of language by hand and its connections with language by ear, mouth, and eye. *Topics in Language Disorders, 20*(4), 65–84. https://doi.org/10.1097/00011363-200020040-00007

Berninger, V. W., & Fuller, F. (1992). Gender differences in orthographic, verbal, and compositional fluency: Implications for assessing writing disabilities in primary grade children. *Journal of School Psychology, 30*(4), 363–382. https://doi.org/10.1016/0022-4405(92)90004-O

Berninger, V. W., & Rutberg, J. (1992). Relationship of finger function to beginning writing: Application to diagnosis of writing disabilities. *Developmental Medicine & Child Neurology, 34*(3), 198–215. https://doi.org/10.1111/j.1469-8749.1992.tb14993.x

Berninger, V., Yates, C., Cartwright, A., Rutberg, J., Remy, E., & Abbott, R. (1992). Lower-level developmental skills in beginning writing. *Reading and Writing: An Interdisciplinary Journal, 4*(3), 257–280. https://doi.org/10.1007/BF01027151

Berninger, V. W., Vaughan, K. B., Abbott, R. D., Abbott, S. P., Rogan, L. W., Brooks, A., Reed, E., & Graham, S. (1997). Treatment of handwriting problems in beginning writers: Transfer from handwriting to composition. *Journal of Educational Psychology, 89*(4), 652. https://doi.org/10.1037/0022-0663.89.4.652

Berninger, V. W., Abbott, R. D., Abbott, S. P., Graham, S., & Richards, T. (2002). Writing and reading: Connections between language by hand and language by eye. *Journal of Learning Disabilities, 35*(1), 39–56. https://doi.org/10.1177/002221940203500104

Berninger, V. W., Abbott, R. D., Jones, J., Wolf, B. J., Gould, L., Anderson-Youngstrom, M., Shimada, S., & Apel, K. (2006). Early development of language by hand: Composing, reading, listening, and speaking connections; three letter-writing modes; and fast mapping in spelling. *Developmental Neuropsychology, 29*(1), 61–92. https://doi.org/10.1207/s15326942dn2901_5

Berninger, V. W., Nielsen, K. H., Abbott, R. D., Wijsman, E., & Raskind, W. (2008). Gender differences in severity of writing and reading disabilities. *Journal of School Psychology, 46*(2), 151–172. https://doi.org/10.1016/j.jsp.2007.02.007

Berninger, V. W., Abbott, R. D., Swanson, H. L., Lovitt, D., Trivedi, P., Lin, S.-J. C., Gould, L., Youngstrom, M., Shimada, S., & Amtmann, D. (2010). Relationship of word and sentence-level working memory to reading and writing in second, fourth, and sixth grade. *Language, Speech and Hearing Services in Schools, 41*(2), 179–193. https://doi.org/10.1044/0161-1461(2009/08-0002)

Berninger, V. W., Nagy, W., Tanimoto, S., Thompson, R., & Abbott, R. D. (2015a). Computer instruction in handwriting, spelling, and composing for students with specific learning disabilities in grades 4-9. *Computers & Education, 81*, 154–168. https://doi.org/10.1016/j.compedu.2014.10.005

Berninger, V. W., Richards, T. L., & Abbott, R. D. (2015b). Differential diagnosis of dysgraphia, dyslexia, and OWL LD: Behavioral and neuroimaging evidence. *Reading and Writing: An Interdisciplinary Journal, 28*(8), 1119–1153. https://doi.org/10.1007/s11145-015-9565-0

Bosga-Stork, I. M., Bosga, J., Ellis, J. L., & Meulenbroek, R. G. (2016). Developing interactions between language and motor skills in the first three years of formal handwriting education. *British Journal of Education Society and Behavioural Science, 12*(1), 1–13. https://doi.org/10.9734/BJESBS/2016/20703

Brooks, A. D., Berninger, V. W., & Abbott, R. D. (2011). Letter naming and letter writing reversals in children with dyslexia: Momentary inefficiency in the phonological and orthographic loops of working memory. *Developmental Neuropsychology, 36*(7), 847–868. https://doi.org/10.1080/87565641.2011.606401

Capodieci, A., Lachina, S., & Cornoldi, C. (2018). Handwriting difficulties in children with attention deficit hyperactivity disorder (ADHD). *Research in Developmental Disabilities, 74*, 41–49. https://doi.org/10.1016/j.ridd.2018.01.003

Connelly, V., Campbell, S., MacLean, M., & Barnes, J. (2006). Contribution of lower order skills to the written composition of college students with and without dyslexia. *Developmental Neuropsychology, 29*, 175–196. https://doi.org/10.1207/s15326942dn2901_9

Cousins, M., & Smyth, M. M. (2003). Developmental coordination impairments in adulthood. *Human Movement Science, 22*(4–5), 433–459. https://doi.org/10.1016/j.humov.2003.09.003

Deuel, R. K. (1995). Developmental dysgraphia and motor skills disorders. *Journal of Child Neurology, 10*(1_suppl), S6–S8. https://doi.org/10.1177/08830738950100S103

Döhla, D., & Heim, S. (2016). Developmental dyslexia and dysgraphia: What can we learn from the one about the other? *Frontiers in Psychology, 6*, 2045. https://doi.org/10.3389/fpsyg.2015.02045

Duiser, I. H., van der Kamp, J., Ledebt, A., & Savelsbergh, G. J. (2013). Relationship between the quality of children's handwriting and the Beery Buktenica developmental test of visuomotor integration after one year of writing tuition. *Australian Occupational Therapy Journal, 61*(2), 76–82. https://doi.org/10.1111/1440-1630.12064

Engel-Yeger, B., Nagauker-Yanuv, L., & Rosenblum, S. (2009). Handwriting performance, self-reports, and perceived self-efficacy among children with dysgraphia. *American Journal of Occupational Therapy, 63*(2), 182–192. https://doi.org/10.5014/ajot.63.2.182

Feng, L., Lindner, A., Ji, X. R., & Joshi, R. M. (2019). The roles of handwriting and keyboarding in writing: A meta-analytic review. *Reading and Writing: An Interdisciplinary Journal*, 1–31. https://doi.org/10.1007/s11145-017-9749-x

Fitzgerald, J., & Shanahan, T. (2000). Reading and writing relations and their development. *Educational Psychologist, 35*, 39–50. https://doi.org/10.1207/S15326985EP3501_5

Graham, S. (1986). A review of handwriting scales and factors that contribute to variability in handwriting scores. *Journal of School Psychology, 24*(1), 63–71. https://doi.org/10.1016/0022-4405(86)90043-9

Graham, S., & Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980 to 1994. *Educational Psychology Review, 8*(1), 7–87. https://link.springer.com/content/pdf/10.1007/BF01761831.pdf

Graham, S., Harris, K. R., & Fink, B. (2000). Is handwriting causally related to learning to write? Treatment of handwriting problems in beginning writers. *Journal of Educational Psychology, 92*(4), 620. https://doi.org/10.1057//0022-0663.92.4.620

Graham, S., Struck, M., Santoro, J., & Berninger, V. W. (2006). Dimensions of good and poor handwriting legibility in first and second graders: Motor programs, visual-spatial arrangement, and letter formation parameter setting. *Developmental Neuropsychology, 29*(1), 43–60. https://doi.org/10.1207/s15326942dn2901_4

Graham, S., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children, 44*(4), 1–12. https://doi.org/10.17161/fec.v44i4.6687

Graham, S., Fishman, E. J., Reid, R., & Hebert, M. (2016). Writing characteristics of students with attention deficit hyperactive disorder: A meta-analysis. *Learning Disabilities Research and Practice, 31*(2), 75–89. https://doi.org/10.1111/ldrp.12099

Graham, S., Aitken, A. A., Hebert, M., Camping, A., Santangelo, T., Harris, K. R., Eustice, K., Sweet, J. D., & Ng, C. (2020). Do children with reading difficulties experience writing difficulties? A meta-analysis. *Journal of Educational Psychology*. https://doi.org/10.1037/edu0000643

Greifeneder, R., Zelt, S., Seele, T., Bottenberg, K., & Alt, A. (2011). Towards a better understanding of the legibility bias in performance assessments: The case of gender-based inferences. *British Journal of Educational Psychology, 82*(3), 361–374. https://doi.org/10.1111/j.2044-8279.2011.02029.x

Kaiser, M. L., Albaret, J. M., & Doudin, P. A. (2009). Relationship between visual-motor integration, eye-hand coordination, and quality of handwriting. *Journal of Occupational Therapy, Schools, & Early Intervention, 2*(2), 87–95. https://doi.org/10.1080/19411240903146228

Kandel, S., & Spinelli, E. (2010). Processing complex graphemes in handwriting production. *Memory & Cognition, 38*(6), 762–770. https://doi.org/10.3758/MC.38.6.762

Kandel, S., Lassus-Sangosse, D., Grosjacques, G., & Perret, C. (2017). The impact of developmental dyslexia and dysgraphia on movement production during word writing. *Cognitive Neuropsychology, 34*(3–4), 219–251. https://doi.org/10.1080/02643294.2017.1389706

Katusic, S. K., Colligan, R. C., Weaver, A. L., & Barbaresi, W. J. (2009). The forgotten learning disability: Epidemiology of written-language disorder in a population-based birth cohort (1976–1982), Rochester, Minnesota. *Pediatrics, 123*(5), 1306–1313. https://doi.org/10.1542/peds.2008-2098

Klein, S., Guiltner, V., Sollereder, P., & Cui, Y. (2011). Relationships between fine-motor, visual-motor, and visual perception scores and handwriting legibility and speed. *Physical & Occupational Therapy in Pediatrics, 31*(1), 103–114. https://doi.org/10.3109/01942638.2010.541753

Limpo, T., & Alves, R. A. (2013). Modeling writing development: Contribution of transcription and self-regulation to Portuguese students' text generation quality. *Journal of Educational Psychology, 105*(2), 401. https://psycnet.apa.org/doi/10.1037/a0031391

MacKay, E. J., Larcohe, A., Parrila, R., & Deacon, S. H. (2019). A beginning exploration of text generation abilities in university students with a history of reading difficulties. *Dyslexia, 25*(2), 207–218. https://doi.org/10.1002/dys.1610

Margalit-Budznak, N. (2017). *Global legibility versus readability scoring methods: The impact on identification of middle-school students with poor handwriting.* Master's Thesis, The Hebrew University of Jerusalem.Unpublished

McCloskey, M., & Rapp, B. (2017). Developmental dysgraphia: An overview and framework for research. *Cognitive Neuropsychology, 34*(3–4), 65–82. https://doi.org/10.1080/02643294.2017.1369016

Medwell, J., & Wray, D. (2014). Handwriting automaticity: the search for performance thresholds. *Language and Education, 28*(1), 34–51. https://doi.org/10.1080/09500782.2013.763819

Meyers, J. E., & Meyers, K. R. (1995). Rey complex figure test under four different administration procedures. In *Professional Manual*. Psychological Assessment Resources.

Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high-and low-level production processes in written composition. *Memory & Cognition, 30*(4), 594–600. https://doi.org/10.3758/BF03194960

Palmis, S., Danna, J., Velay, J. L., & Longcamp, M. (2017). Motor control of handwriting in the developing brain: A review. *Cognitive Neuropsychology, 34*(3–4), 187–204. https://doi.org/10.1080/02643294.2017.1367654

Peverly, S. T., Vekaria, P. C., Reddington, L. A., Sumowski, J. F., Johnson, K. R., & Ramsay, C. M. (2013). The relationship of handwriting speed, working memory, language comprehension and outlines to lecture note-taking and test-taking among college students. *Applied Cognitive Psychology, 27*(1), 115–126. https://doi.org/10.1002/acp.2881

Prunty, M., Barnett, A. L., Wilmut, K., & Plumb, M. (2016). Visual perceptual and handwriting skills in children with developmental coordination disorder. *Human Movement Science, 49*, 54–65. https://doi.org/10.1016/j.humov.2016.06.003

Rodríguez, C., & Villarroel, R. (2017). Predicting handwriting difficulties through spelling processes. *Journal of Learning Disabilities, 50*(5), 504–510. https://doi.org/10.1177/0022219416633863

Roivainen, E. (2011). Gender differences in processing speed: A review of recent research. *Learning and Individual Differences, 21*(2), 145–149. https://doi.org/10.1016/j.lindif.2010.11.021

Rosenberg-Adler, T., & Weintraub, N. (2020). Reliability and preliminary outcomes of a protocol for selection of test accommodations for higher-education students with dysgraphia: A pilot study. *American Journal of Occupational Therapy, 74*, 7404205080. https://doi.org/10.5014/ajot.2020.039826R

Rosenblum, S., Weiss, L., & Parush, P. (2003). Product and process evaluation of handwriting difficulties. *Educational Psychology Review, 15*(1), 41–81. https://doi.org/10.1023/A:1021371425220

Roth, R. M., Gioia, G. A., & Isquith, P. K. (2005). *BRIEF-A: Behavior Rating Inventory of Executive Function—Adult version*. Psychological Assessment Resources.

Salameh-Matar, A., Basal, N., & Weintraub, N. (2018). Relationship between body functions and Arabic handwriting performance at different acquisition stages. *Canadian Journal of Occupational Therapy, 85*(5), 418–427. https://doi.org/10.1177/0008417419826114

Shaywitz, S. E., & Shaywitz, B. A. (2005). Dyslexia (specific reading disability). *Biological Psychiatry, 57*(11), 1301–1309. https://doi.org/10.1016/j.biopsych.2005.01.043

Shin, M. S., Park, S. Y., Park, S. R., Seol, S. H., & Kwon, J. S. (2006). Clinical and empirical applications of the Rey–Osterrieth complex figure test. *Nature Protocols, 1*(2), 892. https://doi.org/10.1038/nprot.2006.115

Smits-Engelsman, B. C. M., Niemeijer, A. S., & Van Galen, G. P. (2001). Fine motor deficiencies in children diagnosed as DCD based on poor grapho-motor ability. *Human Movement Science, 20*, 161–182. https://doi.org/10.1016/S0167-9457(01)00033-1

Sumner, E., Connelly, V., & Barnett, A. L. (2013). Children with dyslexia are slow writers because they pause more often and not because they are slow at handwriting execution. *Reading and Writing: An Interdisciplinary Journal, 26*(6), 991–1008. https://doi.org/10.1007/s11145-012-9403-6

Tal-Saban, M., & Weintraub, N. (2019). Motor functions of higher-education students with dysgraphia. *Research in Developmental Disabilities, 94*, 103479. https://doi.org/10.1016/j.ridd.2019.103479

Tucha, O., & Lange, K. W. (2001). Effects of methylphenidate on kinematic aspects of handwriting in hyperactive boys. *Journal of Abnormal Child Psychology, 29*(4), 351–356. https://doi.org/10.1023/A:1010366014095

Van Galen, G. P. (1991). Handwriting: Issues for a psychomotor theory. *Human Movement Science, 10*(2–3), 165–191. https://doi.org/10.1016/0167-9457(91)90003-G

Weintraub, N., & Graham, S. (2000). The contribution of gender, orthographic, finger function, and visual-motor processes to the prediction of handwriting status. *OTJR: Occupation, Participation and Health, 20*(2), 121–140. http://otj.sagepub.com/content/20/2/121.short

Weintraub, N., Drory-Asayag, A., Dekel, R., Jokobovits, H., & Parush, S. (2007). Developmental trends in handwriting performance among middle school children. *OTJR: Occupation, Participation and Health, 27*(3), 104–113. https://doi.org/10.1177/153944920702700304

Weintraub, N., Gilmour-Grill, N., & Weiss, P. L. T. (2010). Relationship between handwriting and keyboarding performance among fast and slow adult keyboarders. *American Journal of Occupational Therapy, 64*(1), 123–132.

Weintraub, N., Israeli-Ovadia, A., Shoval, H., & Traub-Bar-Ilan, R. (2012). *The handwriting performance of post-secondary students evaluation (HaPPS)*. The School of Occupational Therapy of Hadassah and the Hebrew University.

Wicki, W., Lichtsteiner, S. H., Geiger, A. S., & Müller, M. (2014). Handwriting fluency in children: Impact and correlates. *Swiss Journal of Psychology, 73*(2), 87–96. https://doi.org/10.1024/1421-0185/a000127

Yusilles, N. (2012). *Validity and reliability of the Handwriting Performance of Post-Secondary Students Evaluation (HaPPS)* (Master's thesis). School of Occupational Therapy, Hebrew University of Jerusalem. Unpublished.

# Writing Interventions Using SRSD for Secondary Students with and At-Risk for Learning Disabilities: A Review of Empirical Research


**Amber B. Ray**

**Abstract** Many secondary students find writing challenging and teachers need research-based interventions to help students become successful writers. Seventeen studies using self-regulated strategy development (SRSD) to teach writing were evaluated using the Council for Exceptional Children's *Standards for Evidence-Based Practices in Special Education*. All studies had participants that were in middle and high school and were identified as having or at-risk for a learning disability. The studies were examined by looking at types of writing, writing strategies used, and writing outcomes measured. The studies showed that there are effective strategies for argumentative, narrative, expository, summary, informative, compare and contrast, and paragraph writing. All writing strategies were represented by mnemonics devices. The most commonly used writing outcomes measures were quality, elements, and length. Several implications for practice were identified, most notably that secondary teachers should consider using SRSD writing instruction to enhance the writing skills of students with and at-risk for learning disabilities. Recommendations for future writing instruction research and limitations are discussed.

**Keywords** Writing instruction · Self-regulated strategy development · Learning disabilities · Secondary education

A critical challenge today is preparing more students, including those with and at-risk for learning disabilities (LD), to be ready to write at the college level. On the most recent National Assessment of Educational Progress (NAEP) exam, a staggering 62% of 12th grade students with disabilities scored at the below basic level for writing (U.S. Department of Education, 2011) which indicates minimal to no knowledge or skills in writing. The remaining 38% of 12th grade students with disabilities scored at the basic level for writing, with only partial mastery of fundamental knowledge and skills. Writing is a very difficult task that students are asked to do

A. B. Ray (✉)
University of Illinois at Urbana-Champaign, Champaign, IL, USA
e-mail: amberray@illinois.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_13

every day. At the secondary level, writing is frequently used to facilitate and evaluate students' learning (Kiuhara et al., 2009). To demonstrate knowledge, students are often asked to write paragraphs, short responses, or complete written worksheets. Further, writing is a useful tool for learning (Bangert-Drowns et al., 2004) because it requires making decisions about identifying and organizing central ideas and key details, synthesizing information, and expressing it clearly. Writing to learn provides students the opportunity to reflect on their learning as it creates a concrete record of material students view as important enough to record. Writing helps students internalize information, as putting information into one's own words can make it more memorable. When writing, students must (a) conduct multiple processes; (b) plan, write, and revise papers using strategies; and (c) use self-regulation (Reid et al., 2013). Students with and at-risk for LD need tools to help them become proficient at writing.

When working on academic tasks, such as writing, students with LD typically develop and employ fewer strategies (Stone & Conca, 1993). Difficulties with writing hinder students with LD in entering and succeeding in college (Clinedinst et al., 2015). Given the importance of writing in school and beyond, it is essential to help middle and high school students with and at-risk for LD become better writers. Secondary teachers have the challenging undertaking of preparing their college bound students to be successful at an institution of higher education.

While no one form of instruction will meet the needs of all students, the meta-analysis in the *Writing Next* report found 11 key elements of effective adolescent writing instruction including writing strategies, summarization, collaborative writing, specific product goals, word processing, sentence-combining, prewriting, inquiry activities, process writing approach, study of models, and writing for content learning (Graham & Perin, 2007). The first element described is teaching students strategies for writing. This meta-analysis showed an effect size of 0.82 when students were taught a strategy for writing (Graham & Perin, 2007). Such a dramatic effect demonstrates the importance of using strategy instruction for writing.

Strategies are important in helping students approach and complete schoolwork to the best of their abilities. Strategy instruction can be used to teach students the strategies they need to become successful writers. Strategy instruction is when teachers (a) target specific strategies for students to learn, (b) provide students with the information about the strategy's use and significance, and (c) develop students' self-regulation skills as necessary to deploy, monitor, maintain, and generalize strategies effectively (Sawyer et al., 1992). There are three important aspects of effective strategy instruction, the first being that the strategy is facilitative and essential (Alexander et al., 1998). A strategy should serve as a tool to help students perform academically. Second, students need to be willful and effortful, actively choosing to use the strategy. Third, strategy procedures must have specific tasks associated with them. Students need to be able to recognize and utilize the correct strategy for the task presented to them. Strategies are advantageous tools for students to use in school. Students with and at-risk for LD need to be explicitly taught strategies which can help them improve their academic performance.

## 1 Self-Regulated Strategy Development

Self-Regulated Strategy Development (SRSD) is one model of strategies instruction that has been deemed an evidence-based practice and is effective in improving students' writing quality with an effect size of 1.59 (Graham & Harris, 2018). In the mid-1980s the instructional method of SRSD was developed by Karen R. Harris for elementary school students (Graham & Harris, 2018) and revised over the years to make it more effective. Soon after, SRSD began to be implemented with middle school students (De La Paz & Graham, 1997) and by 2005, studies were conducted using SRSD instruction with high school students (e.g., Chalk et al., 2005).

SRSD is an instructional approach with six stages including (a) developing and activating background knowledge, (b) discussing the strategy, (c) modeling the strategy, (d) memorizing the strategy, (e) supporting the strategy, and (f) independent performance (Reid et al., 2013). The first stage of developing and activating background knowledge is needed to help students learn any pre-skills necessary for the upcoming writing task. Discussing the strategy is the second stage. This is when students and the teacher talk about the strategy together. The teacher works to establish the benefits of using the strategy and how it can be utilized. The third stage is the teacher modeling using the writing strategy and self-regulating the writing process through goal setting, self-instructions, self-monitoring, and self-reinforcement. The students get to see the approach being used when composing an essay and hear the self-instructions the teacher uses to help with the writing process.

The fourth stage in SRSD is memorizing the strategy. This often is introduced and practiced starting in the first instructional stage, but it is important that students have memorized the strategy by this stage in order to achieve independence in utilizing the writing strategy. The fifth stage is supporting the strategy. This is when the teacher scaffolds the students' use of the strategy and provides a gradual release of responsibility from the teacher to the student. The final instructional stage is when students achieve independent performance. Students achieve this when they can self-regulate the writing process and utilize the writing strategy with minimal to no teacher assistance. SRSD is an effective instructional approach because it helps students develop the use of the writing strategy, metacognition, and metastrategy information (Harris & Graham, 1996). This allows students to master the strategies and achieve independence. Using SRSD has been proven highly effective for students with LD (Reid et al., 2013).

The implementation of effective instruction for writing at the middle and high school level is crucial to prepare students to meet college level expectations. The purpose of this literature review was to critically examine writing intervention research that uses SRSD with middle and high school students with and at-risk for LD in order to illuminate effective strategies and identify areas where further research is needed. This review was guided by the following research questions.

Research Question 1: What types of writing are taught using SRSD for middle and high school students with and at-risk for LD?
Research Question 2: What strategies are utilized when teaching writing to middle and high school students with and at-risk for LD using SRSD?
Research Question 3: What writing outcomes are measured when teaching writing to middle and high school students with and at-risk for LD when using SRSD?

Overall, the use of SRSD was examined in this review by looking at the different types of writing taught, strategies utilized, and writing outcomes measured when teaching writing to middle and high school students with and at-risk for LD.

## 2 Method

To answer the research questions, a systematic review of existing literature was conducted. Five inclusion criteria were developed to identify relevant articles. First, the article had to be published in a peer-reviewed journal written in English. Second, the study had to include students in sixth through twelfth grade with or at-risk for LD. The focus of this article was secondary students. Studies were excluded when students were in a sixth-grade classroom that was part of a K – 6 elementary school. Additionally, any studies that combined students from elementary and middle schools were excluded unless disaggregated data was provided. Furthermore, the studies had to include 50% or more of students with or at-risk for LD or provide disaggregated data. Students were considered to be with or at-risk for LD based on individual study criteria. Common phrases within studies to identify students as at-risk for LD included low or poor performing, struggling learners, below average, or in need of supplemental instruction. When disaggregated data was provided, only the data from the participants that met the inclusion criteria was utilized in the review. Third, the intervention needed to teach writing to students using the SRSD instructional model. Fourth, the study needed to include a dependent measure of writing performance. Fifth, the study needed to use an experimental, quasi-experimental, or single-case research design.

### *2.1 Search Procedures*

A computer database search of Education Full Text, Academic Search Ultimate, ERIC, PsycINFO, JSTOR, and EBSCO Professional Development Collection was conducted in September 2020 using the Boolean phrase: (“writ*”) AND (“self-regulat*” OR “SRSD” OR “strategy instruction” OR “self-regulated strategy development”) AND (“child” OR “adolescent” OR “student”) for all previous dates. To ensure that recent articles were not missed, a hand search was conducted of the

following journals for the years 2019 and 2020: *Exceptional Children, Journal of Learning Disabilities, Journal of Literacy Research, Learning Disability Quarterly, Reading and Writing: An Interdisciplinary Journal*, and *Journal of Writing Research*. These journals were selected as they frequently publish research on academic interventions and instruction. Furthermore, an ancestral search of previous literature reviews and meta-analyses was conducted to locate additional studies that met inclusion criteria that had not already been identified (Gillespie & Graham, 2014; Graham et al., 2013, 2016, 2020; Graham & Harris, 2018; Graham & Perin, 2007; Santangelo et al., 2016).

### *2.2 Coding*

All articles were coded based on the recommendations from the *Council for Exceptional Children Standards for Evidence-Based Practices in Special Education* (Cook et al., 2014). The nine overarching categories included participants, setting, intervention agent (teacher, researcher, graduate student, etc.), independent variable (intervention), length of intervention, treatment fidelity, dependent variable (measures), and research design (methodology). The coding sheet included 18 criteria within the nine categories to allow for a systematic way to thoroughly analyze each study.

## 3 Results

The search process previously described resulted in the identification of 1512 articles pertaining to SRSD in writing. The titles and abstracts were then reviewed, and 51 articles were identified. These articles were then read and screened for adherence to the inclusion criteria. From these, 17 studies met the inclusion criteria and were coded (see Table 1). The six reasons studies did not meet the inclusion criteria and were excluded were that the independent variable was not SRSD but rather another form of strategy instruction for writing (n = 10), the participants were typically achieving students (n = 13), participants attended an elementary school that went up through 6th grade or participants were students from elementary and middle schools without disaggregated data (n = 7), less than 50% of the participants were with or at-risk for LD (n = 2), the study focused on teacher outcome measures (n = 1), or the study focused on general learning strategies and not specific writing outcome measures (n = 1). All articles were double coded by the author and a trained graduate student. There were 306 items on which there could be agreement or disagreement (i.e., 17 studies × 18 criteria per study). Agreement was reached on 90% for the initial coding. All disagreements were

**Table 1** Overview of included studies using SRSD to teach writing

| Authors, publication date | Location, sample *N* with or at-risk for LD | Grade | Type of student | Strategies | Type of writing | Design | Writing outcomes and results for post-instruction |
|---|---|---|---|---|---|---|---|
| Asaro-Saddler et al. (2018) | New York, US, *N* = 13 | 9th–12th | LD | WINDOW | Summary | Q | Improved summary writing rubric scores and increased length |
| Benedek-Wood et al. (2014) | Mid-Atlantic, US, *N* = 7 | 6th | LD | POW + TIDE$^2$ | Informative | SCRD | Improved organizational quality, transition words, length, and knowledge units |
| Burke et al. (2017) | Midwest, US, *N* = 6 | 7th | At-risk | PLAN and WRITE | Expository | SCRD | Improved quality and planning, no change in conventions |
| Chalk et al. (2005) | Southeast, US, *N* = 15 | 10th | LD | DARE | Argumentative | SCRD | Quality 100% PND; number of words 100% PND |
| De La Paz (1999) | Southeast, US, *N* = 12 | 7th and 8th | LD and at-risk | PLAN and WRITE | Expository | SCRD | Improved length, elements, and quality; all students planned at post-instruction |
| Eissa (2009) | Egypt, *N* = 34 | 9th | LD | DARE | Argumentative | Q | Quality effect size 5.06; improved quality and length |
| Foxworth et al. (2017) | Northeast, US, *N* = 3 | 7th and 8th | LD | POW + STACS | Narrative | SCRD | Improved quality, story grammar, and elements |
| Geres-Smith et al. (2019) | Vancouver, British Columbia, *N* = 8 | 6th and 7th | At-risk | POW + TREE | Argumentative | E | Improved elements, analytic writing rubric scores, correct word sequence, percent correct word sequence, and writing duration; no improvement on length |
| Hoover et al. (2012) | Eastern US, *N* = 4 | 11th–12th | LD | POW + TREE | Argumentative | SCRD | Elements 55.83% PND; number of words 21.67% PND |
| Konrad and Test (2007) | Southeast, US, *N* = 7 | 7th and 8th | LD | GO 4 IT . . . NOW! | Paragraph | SCRD | Improved quality, no improvement on length |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| MacArthur and Philippakos (2010) | US, $N = 3$ | 6th–8th | LD | TAP and IBC | Compare/ contrast | SCRD | Elements text structure 100% PND; improved elements, quality, and length |
| Mason et al. (2013) | Northeast, US, $N = 33$ | 7th and 8th | LD & at-risk | POW + TREE | Argumentative | Q | Improved elements, overall quality, and organizational quality; no improvement on length |
| Ray and Graham (2020) | Southwest & Midwest, US, $N = 7$ | 10th–12th | LD & at-risk | HIT SONGS[3] | Argumentative | E | Improved planning, quality, elements, and transitions |
| Ray et al. (2019) | Southwest, US, $N = 4$ | 10th | LD & at-risk | HIT SONGS[3] | Argumentative | SCRD | Plans PND 92%, quality PND 100%, elements PND 100%, TWW PND 75%, transitions PND 75% |
| Saddler et al. (2019) | New York, US, $N = 1$ | 6th | LD | WINDOW | Summary | SCRD | Elements PND 100% |
| Straub and Vasquez (2015) | Southeast, US, $N = 4$ | 6th, 8th–10th | LD | POW + TREE | Argumentative | SCRD | Elements PND 66%, improved quality and correct incorrect word sequence |
| Sundeen (2012) | Suburban, US, $N = 11$ | 11th | LD | MIND | Narrative | SCRD | Improved quality |

*Note. PND* percent nonoverlapping data. 90% is considered a large effect, 70–90% is considered a medium effect, and 50–70% is considered a small effect. *LD* learning disability, *SCRD* single-case research design, *E* experimental, *Q* quasi-experimental

discussed until a consensus was reached for 100% of the items. The findings from the 17 studies are synthesized with respect to participants, setting, intervention agent, independent variable, length of intervention, treatment fidelity, dependent variable, and methodology.

## *3.1 Participants*

Across all the studies included, there were a total of 389 participants. The number of treatment students with and at-risk for LD were 123 and 49, respectively. While only studies that included 50% or more of students with or at-risk for LD or provided disaggregated data were included, the characteristics of the other students within the studies included Other Health Impairment, Autism Spectrum Disorder, Emotional Behavior Disorder, Educable Mental Disability, Attention Deficit Hyperactivity Disorder, Speech and Language Impairment, Traumatic Brain Injury, and typically achieving peers. Students' age ranged from 11.0 to 19.2 with an average age of 14.5 across grades 6 to 12. All but two studies reported the gender of students with 91 males and 62 females across the studies. The race of students was provided in 14 studies with a variety of races represented including Caucasian (40%), African American (21.5%), Hispanic (5.4%) Asian (1.3%), Indian (1.3%), and Other (30.9%).

## *3.2 Setting*

There were nine studies that took place in a middle school, seven studies in a high school, and one in a university clinic. The most common setting where the writing intervention took place was in a separate pull-out classroom designed for research purposes (n = 8). This typically occurred when a researcher or graduate assistant was delivering the instruction. Additional settings included a special education resource room (n = 5) and general education classroom (n = 3; language arts and science). One study did not report the setting.

## *3.3 Interventionist*

In a majority of the studies, an individual from the university served as the interventionist (researcher = 5, graduate research assistant = 4, undergraduate research assistant = 1). In the other studies, the classroom teacher delivered the intervention (special education teacher = 4, general education teacher = 2). One study only reported that a teacher delivered the intervention but did not provide any further details about the teacher.

## *3.4 Independent Variable*

**Types of Writing and Strategies Taught Using SRSD**

All studies in this review utilized the SRSD instructional model to provide a writing intervention and incorporated a specific writing strategy. The studies included students composing argumentative essays (n = 8), narrative essays (n = 2), expository essays (n = 2), summaries (n = 2), informative essays (n = 1), compare and contrast essays (n = 1), and paragraph writing (n = 1). All writing strategies were represented by mnemonic devices which aid students in remembering the components of a strategy and reduces the cognitive load. It is important to note that the studies did not merely introduce the mnemonic devices but taught them as a tool to support students' learning within SRSD instruction. The use of mnemonic devices yielded improvements in students' writing due to the strategy being clearly aligned to the desired writing outcome.

**Argumentative Writing Strategies** Argumentative writing is when the author writes to persuade their reader to change their point of view. Eight studies in this review taught students an argumentative writing strategy that helped students during the planning and drafting stages of writing. The writing prompts typically posed a question that was relevant to secondary students (n = 6) such as, "Should students be allowed to eat snacks in the classroom?" (Geres-Smith et al., 2019, p. 43). Students completed a quick-write response within a 10- or 15-min timed session. Two studies utilized sample ACT essay prompts and students completed essays within a 40-min timed session.

The most common strategy utilized when teaching students to write an argumentative essay was POW + TREE (Geres-Smith et al., 2019; Hoover et al., 2012; Mason et al., 2013; Straub & Vasquez, 2015). The first part of the mnemonic introduces students to a general three step planning strategy, POW: (a) Pick an idea or side of a topic, (b) Organize ideas, and (c) Write and say more by modifying and improving the original plan. Students then learn the argumentative writing strategy, TREE: (a) write a convincing Topic sentence, (b) write at least three Reasons why you believe, (c) write Explanations to support each reason, and (d) wrap it up with a good Ending sentence.

Another argumentative strategy was DARE (Chalk et al., 2005; Eissa, 2009). This strategy helps students develop an argument on a topic, acknowledge opposing viewpoints, and provide a rebuttal. DARE taught students to (a) Develop your topic sentence, (b) Add supporting ideas, (c) Reject at least one argument from the other side, and (d) End with a conclusion.

The final argumentative strategy was developed to help students on the ACT college entrance essay exam. The strategy was HIT SONGS[3] (Ray & Graham, 2020; Ray et al., 2019). This strategy outlined essential elements to include in the introduction paragraph, three body paragraphs, and conclusion paragraph. The first part of the mnemonic, HIT, outlined the important parts of the introduction paragraph,

(a) Hook, (b) Introduce the topic, and (c) Thesis. The next part of the mnemonic, SONG, was repeated three times, for each body paragraph. The parts of SONG included (a) State the perspective, (b) Outlook on the perspective, (c) Need examples, and (d) Give your opinion. $S^3$ reminded students what needed to be included in the conclusion paragraph: (a) Support your thesis, (b) State the relationship between your thesis and the perspectives given in the prompt, and (c) Summary.

The studies of argumentative writing spanned grades 6 through 12 and all the approaches led to students with and at-risk for LD improving their argumentative writing abilities. Across the studies implementing an argumentative writing intervention, students improved in the number of argumentative genre elements, quality, organizational quality, length, number of transition words, and planning. The variety of writing strategies demonstrate that students struggling with writing can learn to effectively compose argumentative essays if taught the tools needed through strategy instruction.

**Narrative Writing** Another form of composition students learn is narrative writing, where the author conveys a real or imagined experience. Two studies in this review taught narrative writing strategies (Foxworth et al., 2017; Sundeen, 2012). One study followed the state writing test guidelines of a 45-min timed session and used sample state test prompts like, "A girl arrives home from school to find that the backdoor is ajar. She is not sure if she should go in. What should she do? Write a story about the girl, and what she does next" (Foxworth et al., 2017, p. 233). Students learned to plan their narrative using the mnemonics POW + STACS which stood for (a) Setting, (b) Tension, (c) rising Action, (d) Climax, and (e) Solution. The other study had students write personal narratives responding to prompts such as, "Describe your favorite vacation" (Sundeen, 2012, p. 26) during a 15-min timed session. In this study, students were taught to plan using mind-mapping with the mnemonic MIND: (a) Main, (b) Idea, (c) Numbered subtopics, and (d) Details.

The narrative writing studies spanned 7th through 11th grade. Students who participated in the interventions improved in the number of narrative genre elements, quality, and story grammar within their narrative essays. Overall, students' narrative writing abilities can be enhanced at the secondary level using the SRSD instructional model and narrative writing strategies.

**Expository Writing** Students are often asked to write expository essays where they explain their thoughts on a topic and support their ideas with explanations and evidence. Two studies used SRSD to teach students to write expository essays (Burke et al., 2017; De La Paz, 1999). Both studies used similar types of prompts. For example, "It is exciting to hear about people who win a million dollars in the lotto. Imagine that you were a lotto winner and won one million dollars. Write an essay explaining what you would do with one million dollars if you won it" (Burke et al., 2017, p. 91). Students in the study by Burke et al., (2017) completed 10-min expository quick writes, whereas the students in the study by De La Paz

(1999) wrote five paragraph essays during a 35-min timed session aligned to the state writing test procedures.

Both studies taught students the strategy PLAN and WRITE. The first part of the mnemonic, PLAN, stood for (a) Pay attention to the prompt, (b) List main ideas, (c) Add supporting details, and (d) Number your ideas. PLAN helped students consider the prompt, brainstorm ideas and details, and make an organizational plan for their essay. The second part of the mnemonic assisted students while drafting their essays and was represented by the mnemonic WRITE: (a) Work from your plan to develop thesis statement, (b) Remember your goals, (c) Include transition words in your paragraph, (d) Try to use different kinds of sentences, and (e) Exciting/interesting/$100,000 words. This strategy was utilized with students in seventh and eighth grade across the two studies and helped students improve their planning, expository genre elements, quality, and length of expository essays.

**Summary Writing** Another common writing task at the secondary level is writing a summary after reading text. When writing a summary, students must convey the most important information from a reading in their own words which can help improve both their reading comprehension and writing skills. Two studies taught students to write summaries after reading nonfiction text (Asaro-Saddler et al., 2018; Saddler et al., 2019). Students read a nonfiction story and were asked to write a summary. Students in the study by Saddler et al. (2019) had 20 min to write their summary. There was no time limit for summary writing in the study by Asaro-Saddler et al. (2018).

In both studies, students learned the strategy WINDOW, which stood for (a) Write a topic sentence, (b) Identify important information, (c) Number the pieces of identified information, (d) Develop sentences, (e) Organize sentences using transition words, and (f) Write an ending sentence. When identifying important information, students were taught to only include facts, reasons, and ideas from the author. Students were then instructed to develop sentences using their own words. Across the studies, students in both middle and high school learned to effectively write summaries that were longer and of higher quality as measured by the main ideas, supporting details, organization, conclusion, and mechanics and grammar within the summary.

**Informative Writing** Only one study focused on informative writing (Benedek-Wood et al., 2014) where the writer informs or teaches the reader about a topic. In this study, sixth grade students in a general education science class wrote informative essays about current science content. The prompts followed the following format, "What did you learn about...followed by the main topic (e.g., weather predicting methods)" (Benedek-Wood et al., 2014, p. 76). Students wrote their informative essays in a 10-min timed session. Students informative quick writes improved in length, number of transition words, organizational quality, and number of knowledge units included in the essay.

**Compare and Contrast Writing** Discussing the similarities and differences of concepts is required when writing a compare and contrast essay. One study taught students to write compare and contrast essays by drawing comparisons across categories information (as opposed to the text structure where all the similarities are described and then all the differences; MacArthur & Philippakos, 2010). All the prompts utilized the format, "Write a paper comparing and contrasting ____ and _____." (MacArthur & Philippakos, 2010, p. 443). Students had 35 min to write their essays. Students learned the mnemonics TAP and IBC. When planning compare and contrast essays, students used TAP to consider the (a) Topic, (b) Audience, and (c) Purpose for writing. When writing essays, students used IBC, for (a) Introduction, (b) Body, and (c) Conclusion to guide their text structure. For the introduction, students were taught to include a hook and a statement about what was being compared. The body paragraphs each included a topic sentence and relevant details or comparisons. The essays concluded with a summary of comparisons. The middle school students who were instructed using SRSD for compare and contrast writing improved the text structure and quality of their essays.

**Paragraph Writing** Writing a paragraph is an important academic skill. One study combined teaching self-determination and paragraph writing skills through teaching students to write Individualized Education Program goal paragraphs (Konrad & Test, 2007). Students wrote Individualized Education Program goal paragraphs within a 10-min timed session. Students learned the strategy GO 4 IT . . . NOW!: (a) Goal statement (topic sentence), (b) Objectives (4 of them, supporting details), (c) Identify a Timeline, (d) Name topic, (e) Order details, and (f) Wrap it up and restate topic. Students were instructed that the NOW part of the mnemonic applies anytime they need to write a paragraph. The seventh and eighth grade students that participated in this study improved in the quality of paragraph writing and articulation of their Individualized Education Program goals and objectives.

## *3.5 Length of Intervention*

The duration of the interventions varied from 100 to 900 min with an average amount of instructional time being 311 min. Four of the studies did not provide information about the number of instructional minutes and reported instruction as number of lessons or number of days of instruction.

## *3.6 Treatment Fidelity*

There was high treatment fidelity across the studies ($M$ = 98.9%, Range 94.2% – 100%), with three studies not reporting fidelity.

### *3.7 Dependent Variables: Writing Outcomes*

There are a multitude of approaches for evaluating writing. The most common outcome measures used to assess writing among the studies were quality (n = 13; Benedek-Wood et al., 2014; Burke et al., 2017; Chalk et al., 2005; De La Paz, 1999; Eissa, 2009; Foxworth et al., 2017; Konrad & Test, 2007; MacArthur & Philippakos, 2010; Mason et al., 2013; Ray & Graham, 2020; Ray et al., 2019; Straub & Vasquez, 2015; Sundeen, 2012), elements (n = 11; De La Paz, 1999; Foxworth et al., 2017; Geres-Smith et al., 2019; Hoover et al., 2012; Konrad & Test, 2007; MacArthur & Philippakos, 2010; Mason et al., 2013; Ray & Graham, 2020; Ray et al., 2019; Saddler et al., 2019; Straub & Vasquez, 2015), and length (n = 11; Asaro-Saddler et al., 2018; Benedek-Wood et al., 2014; Chalk et al., 2005; De La Paz, 1999; Eissa, 2009; Geres-Smith et al., 2019; Hoover et al., 2012; Konrad & Test, 2007; MacArthur & Philippakos, 2010; Mason et al., 2013; Ray et al., 2019). The quality outcome measures were a holistic, general impression of the students' writing whereas the elements outcome measures evaluated student writing by whether the essay included all the essential genre elements as outlined in instruction. Studies also looked to see if students wrote longer essays after participating in the writing intervention. To illustrate, in the study by Mason et al. (2013) students' persuasive quick writes were scored for quality, elements, and length. The organizational quality of the students' writing was scored on a seven-point holistic scale that looked at the inclusion of persuasive quick write elements and the organizational structure of the writing into paragraphs. The elements outcome measure was scored by students earning one point for each argumentative element included in their essay (i.e., one point for a topic sentence, one point for each reason included, one point for each explanation included, one point for a counter reason, one point for the rebuttal to the counter reason, and one point for an ending sentence). Additionally, the total number of words written were scored by using the word count feature of a word processing program. Some studies scored essays for an overall holistic quality and scored additional subcategories for quality. For example, the studies by Ray and Graham (2020) and Ray et al. (2019) used four quality subcategories including (a) ideas and analysis, (b) development and support, (c) organization, and (d) language use.

Additional writing outcome measures that occurred within multiple studies included planning (n = 5; Burke et al., 2017; De La Paz, 1999; MacArthur & Philippakos, 2010; Ray & Graham, 2020; Ray et al., 2019), transition words and phrases (n = 3; Burke et al., 2017; Ray & Graham, 2020; Ray et al., 2019), rubric (n = 2; Asaro-Saddler et al., 2018; Geres-Smith et al., 2019), and CWS-IWS (n = 2; Konrad & Test, 2007; Straub & Vasquez, 2015). The planning outcome measures were designed to examine whether students utilized the strategies they were taught during SRSD instruction when writing independently. This was often done by evaluating the planning pages students utilized to create a plan for their essay. Students' essays were also evaluated to examine the number of transition words and phrases that were used at the beginning and within paragraphs. Two studies developed or utilized rubrics from an already developed writing test to examine various aspects

of students' writing. Two studies examined writing mechanics using CWS - IWS which is the number of correct word sequences minus the number of incorrect word sequences. This measure evaluates correct spelling and semantic and syntactic accuracy within sentences.

Finally, there were several outcome measures that were only used in one study including knowledge units (n = 1; Benedek-Wood et al., 2014), story grammar (n = 1; Foxworth et al., 2017), CWS-IWS with spelling errors excluded (n = 1; Konrad & Test, 2007), conventions (n = 1; Burke et al., 2017), CWS (n = 1; Geres-Smith et al., 2019), percent CWS (n = 1; Geres-Smith et al., 2019), and writing duration (n = 1; Geres-Smith et al., 2019). The measure of knowledge units was utilized when evaluating informative writing. Students earned a point for each piece of accurate information about the topic that they included in their informative essays. When writing narratives, story grammar was measured on a 0 to 2-point scale to evaluate the development of the story structure (e.g., a student could earn 2 points for each action event that was logical and highly developed). The study by Konrad and Test (2007) utilized the measure CWS-IWS with spelling errors excluded. The authors wanted to evaluate students' abilities to write sentences that made sense and were grammatically correct, thus the use of CWS-IWS. However, the intervention did not address spelling skills, so they used the CWS-IWS measure but did not penalize students for misspelled words. Burke et al. (2017) developed their own conventions measure where students scored on a scale of 1 to 6 based on spelling and mechanics as well as paragraph breaks and sentence structure. CWS, percent CWS, and writing duration were all measures used in the study by Geres-Smith et al. (2019). CWS stands for correct word sequences that examines the number of adjacent, correctly spelled words in a sentence that are semantically correct. The percent of CWS is the CWS divided by the total number of word sequences written by the student. While many of the studies had a writing time limit, the students in the study by Geres-Smith et al. (2019) were not constrained by a time limit rather the amount of time students spent planning and writing their essays was recorded. Altogether, there was a wide variety of measures utilized to evaluate students' writing allowing for an in-depth examination of the specific aspects of writing where students made progress.

### *3.8 Methodology*

A variety of designs were implemented when evaluating the effectiveness of writing interventions using SRSD instruction with secondary students with and at-risk for LD. The majority of studies utilized a single-case research design (n = 12; see Table 1), which is commonly used in special education as an alternative to group designs. Additionally, three studies utilized a quasi-experimental design and two studies used an experimental design.

## 4 Discussion

Secondary teachers must prepare students to meet college level writing expectations. Research indicates that SRSD writing instruction is effective for secondary students with and at-risk for LD for a variety of writing genres and purposes. After examining the 17 articles included in this review, effective SRSD interventions for improving secondary students' with and at-risk for LD argumentative, narrative, expository, summary, informative, compare and contrast, and paragraph writing were identified. There are a variety of writing strategies represented by mnemonics that can be used to support students when planning and drafting within various genres and for an array of purposes.

### *4.1 Implications for Practice*

Teachers who are instructing students with and at-risk LD should utilize SRSD writing instruction to help improve their students' writing. Not only is SRSD an evidence-based practice with an effect size of 1.59 for improving the quality of students' writing (Graham & Harris, 2018), but SRSD is also a research-based strategy for secondary students with and at-risk for LD, as demonstrated by the 17 articles included within this review. SRSD instruction can also improve the quality, genre elements, total words written, and many additional aspects of secondary students' writing. To best support students' writing, teachers must select a writing strategy that directly aligns with the writing task. Teachers can select from a variety of strategies in the core genre areas (argumentative, informative, and narrative) as well as for more general writing purposes such as summary or paragraph writing that they feel will be meet the needs of their students. All in all, selecting a strategy that supports the writing task and teaching the strategy using the six stages of SRSD instruction along with self-regulation strategies is highly effective in improving the writing of secondary students and should be utilized by secondary teachers when teaching writing.

Furthermore, the researchers within the studies included in this review most commonly evaluated students' writing based on quality, elements, and length. When teachers are evaluating students' compositions, they need to be sure they are measuring the writing based on the strategies taught and practiced. For example, the argumentative writing strategy of TREE teaches students to include a topic sentence, reasons, explanations, and an ending. To evaluate an essay by a student who learned and practiced TREE on mechanics would be evaluating a student on something they had not been prepared to compose. In order to evaluate on mechanics, students would need to learn an additional strategy for revising or receive explicit instruction on mechanics. To support students' self-regulation of the writing

process, students need to be a part of the evaluation process through setting writing goals, self-evaluating their work, and monitoring and graphing their progress. When students are taught and practice these self-regulation strategies, they are able to recognize advances and areas of their writing still in need of improvement. Recording progress toward a writing goal on a graph provides students with a visual representation of their writing development and can increase students' writing motivation and buy-in to continue using the strategies learned when writing.

## *4.2 Limitations of Studies Reviewed*

Several limitations of the studies reviewed were identified. The first was that half of the studies took place in a separate pull-out classroom designed for research instruction. These students were all instructed by researchers. This is important to note as students may perform differently based on the instructor. While this is a crucial first step in developing interventions, additional studies need to be conducted where classroom teachers implement the intervention as teachers are ultimately responsible for the writing instruction of students. Moreover, many of the studies did not include generalization measures. Collecting information about how the writing skills taught using SRSD instruction transfers to other writing tasks is important (Graham & Harris, 2014).

Another limitation was that many studies had a small number of participants. Six of the studies had only three to six participants and five studies had 11–15 participants. Ten of these studies involved single-case research design. While a large number of participants is not required for studies using a single-case research design, greater replication of studies is needed to increase the generalizability of the findings. Additionally, across the studies many participants were Caucasian (40%). It is important to examine the effectiveness of interventions with students from a variety of racial and ethnic groups. Nevertheless, the findings overall are robust enough to make a strong claim that SRSD can be used to teach writing strategies to secondary students with and at-risk for LD, and that such instruction improves their writing performance.

## *4.3 Future Research*

After critically reviewing the research, it was clear much more research testing the effectiveness of SRSD writing interventions for secondary students with an at-risk for LD is needed. A majority of studies to date have been conducted with elementary school students with disabilities (Graham et al., 2013). This may be due to a focus on intervening early in students' school careers and that it can be easier to gain

accessibility to elementary students. However, many students in middle and high school have minimal to no basic skills in writing and need effective writing interventions.

Additionally, only two randomized control studies testing SRSD writing strategies with secondary students with LD were located. More randomized control trials are needed because they allow for greater validation and a more rigorous test of the effects of SRSD with larger groups of students. Additional replications of the single-case research design studies in this review are also needed to better generalize the results (Cook et al., 2014).

Furthermore, research needs to be conducted using writing strategies and SRSD instruction with students from a variety of racial and ethnic backgrounds with classroom teachers as the instructors. Studies also need to be conducted with editing strategies and with writing strategies used in conjunction with reading strategies to incorporate information from source text. Finally, future research should include examining the effectiveness of SRSD instruction with strategies for different writing purposes including writing to learn, writing in different content areas, and writing longer essays.

## 5 Concluding Remarks

A majority of high school graduates with disabilities do not have the writing skills needed to succeed in college-level English composition courses (U.S. Department of Education, 2011). Increasing the writing capacities of secondary students with and at-risk for LD is essential to improving their educational outcomes in middle and high school, college, and beyond. SRSD is an evidence-based practice that needs to be utilized when teaching writing to secondary students with and at-risk for LD. Students can make significant improvements in their writing ability when taught writing and self-regulation strategies using SRSD. Many writing strategies taught using SRSD have been researched to improve the writing of secondary students with and at-risk for LD including strategies for composing argumentative, narrative, expository, informative, and compare and contrast essays and writing summaries and paragraphs. Teachers must select a strategy that aligns with the genre or writing skill they are teaching and implement SRSD in the classroom to teach the writing strategy. Students with and at-risk for LD need strategies to help them be successful writers. Additionally, teachers need to provide explicit instruction on generalization of writing strategies learned to help prepare students for writing in a variety of contexts, including college level writing. Through examining the types of writing, strategies used, and writing outcomes within studies included in this literature review, it can be concluded that teaching students writing and self-regulation strategies using SRSD can enhance the crucial skill of writing for secondary students with and at-risk for LD.

## References


Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review, 10*, 129–154. https://doi.org/10.1023/A:1022185502996

Asaro-Saddler, K., Muir-Knox, H., & Meredith, H. (2018). The effects of a summary writing strategy on the literacy skills of adolescents with disabilities. *Exceptionality, 26*(2), 106–118. https://doi.org/10.1080/09362835.2017.1283626

Bangert-Drowns, R. L., Hurley, M. M., & Wilkinson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*(1), 29–58. https://www.jstor.org/stable/3516060

Benedek-Wood, E., Mason, L. H., Wood, P. H., Hoffman, K. E., & McGuire, A. (2014). An experimental examination of quick writing in the middle school science classroom. *Learning Disabilities: A Contemporary Journal, 12*(1), 69–92.

Burke, L., Poll, G., & Fiene, J. (2017). Response to an expository writing strategy across middle school RtI tiers. *Learning Disabilities: A Contemporary Journal, 15*(1), 85–101.

Chalk, J. C., Hagan-Burke, S., & Burke, M. D. (2005). The effects of self-regulated strategy development on the writing process for high school students with learning disabilities. *Learning Disability Quarterly, 28*(1), 75–87. https://doi.org/10.2307/4126974

Clinedinst, M., Koranteng, A. M., & Nicola, T. (2015). *2015 state of college admission*. National Association for College Admission Counseling. https://indd.adobe.com/view/c555ca95-5bef-44f6-9a9b-6325942ff7cb

Cook, B., Buysse, V., Klingner, J., Landrum, T., McWilliam, R., Tankersley, M., & Test, D. (2014). *Council for Exceptional Children standards for evidence-based practices in special education*. Council for Exceptional Children.

De La Paz, S. (1999). Self-regulated strategy instruction in regular education settings: Improving outcomes for students with and without learning disabilities. *Learning Disabilities Research and Practice, 14*(2), 92–106. https://doi.org/10.1207/sldrp1402_3

De La Paz, S., & Graham, S. (1997). Effects of dictation and advanced planning instruction on composing of students with writing and learning problems. *Journal of Educational Psychology, 89*(2), 203–222. https://doi.org/10.1037/0022-0663.89.2.203

Eissa, M. A. (2009). The effectiveness of a program based on self-regulated strategy development on the writing skills of writing-disabled secondary school students. *Electronic Journal of Research in Educational Psychology, 7*(1), 5–24.

Foxworth, L. L., Mason, L. H., & Hughes, C. A. (2017). Improving narrative writing skills of secondary students with disabilities using strategy instruction. *Exceptionality, 25*(4), 217–234. https://doi.org/10.1080/09362835.2016.1196452

Geres-Smith, R., Mercer, S. H., Archambault, C., & Bartfai, J. M. (2019). A preliminary component analysis of self-regulated strategy development for persuasive writing in grades 5 to 7 in British Columbia. *Canadian Journal of School Psychology, 34*(1), 38–55. https://doi.org/10.1177/0829573517739085

Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children, 80*(4), 454–473. https://doi.org/10.1177/0014402914527238

Graham, S., & Harris, K. R. (2018). Evidence-based writing practices: A meta-analysis of existing meta-analyses. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Design principles for teaching effective writing: Theoretical and empirical grounded principles* (pp. 13–37). Brill Editions.

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools*. Alliance for Excellent Education.

Graham, S., Harris, K. R., & McKeown, D. (2013). The writing of students with LD and a meta-analysis of SRSD writing intervention studies: Redux. In L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 405–438). Guilford Press.

Graham, S., & Harris, K. R. (2014). Conducting high quality writing intervention research: twelve recommendations. *Journal of Writing Research, 6*(2), 89–123. https://doi.org/10.17239/jowr-2014.06.02.1

Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., Kim, J., Lyskawa, J., Olson, C. B., & Smither Wulsin, C. (2016). *Teaching secondary students to write effectively* (NCEE 2017-4002). National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education. https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/wwc_secondary_writing_110116.pdf

Graham, S., Banales, G., Ahumada, S., Munoz, P., Alvarez, P., & Harris, K. R. (2020). Writing strategies interventions. In D. L. Dinsmore, L. K. Fryer, & M. M. Parkinson (Eds.), *Handbook of strategies and strategic processing* (pp. 141–158). Routledge.

Harris, K. R., & Graham, S. (1996). *Making the writing process work: Strategies for composition and self-regulation* (2nd ed.).

Hoover, T. M., Kubina, R. M., & Mason, L. H. (2012). Effects of self-regulated strategy development for POW +TREE on high school students with learning disabilities. *Exceptionality: A Special Education Journal, 20*(1), 20–38. https://doi.org/10.1080/09362835.2012.640903

Kiuhara, S. A., Graham, S., & Hawken, L. S. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology, 101*(1), 136–160. https://doi.org/10.1037/a0013097

Konrad, M., & Test, D. W. (2007). Effects of GO 4 IT ... NOW! Strategy instruction on the written IEP goal articulation and paragraph-writing skills of middle school students with disabilities. *Remedial and Special Education, 28*(5), 277–291. https://doi.org/10.1177/07419325070280050301

MacArthur, C. A., & Philippakos, Z. (2010). Instruction in a strategy for compare-contrast writing. *Exceptional Children, 76*(4), 438–456. https://doi.org/10.1177/001440291007600404

Mason, L. H., Kubina, R. M., Kostewicz, D. E., Cramer, A. M., & Datchuk, S. (2013). Improving quick writing performance of middle-school struggling learners. *Contemporary Educational Psychology, 38*(3), 236–246. https://doi.org/10.1016/j.cedpsych.2013.04.002

Ray, A. B., & Graham, S. (2020). A college entrance essay exam intervention for students with high-incidence disabilities and struggling writers. *Learning Disability Quarterly.* Advanced online publication. https://doi.org/10.1177/0731948720917761.

Ray, A. B., Graham, S., & Liu, X. (2019). Effects of college entrance essay exam instruction for high school students with disabilities or at-risk for writing difficulties. *Reading and Writing: An Interdisciplinary Journal, 32*(6), 1507–1529. https://doi.org/10.1007/s11145-018-9900-3

Reid, R., Ortiz-Lienemann, T., & Hagaman, J. L. (2013). *Strategy instruction for students with learning disabilities* (2nd ed.). Guilford.

Saddler, B., Asaro-Saddler, K., Moeyaert, M., & Cuccio-Slichko, J. (2019). Teaching summary writing to students with learning disabilities via strategy instruction. *Reading & Writing Quarterly, 35*(6), 572–586. https://doi.org/10.1080/10573569.2019.1600085

Santangelo, T., Harris, K. R., & Graham, S. (2016). Meta-analysis of the self-regulation processes in Zimmerman and Risemberg's model. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 174–193). Guilford.

Sawyer, R. J., Graham, S., & Harris, K. R. (1992). Direct teaching, strategy instruction, and strategy instruction with explicit self-regulation: Effects on the composition skills and self-efficacy of students with learning disabilities. *Journal of Educational Psychology, 84*(3), 340–352. https://doi.org/10.1037/0022-0663.84.3.340

Stone, C. A., & Conca, L. (1993). The origin of strategy deficits in children with learning disabilities: A social constructivist perspective. In L. J. Meltzer (Ed.), *Strategy assessment and instruction for students with learning disabilities* (pp. 23–59). PRO-ED.

Straub, C. L., & Vasquez, E. (2015). Effects of synchronous online writing instruction for students with learning disabilities. *Journal of Special Education Technology, 30*(4), 213–222. https://doi.org/10.1177/0162643415618929

Sundeen, T. H. (2012). Explicit prewriting instruction: Effect on writing quality of adolescents with learning disabilities. *Learning Disabilities: A Multidisciplinary Journal, 18*(1), 23–33.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, & National Assessment of Educational Progress. (2011). *Writing 2011: Grade 12 national results*. The Nation's Report Card. https://www.nationsreportcard.gov/writing_2011/g12_national.aspx?tab_id=tab2&subtab_id=Tab_7#chart

# Part V
# Teacher Practice and Professional Development

# "Nothing Continued to Happen": Addressing Attrition, Lack of Fidelity, and Other Barriers to Implementation in High-Poverty, High-Mobility Urban Settings


**Erin FitzPatrick, Debra McKeown, Megan C. Brown, and Nicole Patton-Terry**

**Abstract** The demand for writing skills is increasing in academic settings and careers. Recent changes to learning standards and high-stakes testing have placed writing at the center of learning. In response to principals' request, a complex, elegant five-school study was conceived, but collapsed during rollout. In the aftermath, one principal requested support for teacher development for writing instruction. To this end, six 3rd, 4th, and 5th grade teachers of English and Language Arts at an urban, public elementary school received 5 days (30 h) of professional development in evidence-based writing instruction. The professional development included student strategies for using the writing process, strategies for analyzing prompts, and foundational writing instruction for both persuasive and informational essays. The writing instruction provided was self-regulated strategy development, an evidence-based writing intervention (What Works Clearinghouse, *Teaching elementary school students to be effective writers: Educator's practice guide*. http://ies.ed.gov/ncee/wwc/practiceguide.aspx?sid=17, 2012). Researchers measured fidelity of implementation; student writing outcomes including essay length, genre elements, and holistic quality; and social validity. In this chapter, along with the results of this study, we discuss the challenges of conducting rigorous scientific research in

E. FitzPatrick (✉)
University of North Carolina at Charlotte, Charlotte, NC, USA
e-mail: Erin.FitzPatrick@charlotte.edu

D. McKeown
Texas A&M University, College Station, TX, USA
e-mail: DebraMcKeown@tamu.edu

M. C. Brown
Stats Fabulous, L.L.C, Burien, WA, USA
e-mail: MeganCBrown@gmail.com

N. Patton-Terry
Florida State University, Tallahassee, FL, USA
e-mail: NPattonTerry@fsu.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_14

high-poverty, high-mobility urban settings with a focus on the ethical questions and considerations when working in settings with multiple competing priorities.

**Keywords** Writing · Professional development · Writing instruction · Elementary education · Self-regulated strategy development

The demand for writing skills is increasing in academic settings and careers. Changes to learning standards and the emphasis on high-stakes testing have increased the focus on the need to write well in and beyond school settings. Nearly three out of four students are not meeting basic proficiency in writing performance and the written performance of students with learning disabilities and those from marginalized racial and ethnic groups are even more dire (National Center for Education Statistics, 2011). Moreover, most teachers report not being well prepared to teach writing and do not view themselves as writers (Brindle et al., 2016).

## 1 Practice-Based Professional Development

Practice-based professional development (PBPD) provides teachers with an environment of supported practice to add new skills and materials to their instructional practices (Ball & Cohen, 1999). PBPD emphasizes the development of teacher content knowledge, pedagogy, collaboration, and the practice of new skills in an environment with responsive expert feedback. Teachers practice lessons and receive feedback in this supportive environment which ensures active participation and facilitates real-time problem-solving (Ball & Cohen, 1999).

PBPD is characterized by (a) working with colleagues, (b) personalized support in response to teachers' classroom context, (c) expert assessment and instruction of content knowledge, (d) explicit modeling and opportunities for teachers to model every activity and receive peer feedback, (e) identical materials that will be used in the classroom, and (f) guidance in the areas of differentiation (Ball & Cohen, 1999; Harris et al., 2012). Several studies have demonstrated that following PBPD for writing, teachers can offer evidence-based practices with high fidelity, report high social validity for the professional development experience, and meaningfully impact student writing outcomes (FitzPatrick & McKeown, 2020; Harris et al., 2012; McKeown et al., 2016).

## 2 Self-Regulated Strategy Development

Self-regulated Strategy Development (SRSD) is a multi-component, criterion-based, strategic instructional method which fosters acquisition and adoption of new learning and has been validated as an evidence-based practice by multiple

independent agencies (e.g., What Works Clearinghouse, 2012). SRSD consistently has the highest effect sizes among writing interventions in meta-analyses (Graham et al., 2012, 2013).

SRSD for writing provides skills instruction with an emphasis on genre-knowledge as students are explicitly taught an essay writing strategy through a recursive instructional approach featuring these six stages: (a) Develop Background Knowledge, (b) Discuss It, (c) Model It, (d) Memorize It, (e) Support It, and (f) Independent Performance (Harris et al., 2008). SRSD includes explicit instruction to support students in self-regulation throughout the cognitively demanding task of writing while addressing discrepancies in initial schema, genre knowledge, motivation, and self-efficacy. Through SRSD, student performance is actively scaffolded by way of gradual release of responsibility following teachers explicitly modeling lessons including detailed think alouds exposing the cognitive processes (Harris et al., 2008).

## 3 Prelude to the Current Study

We were approached by the leadership of a five-school cluster of urban schools to help them address low student writing achievement. Due to low Annual Yearly Progress (AYP), the cluster had been targeted for improvement and risked impending takeover by the state. The principals were working collaboratively to develop their annual professional learning program and after reviewing data, had designated writing as a target skill. At the initial meeting, the research team explained both PBPD and SRSD, past results, and how implementation could work. All five principals agreed to move forward.

### *3.1 Plan 1*

The research team designed a wait-listed randomized controlled trial in which the five schools would be randomly assigned to treatment or control. Then, after completing implementation and post-testing, the control group would receive the intervention.

However, after learning about the requirements of participating in a scientific study (e.g., delayed access to the intervention for some schools, student assessment before beginning instruction, required classroom observations), three of the five schools decided against moving forward with the study. The principals wanted guaranteed access to the intervention immediately as their students needed effective writing instruction.

### 3.2 Plan 2

The remaining schools' leaders wanted us to address both persuasive and informational writing, so we required additional development days. The principal from one of the two remaining schools indicated her teachers could not all come together for the required number of days, but she still wanted to participate. To address this, we proposed offering a hybrid model of professional learning whereby her teachers would engage in computer modules for half of the instruction as well as in-person learning. This altered our study and lessened the scientific rigor of the design, but would allow us to provide the intervention to both schools simultaneously and still contain a scientific comparison – in-person versus hybrid professional learning presentations. With this change, the research team anticipated a smaller effect size than that which is typically found in SRSD studies.

We created the online learning modules, recorded modeling videos, and designed formative assessments throughout the online program. We also scheduled dates for school visits to obtain teacher consent, in-person professional learning, hybrid professional learning, school visits to seek student assent and parent permission, pretesting, implementation and observation, and post-testing.

After consenting procedures were completed, we provided the first day of professional learning, in-person sessions for both schools. Days 2 and 3, which completed the persuasive genre, were in-person for one school (Governors Elementary; all schools names are pseudonyms) and online learning modules for the other (Edwards Elementary) to be completed within 2 weeks. Meanwhile, pretesting of both the persuasive and informational genre took place in both schools. Then, teachers began to implement persuasive lessons, and research assistants observed teachers for fidelity.

Observers were trained to contact teachers and schedule observations in advance during writing times. However, when observers arrived at scheduled times, they were often met with teachers saying they were unable to teach writing that day or had already taught writing. Two teachers at Edwards Elementary reported they had not completed the online professional learning, so they did not plan to implement SRSD. We contacted the principal to ask how we could support her teachers. The principal reported that she had attended the teacher planning meeting that week, and the lead teacher indicated that despite low writing scores on standardized assessments which had caused them to target writing for improvement, the teachers were overwhelmed with initiatives, contented with teaching writing the way they had been which resulted in those scores, and the principal felt compelled to support her teachers. Thus, that school opted out of the study.

### 3.3 Plan 3: The Present Study

Still intent to support the remaining teachers and students, we continued, altering the research design to pre/post only, using students as their own control, a weaker within-subject design with too few teachers to account for differences between

classrooms and teachers. Again, with this change, the research team anticipated a smaller effect size than found in previous SRSD studies.

The following research questions were addressed: (a) To what extent does PBPD for SRSD writing in persuasive and informational genres result in teacher implementation with fidelity? (b) To what extent does PBPD for SRSD writing in persuasive and informational genres improve student writing in terms of genre elements, holistic quality, and length of third-, fourth-, and fifth-grade student writing when taking timed writing assessments? (c) To what extent do teachers find this intervention to be socially valid?

## 4 Method

The intervention took place in an elementary school located in a major metropolitan area in the Southeast. Governors Elementary had approximately 425 students enrolled from Pre-K to 5th grade (46% female, 54% male; 99% African American/Black, 1% Hispanic, and 0.4% White). All of the students (100%) received free or reduced lunch.

At Governors Elementary, three 3rd grade teachers, two 4th grade teachers, and one 5th grade teacher consented to participate (see Table 1 for details). Five of the six teachers held graduate degrees and were certified in the area in which they were teaching. These teachers had been teaching from 4 to 23 years and had been teaching the grade level at the current school between 1 and 5 years. Teachers were required by their principal to attend the professional development sessions; they were not volunteers.

All students from the participating teachers' classrooms were invited to participate. We received 124 consent/assent forms from students. There were high rates of transfer and/or absenteeism in the school. These factors detrimentally affected the sample size for analysis. Of participating students, only 34 had a pre- and post-test for the persuasive genre and only 29 had a pre- and post-test for the informational genre. Since teachers presented the genres in succession, there is some, but not total overlap in the populations reported for each genre. The results reported are on the small population of students who returned the required forms for participation in the study and who were present for both the pre- and the post-test for a genre (see Table 2 for student demographic information).

### *4.1 Procedures*

This intervention involved two different levels: teacher-level and student-level. First, teachers participated in PBPD in the persuasive genre. Second, teachers pretested students in each writing genre (persuasive and informational). Third, teachers implemented the writing intervention in their classrooms, and trained observers

**Table 1** Teacher demographic information

| | Teacher 1 | Teacher 2 | Teacher 3 | Teacher 4 | Teacher 5 | Teacher 6 |
|---|---|---|---|---|---|---|
| Grade taught | 3 | 3 | 3 | 4 | 4 | 5 |
| Gender | M | M | F | F | F | M |
| Race | AA | AA | AA | AA | AA | AA |
| Undergraduate education | BA, Psychology | BS, Elem. Ed. | BS, Psychology/ Sociology[a] | BS, Elem. Ed. | BA, Early Childhood Ed. | BA, Elem. Ed. |
| Graduate education | MA, Elem. Ed[a] | None | MS, Elem. Ed M.Ed., Ed. Leadership Ed.D., Instructional Leadership | MA, Elem. Ed. | MA, Early Childhood Ed. w/ Reading | MA, Early Childhood Ed. |
| Certified in current teaching assignment? | Yes | Yes | Yes | No | Yes | Yes |
| Years teaching | 5 | 15 | 17 | 23 | 4 | 17 |
| Years in current grade level | 2 | 5 | 1 | 3 | 4 | 7 |
| Years in current school | 2 | 5 | 1 | 3 | 4 | 7 |
| Years in current grade at current school | 2 | 5 | 1 | 3 | 4 | 7 |

Note: [a]Online degree
*AA* African American, *BA or BS* bachelor degree, *MA* or *MS* master's degree

visited the classes to collect fidelity of implementation data. Next, after intervention of the persuasive genre was complete, teachers post-tested on the persuasive genre. Then, teachers participated in PBPD in the informational genre. Teachers implemented the intervention classwide. Afterward, teachers tested students on the informational genre. Finally, after all post-tests were complete, researchers invited teachers to participate in focus group interviews. We further explain these procedures below.

## *4.2 Professional Development for SRSD*

We provided 5 days (30 h) of PBPD in the area of writing instruction to the six 3rd, 4th, and 5th teachers of English and Language Arts at Governors Elementary School. Four teachers attended the first session and two attended the makeup session. To counteract teacher absenteeism on days of professional development, researchers actually provided 10 days of PBPD in total; that is, we conducted

**Table 2** Student demographic information

| | Participated in study | | | | Did not participate | |
|---|---|---|---|---|---|---|
| | Persuasive | | Informational | | | |
| | N | % | n | % | n | % |
| **Total n** | 34 | | 29 | | 90 | |
| **Grade** | | | | | | |
| 3rd | 21 | 62% | 21 | 72% | 42 | 47% |
| 4th | 11 | 32% | 7 | 24% | 37 | 41% |
| 5th | 2 | 6% | 1 | 3% | 11 | 12% |
| **Gender** | | | | | | |
| Female | 21 | 62% | 17 | 59% | 43 | 47% |
| Male | 13 | 38% | 12 | 41% | 47 | 52% |
| **Race/Ethnicity** | | | | | | |
| African American | 33 | 97% | 28 | 97% | 90 | 100% |
| Biracial | 1 | 3% | 1 | 3% | 0 | 0% |
| **Free Lunch** | 34 | 100% | 29 | 100% | 90 | 100% |
| **Special Education Designation** | | | | | | |
| Speech | 2 | 6% | 2 | 7% | 3 | 3% |
| Emotional/behavior | 3 | 9% | 3 | 10% | 7 | 8% |
| Specific learning disability | 0 | 0% | 0 | 0% | 2 | 2% |
| **Early Intervention** | 11 | 32% | 11 | 38% | 12 | 13% |
| **Referred for intervention** | 2 | 6% | 2 | 7% | 18 | 20% |
| **Designated homeless** | 0 | 0% | 0 | 0% | 4 | 4% |

makeup sessions only for teachers who were absent for originally scheduled sessions. The instructional coach, special education teachers, and paraprofessionals attended the PD as well. The professional development included student strategies for using the writing process, analyzing writing prompts, and SRSD writing instruction for both persuasive and informational essays.

## Persuasive Genre

Teachers participated in three consecutive days of PBPD in teaching SRSD for the persuasive genre (Ball & Cohen, 1999; Harris et al., 2012) and 1 day of supported preparation, an element requested by teachers in prior studies (McKeown et al., 2014). In total, there were seven persuasive writing lessons (Lesson 0 through 6), each oriented around the stages of SRSD. Each lesson varied in the number of learning sessions it may take to implement, giving teachers flexibility in their schedules. For each lesson, a researcher/facilitator modeled teaching the lesson to the group, using the same materials as teachers would in the classroom. Then, the teacher participants practiced teaching the lesson to peers, with facilitators and peers observing and giving feedback. At the end of each practice lesson, teachers and facilitators discussed ideas on how to differentiate for individual students.

On the first day, we set ground rules for collaborative learning, made introductions, and the teachers discussed their beliefs and practices in writing instruction, their students' strengths and weaknesses in writing, and learned the theoretical foundations of SRSD. After this, the facilitators modeled Lesson 0 followed by teacher practice. This lesson focused on analyzing prompts to determine the genre and purpose of the task. The state writing test was prompt-based and could be a persuasive, informational, or narrative writing task; students had to correctly identify the genre of the prompt to correctly respond. Then, facilitators modeled Lesson 1, focused on building background knowledge of good writing, the writing process, and the persuasive genre. After teachers practiced Lesson 1, they received their homework assignment. We asked teachers to write a persuasive essay in response to a prompt, write at their students' grade level, and use local names and places students would find engaging. Later, researchers shared each teacher's essay with others. This way, teachers had private practice in writing, but also had a set of exemplar essays for their lesson plans.

On the second day, the modeling and practice cycle was repeated for the remainder of the lessons – 2 through 6 – for persuasive writing. Lesson 2 focused on the Discuss It stage of SRSD. In this lesson, teachers discussed each step of the writing strategy, viewed examples, learned about self-regulation, and through goal setting made a commitment to use the strategy to secure better writing outcomes. There was special emphasis on Lesson 3, the teacher modeling lesson, as it is typical for teachers to struggle the most with this skill. Both explicit modeling and teacher presentation of Lessons 4 – a collaborative model where students take additional ownership of the tools while still fully engaged with the teacher; Lesson 5 – supported practice in which students use the tools more independently with support as needed; and Lesson 6 – independent performance – followed.

On the third day, we discussed rubric sensitivity and reliability. Teachers expressed concern about district-provided rubrics, so, after providing examples, we asked teachers to develop rubrics that would be sensitive to their instruction but also quick and easy to score regularly. We also discussed data-based instruction and how to use frequent checks on writing to inform instruction. Finally, we discussed the research process and fidelity observation visits.

The fourth day was set aside for teachers to plan for implementation. This included making anchor charts and posters for their classrooms (materials were provided), scheduling lessons across teachers, and aligning rubrics across grades such that one grade built on the next. Teachers then agreed to implement SRSD for writing for at least 30 minutes twice weekly.

### Informational Genre

We completed PBPD for the informational genre in a single day due to the teachers' experience with SRSD. The research team replicated the routines of the first PD modeling each informational lesson, and the teachers practiced teaching each

lesson. The day ended with questions and scheduling implementation. Again, we asked teachers to write example essays to use in their lessons.

#### Materials

We provided teachers with a binder that contained detailed lesson plans, all required student materials, example essays, a bank of prompts, a pacing calendar, example rubrics, and student writing folders. We also provided electronic copies of all materials.

### *4.3 Classroom Observations*

To assess fidelity of implementation, a member of the research team observed every third writing lesson and rated using the fidelity checklist for the given lesson. Teachers used these checklists during PBPD, saw them used during lesson modeling, and received copies in their binders. It was important that teachers felt our visits to the classroom were supportive and not critical. Thus, we were committed to transparency about the nature of the observations, the feedback, as well as the confidential nature of observations (i.e., we do not report on classroom activities to anyone; all data reported are de-identified).

#### Observation Feedback

Following each classroom observation, the teacher received an email with feedback. We trained graduate research assistants (GRAs) to use the following protocol: (a) Dear [teacher's name]; (b) Thank you for allowing us to visit/observe (or any version of this that seems natural to the observer); (c) State the positive aspects of the lesson, steps completed correctly; (d) positively note any appropriate changes they made to the lesson plan; (e) state the lesson steps the teacher missed, did not complete, or did not do well in the form of: "For future lessons, we have a few pointers we would like to share"; (f) Thank you again; (g) We will see you [next scheduled observation]; (h) Signature and contact information.

We instructed observers to write each email naturally, with a friendly tone, and without formulaic phrases. Observers based email content on elements of the fidelity checklist. Observers sent the email to the feedback coordinator who checked for adherence to protocol. Following confirmation, observers sent the email to the teacher and principal investigator. This process was completed within 24 h of observations, and emails were typically sent by 5 pm the same day. In focus groups, teachers indicated they appreciated the timely feedback as it gave them a boost to know what they were doing right and helped them focus

on how to improve the next time. There were no negative statements in the focus group about email feedback.

## *4.4 Teacher Measures*

Teacher outcomes were measured by fidelity observations and post-intervention interviews to determine social validity. We report fidelity to intervention below followed by preliminary results from the focus group interviews regarding social validity.

### Fidelity to Intervention Observations

A member of the research team provided teachers fidelity checklists for each genre during PBPD to ensure they understood the key lesson parts as well as to be transparent about researcher observations. Teachers were encouraged to use the fidelity checklists themselves to help stay on track with the lessons. Trained GRAs observed teachers during writing instruction. Observers attended 33.3% of planned writing instruction sessions for each teacher. During these observations, teacher fidelity was tracked using a checklist for the lesson being taught that day. Each checklist operationalizes the lesson parts. Of the observed lessons, a second trained GRA observed another 33.3% to obtain inter-observer agreement (IOA). This means 11.1% of the lessons taught were observed independently but simultaneously by two GRAs; agreement was calculated by dividing the number of agreements by the total number of observed lesson parts. IOA was 95.6%.

### Social Validity

After intervention was complete, we invited teachers to participate in a one-hour focus group interview. Using open ended questions, we queried what the teachers found useful and what needed to be changed in both the professional development and the intervention. Focus groups were audio recorded, and we took field notes.

## *4.5 Student Measures*

To evaluate student writing outcomes, members of the research team scored a pre- and post-test in the same genre. Students wrote to a persuasive and an informational writing prompt prior to intervention. Following the completion of the respective genre instruction, students took a post-test on a different persuasive or informational prompt.

### Writing Prompts

We based writing prompts on common topics students were expected to know without conducting additional research or consulting other sources. As an example, we used this prompt for the persuasive genre, "What should your bedtime be? Write a letter to your parents to explain what time you think you should go to bed each night. Make sure to give reasons for your opinion." and this for the informational genre, "What would be the best party ever? Describe what the party would be like."

**Controlling for Scoring Bias** Trained GRAs typed each student essay to avoid surface-level errors interfering with objective scoring of student writing (Graham et al., 2011). They corrected spelling, basic punctuation, and capitalization in the typed versions. A different trained GRA checked the typing for accuracy. Disagreements were resolved through discussion and agreed-upon changes made.

Prior to scoring, we developed a rubric and a codebook of definitions for each genre. We trained GRAs to criterion (>0.90 agreement) on one genre at a time. GRAs scored 66.7% of the essays. To determine the reliability of scoring, a different scorer rated 33.3% of the other's essays. These secondary raters worked independently. Reliability of scoring was 0.94 for persuasive essays and 0.88 for informational essays. In all cases, scorers were blind to the time point in which students wrote the essays (pre or post). Additionally, they were blind to the grade level, teacher, and gender.

**Persuasive Genre Elements** The research team scored persuasive essays using a rubric aligned with the key elements of the genre. This rubric included the following elements and possible points indicated in parentheses: position (0–1); elaborated position (0–1); hook (0–1); reason 1 (0–2); evidence 1 (0–1); reason 2 (0–2); evidence 2 (0–1); reason 3 (0–2); evidence 3 (0–1); extras (0–2); conclusion; and transition words (0–2). See Appendix A for detailed scoring procedures for the persuasive genre.

**Informational Genre Elements** The research team scored informational essays with a rubric aligned with informational genre elements. This included the following elements: premise/purpose (0–1); refined purpose (0–1); hook (0–1); idea 1 (0–2); support 1 (0–1); idea 2 (0–2); support 2 (0–1); idea 3 (0–2); support 3 (0–1); conclusion (0–2); and linking words (0–2). See Appendix B for detailed scoring procedures for the informational genre.

**Holistic Scores** Holistic quality scores are based on a scale from 0 to 6 points and are designed to measure organization, development, sentence fluency, word choice, and audience awareness. Despite evidence that it is difficult to show improvements in holistic scores in short term interventions (Mckeown et al., 2019a, b), we include these scores because holistic scores are the most frequently used measure of writing quality (Graham & Perin, 2007). However, holistic scores are influenced more by

length of an essay, flow, and depth of essay development than by number of essay elements. This means that while we would expect holistic quality to improve across time as the child develops skills with the strategy, holistic quality is not as sensitive to strategies-based instruction as genre elements scores are.

To score the essays holistically, each essay was judged as a whole, not weighting any one category more than another. We trained scorers to rate essays holistically in a six-hour session led by the second author. Scorers evaluated persuasive essays and informational essays independently and we held separate training sessions with distinct anchor papers for each. Raters were blind to the time of testing; that is, they did not know if they were scoring a pre-test or a post-test.

We used a random selection of essays from the current participants during training to collectively establish the anchor papers representing each score from 0 to 6. At first, scorers read an essay and then discussed its various attributes with one another and what score might be represented by the essay. This is what the second author calls establishing "group think." Then, scorers rated papers individually. We compared scores and defended our reasons to one another. Everyone talked until the group achieved consensus on a score. When an essay clearly represented a given score to the group, we made it an anchor paper until we had 1–2 examples for each scoring point. Once the group established a collection of anchor papers, we trained scorers to rate essays using the following procedure: (a) read the essay carefully, but not laboriously; (b) consider what rating the essay might receive; (c) read the anchor paper representing that score as well as the score below and above it; (d) determine which anchor paper is most similar to the essay being rated; (e) assign a score. At the end of training, agreement between scoring pairs was >.90. We assigned each rater 66.7% of the essays assigned to the pair (33.3% of all essays were scored twice for agreement). We calculated agreement across time to identify and control for drift.

### *4.6 Study Design and Data Analysis*

The study was a within-participant design, with pre- and post-tests for both genres (informational and persuasive). Students acted as their own baseline so growth in essay elements, holistic scores, and length could be assessed. Four one-way ANOVAs were run for each outcome and essay type to determine the effects of pre-test scores on holistic and genre elements scores. We report fidelity with simple descriptive statistics and social validity through preliminary analyses of focus group responses.

## 5 Results

In this pre-/post-design with students as their own control, we explored the extent to which teachers could implement SRSD following PBPD with high fidelity in persuasive and informational genres; the impact of PBPD for SRSD on student writing outcomes

**Table 3** Average % fidelity by teacher

| | Average % fidelity |
|---|---|
| Teacher 1 | 65% |
| Teacher 2 | 66% |
| Teacher 3 | 88% |
| Teacher 4 | 91% |
| Teacher 5 | 83% |
| Teacher 6 | 59% |
| Overall average | 75.3% |

**Table 4** Average % fidelity by lesson

| Lesson 0 | Lesson 1 | Lesson 2 | Lesson 3 | Lesson 4 | Lesson 5 | Lesson 6 | Avg. |
|---|---|---|---|---|---|---|---|
| 71.4% | 89.9% | 81.1% | 59.9% | 100% | 70.6% | 73% | 77.9% |

in terms of genre elements, holistic quality, and length; as well as teachers' evaluation of social validity of both the professional development and the SRSD writing intervention.

## *5.1 Teacher Outcomes*

### Fidelity to Intervention

Although each ELA teacher from Grades 3–5 attended and implemented the strategies in their classrooms, these teachers varied in their fidelity to the intervention. In this study, teacher fidelity of implementation ranged from between 59–91% with an average of 75.3% (see Table 3). Fidelity per lesson ranged from 59.9% to 100% with Lesson 3 posing the most difficulty in terms of including the critical aspects of the lesson (see Table 4).

### Social Validity

Teacher perceptions of PBPD and the writing interventions were measured via focus group interviews. Every teacher believed the professional development was "time well spent." Every teacher also said their students benefited from the instruction and they would teach SRSD again. One of the most frequent ideas stated was that teachers appreciated having writing instruction that they could implement comfortably and wished they had more time to spend teaching writing but testing and test preparation frequently got in the way. High social validity is consistent with prior work in SRSD and PBPD (FitzPatrick & McKeown, 2020; McKeown et al., 2019b).

### 5.2 Student Results

**Persuasive Writing**

Persuasive essays were scored for both genre elements and holistically. The change in score on the genre elements rubric was an average increase of 2.56 elements (*SD* = 2.89, range = 14) on a 14-point rubric. Due to large standard deviations (likely due to variation in implementation) and small sample sizes (and limited power to detect effects), there was no evidence of a main effect of pre-test genre elements scores ($M = 3.27$, $SD = 2.18$) on post-test genre elements scores ($M = 5.82$, $SD = 2.39$), $F(1, 32) = 1.35$, $p = 0.24$; however, there was a practical increase with student including, on average, nearly twice as many elements in their post-test essays. Holistic scores rose on average by 0.79 points (*SD* = 1.49, range = 6) on a 0–6 scale. Again, while there was no evidence of a main effect of pre-test holistic scores ($M = 1.41$, $SD = 0.99$) on post-test holistic scores ($M = 2.21$, $SD = 1.12$), $F(1, 32) = 0$, $p = 0.99$, there was a considerable practical increase. Students had a mean decrease in essay length of 26.71 words (*SD* = 64.64, range = 296; See Table 5). To summarize, students included more genre elements, produced higher quality essays, and did so in fewer words following SRSD instruction for writing.

**Informational Writing**

Informational essays were scored for both genre elements and holistically. The genre elements rubric score had an average increase of 1.83 elements included, (*SD* = 3.27, range = 14). There was no significant difference between pre-test ($M = 4.76$, $SD = 2.96$) and post-tests genre elements scores ($M = 6.59$, $SD = 2.52$), $F(1, 27) = 2.54$, $p = 0.12$. Holistic scores revealed an average increase of 0.62 (*SD* = 1.45, range = 6). There was no significant change between pre-test ($M = 2.41$, $SD = 1.45$) and post-test holistic scores ($M = 2.45$, $SD = 1.48$), $F(1, 27) = 1.46$, $p = 0.24$. Students had a mean change in essay length of 10.31 words (*SD* = 110.68, range = 486). See Table 5.

## 6 Discussion

Below, we discuss implications of the findings from this study as well as limitations. First, we present the implications of the teacher-level outcomes. Then, we present the implications of the student-level outcomes. Finally, we suggest future directions of both research and practice.

**Table 5** Student writing measures

| | Mean change | SD |
|---|---|---|
| *Persuasive (n = 34)* | | |
| Length | −26.70 | 63.64 |
| Genre Elements | 2.56 | 2.89 |
| Holistic | 0.79 | 1.49 |
| *Informational (n = 29)* | | |
| Length | 10.31 | 110.67 |
| Genre Elements | 1.83 | 3.27 |
| Holistic | 0.03 | 1.82 |

*Note*: Of the 124 consented participants, 51 and 59 completed either a pre- or post-test in the persuasive and/or informational genres respectively; 14 students who agreed to participate were never present for testing

## *6.1 Teacher Outcomes*

### Fidelity

We were concerned with the effectiveness of the professional development experience on teacher fidelity of implementation as fidelity is typically highly correlated with outcomes (Cordray, 2007). The lesson with the lowest fidelity was Lesson 3. This finding is not surprising as it is in Lesson 3 that teachers are required to model writing an essay from start to finish while using think-alouds as well as self-statements. This lesson has been found to have the lowest fidelity in prior implementations (McKeown et al., 2019b). While each teacher saw this lesson modeled and practiced teaching this lesson to peers during PBPD, it still posed a challenge in the classroom. Teachers reported they are not comfortable writing essays themselves, thinking aloud as they write, using self-statements aloud, and modeling while students watch/listen. These findings are consistent with prior research (McKeown et al., 2016, 2019a). The research team addressed each of these struggles in professional development; however, this finding indicates trainers and teachers may need to spend more time practicing prior to implementation.

While fidelity is typically found to be in the 90% range in studies of SRSD for writing, 78% average fidelity with a range of 59–91% is comparable to a larger study conducted in the same metropolitan center (McKeown et al., 2019b). Only Lessons 1 and 4 achieved fidelity commensurate with most prior research. The variability in fidelity is likely due to several factors.

First, teachers who volunteer to participate in professional development are more likely to implement interventions, to implement more quickly, and for longer periods of time with higher fidelity than non-volunteers (Johnson et al., 2013). The teacher participants in this study were not volunteers in the strictest sense of the word. They agreed to participate in the study voluntarily, but their principal required them to participate in professional learning and to implement the intervention

regardless of their participation in the research or not. Additionally, the initial round of professional development was prior to school starting and overlapped with before-school preparation days. Understandably, teachers were not able to be fully present during professional development because they were concerned with getting their classroom and curriculum ready for the year. The teachers who were most distracted (e.g., were in and out of PD attending to other issues, on computer for reasons unrelated to PD) were also the most inconsistent in implementation and had lower fidelity. This may indicate increased engagement in professional development impacts implementation. However, it may also indicate that teachers who feel overwhelmed may not yet be ready to effectively implement new interventions.

Some variability may also be explained by the unique context of urban education environments. In urban school settings, teachers are likely to have less control over curriculum and class time than suburban or rural teachers (Center for Technology and Education [CTE], 2016). Urban schools are more likely to have lower achievement scores, higher rates of poverty, and higher rates of incidents affecting academics (e.g., absenteeism, classroom discipline; CTE, 2016). In this study, the participating school had not met AYP, causing a disproportionate amount of class time to be spent on testing and test preparation. We recorded multiple instances whereby the writing instruction was delayed or not completed due to other demands on instructional schedules. These demands, or barriers to implementation, included field trips, preparation for statewide testing, class tests, career day, assembly practice, teacher sickness, and other state mandated benchmark testing. These results align with other studies conducted in the metropolitan area (e.g., McKeown, et al., 2019b). Additionally, research initiatives are often a low priority for teachers (Gersten et al., 1995) and identifying time within a very busy school schedule to implement interventions is a crucial challenge for implementation (Foorman et al., 2016).

Even with lower fidelity than is typically found in similar studies, student outcomes were positive. While high intervention fidelity is correlated with positive outcomes (Durlak & DuPre, 2008), Balu & Doolittle (2016) has addressed intervention implementations stating where there is a "threshold level of fidelity" and a "stark contrast" with the usual program, it is still possible to achieve positive outcomes even with lower fidelity (p. 107). While there is no comparison condition in this study to provide the "stark contrast" this phenomenon has been seen in another study in the region in which fidelity of implementation was similar (McKeown et al., 2019b).

### Social Validity

While it is reassuring teachers believed the professional learning was worthwhile and the intervention was effective, a critical finding from the focus group is that teachers reported they were spending an inordinate amount of time on testing and/

or test preparation required by their administration, to the detriment of academic instruction. Every teacher reported being deeply concerned about the emphasis on testing, and the related loss of agency in terms of curriculum and classroom scheduling. This finding supports those of others in other content areas (e.g., LaRusso et al., 2016). The culture of testing has taken hold at the national level, so interventionists and researchers must find ways to deliver effective instructional practices that also serve these purposes. One shift in practice may be adjusting professional development and related interventions to achieve quick improvement of high-stakes testing skills and outcomes. This is the reality of the schools we serve.

## *6.2 Student Outcomes*

Despite the wide variation in teacher fidelity and the relatively low average fidelity rating, student results from this study were positive, though not statistically significant. We witnessed growth across two genres as a result of the intervention and in both genre elements and holistic writing quality. Yet, caution must be exercised when interpreting these results, as the students included in the study sample are students who returned a parent permission form, signed the student assent form, and were present for both pre- and post-testing. These students may be different from their non-participating peers in ways we cannot identify. There was a slightly larger percentage of non-participants (2%) who were diagnosed with a learning disability though a higher percentage of participants were receiving early intervention for academics. The small sample size limits our ability to generalize. However, these positive student results do give reason to believe SRSD for persuasive and informational writing can be effective in classwide implementation in urban schools. Teacher fidelity varied both between teachers and within teacher, which adds noise to measures of student learning and outcomes.

## *6.3 Science in Challenging Settings*

The story of this study may not be in the fidelity nor in the improvements in student performance, but rather in difficulty of scaling up evidence-based practices into complicated educational settings. The research team involved in this study was experienced, well-supported, and well-funded. We had been trained at leading research universities and completed multiple previous large trials successfully. Thus, even with a well-practiced machine behind us, even with high-quality technical implementation, external factors interrupted our implementation of this study.

### Constrained Effects

There were high rates of transfers and/or absenteeism in the school. The research team made at least three efforts to obtain make-up tests for absent students. Approximately one-fourth of all consented and assented students completed pre- and post-test within a genre. The school possessed a highly mobile student population. A mobile student population affects more than urban, public schools, but can also be observed in schools that serve students with parents who are migrant workers, homeless students, students in foster care, and indigenous students. As our effective sample size shrank due to student mobility, it also reduced our power to detect an effect.

As evidence-based practices move from controlled settings to implementation in high-poverty, high-mobility school settings and more variables that impact performance are introduced, it is expected that effects may be reduced (Raudenbush, 2007; Raudenbush et al., 2007). Cohen et al. (2003) suggest that effects are moderated by the characteristics unique to the setting; thus, as the settings introduce more variables that can impact implementation, outcomes will likely diminish. Raudenbush (2007) suggests baseline effect sizes in researcher-controlled environments represent the upper limit of the intervention's impact. The impacts of poverty, mobility, teacher commitment, and alignment of school and research goals may be factors that reduce student outcomes as evidence-based practices are implemented in more authentic settings (Baker, 2007).

The results of this study should be interpreted with caution. Interventions implemented with lower fidelity may not represent the full potential of the intervention. Effects due to the unique school environment cannot be separated from the effects of the intervention itself. This study was carried out in a high poverty, high mobility school located in a metropolitan area in the Southeast. Teachers rated the intervention with high social validity; still, student outcomes, though positive, were not significant. This is attributable to many things, but in part, a righteous resistance to the demands of rigorous science in otherwise fragile settings. In the end, despite improvements in student writing outcomes, the sample was too small and there was too much variability to pull in the standard deviations around that effect to deem it significant.

### Righteous Resistance

The five principals in the school cluster and the research team sincerely wanted to engage with evidence-based practices to support students who need high quality instruction in literacy. However, the requirements of rigorous science and the mandated timelines for school improvement were at odds with one another causing three schools to drop out of the study before it began. After another experimental design was conceived with remaining participants, teachers engaged in righteous resistance to the implementation due to the numerous other demands they were under as

schools faced radical reorganization from state and national performance mandates. This caused yet another school to opt out of the study and left the researchers with a third, weaker study design. It is imperative, as researchers, we recognize the practical and ethical principles of conducting research in more fragile settings by balancing tight experimental control, social validity, and the ability to address the other demands stakeholders are expected to meet (Raudenbush, 2007). Beyond the experimental design, other factors negatively impacted implementation such as teacher absenteeism during professional development and student mobility and absenteeism during instruction.

As interventions move from efficacy studies to authentic implementation, researchers must accept having less control over experimental conditions and release some degree of responsibility in decision-making (Kennedy, 2005). Within both teacher interviews and our observations, we noted ongoing mandated district and state testing as well as school, district, and state level pre-testing to assess students' readiness for upcoming assessments. Teachers were met with many competing priorities and had to decide for themselves which would have to be discarded as the school, district, and state that would normally do that for them were actually the entities creating the demands. In this case, an evidence-based intervention for literacy with the highest effect size across multiple meta-analyses was the priority that fell below the threshold for inclusion in favor of test preparation.

## 7 Future Directions

Based on the results of this study, researchers must engage with implementation science principles to better understand and facilitate effective implementation of evidence-based practices in complex school settings. Every actor throughout this study wanted to improve the lives of children and their actions were rational within their context. This means that researchers must find ways to make participation in high quality research more acceptable and appealing to schools, teachers, students, and parents.

While the requirements of the research study were clearly defined for the school administrators and explained to teachers during professional learning, four of the five schools withdrew because of research-related demands which they viewed as having a larger net negative impact than the potential gains to be realized. It may help future researchers to approach schools with an array of acceptable research designs to take a more collaborative approach. It may also help if teachers participated in co-designing the experiments to increase teacher buy-in and also improve the fit of the study to the local context. Of course, where there are too many competing demands, researchers must rely on administrators and teachers to set priorities in their own schools.

When conducting research in school settings, especially high need schools, researchers will be well-served by planning for diluted effects and high attrition. As such, they should size the research accordingly, when possible.

Fidelity of implementation varied and was relatively low in comparison with other studies of SRSD, but such variability has been seen in other studies in which interventions were implemented in urban school settings (e.g., Nese et al., 2016). This suggests that researchers may want to implement in a wide spectrum of school contexts to try to differentiate the effects of the intervention from those of the school environment. Comparing urban, public schools under state mandates with other schools can create a new list of student, teacher, and instructional confounds, but may help researchers describe the range of implementation quality that is possible in diverse school environments.

## References

Baker, E. L. (2007). Principles for scaling up: Choosing, measuring effects, and promoting widespread use of educational innovation. In B. Schneider & S. McDonald (Eds.), *Scale-up in education: Ideas in principle* (Vol. I, pp. 37–54). Rowman and Littlefield.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond & G. Sykes (Eds.), *Teaching as a learning profession: Handbook for policy & practice* (pp. 3–31). Jossey-Boss.

Balu, R., & Doolittle, F. (2016). Commentary: Learning from variations in fidelity of implementation. *New Directions for Child and Adolescent Development, 154*, 105–108. https://doi.org/10.1002/cad.20173

Brindle, M., Graham, S., Harris, K. R., & Hebert, M. (2016). Third and fourth grade teacher's classroom practices in writing: A national survey. *Reading and Writing, 29*, 929–954. https://doi.org/10.1007/s11145-015-9604-x

Center for Technology and Education. (2016). *Characteristics of urban education*. http://cte.jhu.edu/urbaneducation/ses1_act2_pag1.shtml

Cohen, D. K., Raudenbush, S. W., & Loewenberg Ball, D. (2003). Resources, instruction, and research. *Educational Evaluation and Policy Analysis, 25*(2), 119–142. https://doi.org/10.3102/01623737025002119

Cordray, D. (2007, June 17–29). *Fidelity of intervention implementation* (Power Point Slides). IES Summer Training Institute on Cluster Randomized Control Trials, Vanderbilt University.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3/4), 327–350. https://doi.org/10.1007/s10464-008-9165-0

FitzPatrick, E., & McKeown, D. (2020). Meeting the needs of middle school writers in a special education classroom: SRSD for the informational genre citing text-based evidence. *Education and Treatment of Children, 43*, 71–84. https://doi.org/10.1007/s43494-020-00006-2

Foorman, B., Dombek, J., & Smith, K. (2016). Seven elements important to successful implementation of early literacy intervention. *New Directions for Child and Adolescent Development, 154*, 49–65. https://doi.org/10.1002/cad.20178

Gersten, R., Morvant, M., & Brengelman, S. (1995). Close to the classroom is close to the bone: Coaching as a means to translate research into classroom practice. *Exceptional Children, 62*(1), 52–66.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Graham, S., Harris, K. R., & Hebert, M. (2011). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children, 44*(4), 1–12. https://doi.org/10.17161/foec.v44i4.6687

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879–896. https://doi.org/10.1037/a0029185

Graham, S., Harris, K. R., & McKeown, D. (2013). The writing of students with LD and a meta-analysis of SRSD writing intervention studies: Redux. In L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities*. Guilford Press.

Harris, K. R., Graham, S., Mason, L. H., & Friedlander, B. (2008). *Powerful writing strategies for all students.* Paul H. Brookes Publishing Co., Inc.

Harris, K. R., Lane, K. L., Graham, S., Driscoll, S. A., Sandmel, K., Brindle, M., & Schatschneider, C. (2012). Practice-based professional development for SRSD in writing: A randomized controlled study. *Journal of Teacher Education, 63*, 103–119. https://doi.org/10.1177/0022487111429005

Johnson, L. D., Wehby, J. H., Symons, F. J., Moore, T. C., Maggin, D. M., & Sutherland, K. S. (2013). An analysis of preference relative to teacher implementation of intervention. *The Journal of Special Education, 48*, 214–224. https://doi.org/10.1177/0022466913475872

Kennedy, M. M. (2005). *Inside teaching: How classroom life undermines reform*. Harvard University Press.

LaRusso, M., Donovan, S., & Snow, C. (2016). Implementation challenges for tier one and tier two school-based programs for early adolescents. *New Directions for Child and Adolescent Development, 154*, 11–30. https://doi.org/10.1002/cad.20179

McKeown, D., FitzPatrick, E., & Sandmel, K. (2014). SRSD in practice: Creating a professional development experience for teachers to meet the writing needs of students with EBD. *Behavioral Disorders, 40*(1), 15–25. https://doi.org/10.17988/0198-7429-40.1.15

McKeown, D., Brindle, M., Harris, K. R., Graham, S., & Collins, A. A. (2016). Illuminating growth and struggles using mixed methods: Practice-based professional development and coaching for differentiating SRSD instruction in writing. *Reading and Writing: An Interdisciplinary Journal, 29*, 1105–1140. https://doi.org/10.1007/s11145-016-9627-y

McKeown, D., Brindle, M., Harris, K. R., Sandmel, K., Steinbrecher, T., Graham, S., Lane, K., & Oakes, W. (2019a). Teachers' voices: Understanding effective practice-based professional development for elementary teachers on SRSD in writing. *American Educational Research Journal, 56*, 753–791. https://doi.org/10.3102/0002831218804146

McKeown, D., FitzPatrick, E., Brown, M., Brindle, M., Owens, J., & Hendrick, R. (2019b). Urban teachers' SRSD implementation following PBPD: Positive effects mediated by compromised fidelity. *Reading and Writing, 32*, 1483–1506. https://doi.org/10.1007/s11145-018-9864-3

National Center for Education Statistics. (2011). *National assessment of educational progress: Writing assessment*. Institute of Education Sciences, U.S. Department of Education.

Nese, R., McIntosh, K., Nese, J., Hoselton, R., Bloom, J., Johnson, N., Richter, M., Phillips, D., & Ghemraoui, A. (2016). Predicting abandonment of school-wide positive behavioral interventions and supports. *Behavioral Disorders, 42*(1), 261–270. https://doi.org/10.17988/BD-15-95.1

Raudenbush, S. W. (2007). Designing field trials of educational innovations. In B. Schneider & S. McDonald (Eds.), *Scale-up in education: Issues in practice* (Vol. II, pp. 23–40). Rowland and Littlefield.

Raudenbush, S. W., Martinez, A., & Spybrook, J. (2007). Strategies for improving precision in group-randomized experiments. *Educational Evaluation and Policy Analysis, 29*, 5–29. https://doi.org/10.3102/0162373707299460

What Works Clearinghouse. (2012). *Teaching elementary school students to be effective writers: Educator's practice guide*. http://ies.ed.gov/ncee/wwc/practiceguide.aspx?sid=17

# Using Generalizability Theory to Explore Sources of Variance on an Observation Tool of Writing Instruction


**Alyson A. Collins, Stephen Ciullo, and Micheal P. Sandbank**

**Abstract** Observations of classroom practice provide guidance in identifying areas for teacher professional development and adjustments to instructional approaches. The present study used generalizability theory to investigate the dependability of scores from an observation tool designed to measure writing instruction in upper elementary classrooms. A total of 25 fourth and fifth grade general education teachers were observed during classroom writing lessons three times across a school year. The Writing Instruction Observation Protocol (WIOP) documented lesson components within two scales: (a) instructional practices, and (b) quality of implementation. G studies of both scales estimated variance across teachers and observation occasions to investigate the proportion of variance accounted for by each measurement facet. When comparing the two scales, variance estimating differences in teachers was larger for the quality of implementation scale (35.4%) when compared to person variance for total instructional practices (16.6%). For the two scales, observation occasion accounted for only a small portion of variance (7.8% and 12.4%), and more than half of variance was attributable to unmeasured facets (52.2% and 75.6%). D studies further estimated the number of observations necessary to obtain a dependable estimate on the two WIOP scales when using an average of scores from 1 to 10. Results indicated at least 8 and more than 10 observations may be needed to obtain dependable scores identifying differences among teacher instructional practices and quality of implementation. Findings suggest examination of other sources of variance may need investigation in future research on writing observation tools.

**Keywords** Writing instruction · Observation · Reliability · Elementary · Generalizability theory

A. A. Collins (✉) · S. Ciullo
Department of Curriculum and Instruction, College of Education, Texas State University, San Marcos, TX, USA
e-mail: alysonacollins@txstate.edu

M. P. Sandbank
University of North Carolina at Chapel Hill, Chapel Hill, NC, USA



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_15

## 1 Using Generalizability Theory to Explore Sources of Variance on an Observation Tool of Writing Instruction

Education stakeholders continuously strive to improve classroom literacy instruction. These key education stakeholders who influence literacy instruction in schools include pre-service teacher faculty, school administrators, professional development providers, educational researchers, and classroom teachers. Most stakeholders focus on analyzing student performance data to evaluate the effectiveness of literacy instruction. Because improving student outcomes serves as the primary and ultimate goal for education stakeholders, improving teacher instructional practices may serve as a pathway towards this goal. To this end, observations of typical, everyday classroom instruction can inform decisions about effective literacy instructional practices and professional development needs, reforms to teacher training programs, and directions for future research (Durkin, 1978).

As such, observation tools provide value in educational settings, yet the purposes of such tools and constructs they are meant to capture vary widely. Some classroom observation systems provide teachers with specific feedback to help them immediately adapt and adjust their daily instruction (e.g., Biancarosa et al., 2020; Pianta & Hamre, 2009). Other tools document whether teachers use empirically-sound practices in the classroom, with the aim of guiding future professional development considerations (e.g., Swanson, 2008). In either case, data derived from these tools can help stakeholders identify areas of strength in instructional practices and areas in need of instructional improvement (Connor et al., 2004; Durkin, 1978, 1984). Thus, these tools are critical to identifying professional development needs and ultimately improving the outcomes of students (Kelcey & Carlisle, 2013).

In order to collect useful information about teacher practices that can improve student literacy outcomes, educational stakeholders need access to reliable and valid observation tools that specifically focus on writing instruction. A challenge, however, is that most observation tools focus on measuring literacy instruction in reading (e.g., Automated Classroom Observation System for Reading [ACOS-R], Kelcey & Carlisle, 2013; Instructional Content Emphasis-Revised [ICE-R], Edmonds & Briggs, 2003). Despite increased attention to the need for integrated literacy instruction in Common Core and state initiatives (Graham & Harris, 2015; Troia & Olinghouse, 2013), few observation tools solely focus on writing-specific instruction (e.g., Coker et al., 2018a, b), and teacher self-report surveys (e.g., Graham et al., 2014) account for the majority of information currently available about typical writing instructional practices. Even less research exists examining the reliability and validity of the few writing instruction observation tools available.

Given few tools exist, foundational studies of literacy observation tools have aimed to establish the reliability and validity of the measures to determine if specific tools represent the nature of teacher behaviors and student interactions. For instance, Johnson et al. (2020) estimated reliability on a special education teacher

observation tool called the Recognizing Effective Special Education Teachers (RESET) tool. Analyses revealed that raters were generally consistent when scoring low-inference items, but rater stability was more varied on high-inference items. Other studies have established reliability among observers when comparing coding between two lessons to explore agreement among raters and variance across items such as modeling, practice opportunities, and feedback (Smolkowski & Gunn, 2012). These studies suggest reliable and valid literacy observation tools can be developed to capture teacher instruction and student interactions, while also underscoring the need for investigations of reliability and validity as new tools are developed.

When investigating the reliability and validity of writing instruction observation tools, it is important to identify common variations in measurement contexts that can function as sources measurement error (e.g., Johnson et al., 2019, 2020; Smolkowski & Gunn, 2012). Although reliability investigations tend to focus on measurement error contributed by items or raters (i.e., by documenting interitem or inter-rater reliability), other aspects of measurement contexts, such as occasion, can function as sources of measurement error in scores derived from observation tools (e.g., Semmelroth & Johnson, 2014). In other words, day-to-day variations in teacher instructional practices can contribute error to scores derived from these tools, reducing the extent to which such scores are truly representative of teachers' overall writing instruction quality. Unreliable and unrepresentative scores increase the likelihood of Type II statistical error, potentially preventing the detection of associations between teacher practices and student achievement even when such links truly exist (Doabler et al., 2021; Smolkowski & Gunn, 2012). Consequently, there is a need for the development and evaluation of observation tools of writing instruction that capture the components of effective instruction and quality of teacher implementation. The purpose of this study was to investigate sources of measurement error and dependability of an observation tool used to capture writing instructional practices and quality of implementation of general education teachers in Grades 4 and 5.

## *1.1 Observations of Classroom Writing Instruction: What We Know and Why It Matters*

The importance of exploring writing instruction by using systematic observation tools, and the basis for the present chapter is substantiated by the need to enhance teachers' use of effective instructional practices in writing (Coker et al., 2018a). National writing achievement data consistently shows many students could benefit from additional support with written expression. For decades, the majority of students in the United States (including Grades 4, 8, and 12) have not attained a proficient score on the National Assessment of Education Progress in Writing assessment

(National Center for Education Statistics, 2012). Additional disparities and inequities in writing performance exist between students with and without learning disabilities (Graham & Perin, 2007). Thus, collecting classroom observation data may help educators determine the extent to which teachers utilize effective instructional practices and ensure teachers are addressing the writing needs of a range of learners.

Recent research examining observations of classroom writing instruction suggests increasing practice opportunities for generating and composing connected text positively enhances student writing achievement, particularly for boys in first grade classrooms (Coker et al., 2018a). Moreover, systematic teacher observations indicate practice opportunities of generative writing mediate effects of writing instruction on student reading achievement of young students (Coker et al., 2018b). Other findings from systematic observations, however, document negative and null effects of writing instruction on student achievement (Coker et al., 2018a). Mixed findings suggest further exploration is needed on the extent to which elementary teachers use components of effective writing instruction and quality with which they implement them with students. Additional research to understand current classroom practice could guide professional development providers in developing high-quality training to support teachers in how to teach writing.

### *1.2 The Writing Instruction Observation Protocol (WIOP): A Writing-Specific Observation Tool*

The Writing Instruction Observation Protocol (WIOP) is an observation tool (and the focus of the current study) that was developed to extend research on classroom writing instruction and document teachers' use of effective instructional practices (Kotula et al., 2014). The WIOP was initially developed and piloted over two years in 259 fourth- and fifth-grade classrooms to identify teacher practices used in 'business as usual' writing instruction to compare typical practices to a K-12 writing curriculum. The WIOP contains two scales measuring the total instructional practices and quality of implementation of those practices. All items associated with effective writing instruction were developed from findings from systematic reviews of writing research (Graham & Perin, 2007). Instructional practice items specifically identify the presence or absence of instructional practices such as stating lesson objectives, providing examples of written text, giving feedback to students, and engaging in other practices associated with improved writing achievement. A unique feature of the WIOP is the inclusion of a scale that evaluates quality of implementation of instructional practices because few available measures contrast quality with instructional practices. Collectively, the WIOP offers a comprehensive tool to compare instruction on skills, strategies, and composing along with quality of implementation.

One limitation of the WIOP, however, is the limited information available regarding the reliability and validity of the instrument. Information presently available on the tool pertains to percentage of inter-rater agreement on dichotomous items (e.g., Did the teacher activate prior knowledge relevant to today's writing session? Yes or No), with evidence observer agreement may range from 89% to 96% (Kotula et al., 2014). Some research also suggests observers can achieve greater than 80% inter-rater agreement on the quality of implementation scale, and internal consistency of individual quality ratings versus overall effectiveness is estimated as 0.87 (Kotula et al., 2014). What remains unknown is how reliably the WIOP documents differences in teacher use of effective writing instructional practices that is attributable to true differences between teachers, or the number of observations necessary to derive dependable estimates of teachers' typical writing instruction practices.

## 2 How Generalizability Theory Informs Observation Research

Generalizability theory (Cronbach et al., 1963), often referred to as G theory, offers an empirical framework that researchers can use to better understand sources of variance that contribute measurement error to scores derived from observation tools. Dependability studies (D studies) use variance estimates from generalizability (G) studies to extrapolate the number of observations which are necessary to obtain an average score quantifying a given construct (e.g., an instructional practice, student responses) that is dependable for making either relative or absolute decisions. D studies yield two coefficients: (a) the generalizability (G) coefficient, which reflects the dependability of a score for making relative decisions (such as ranking), and (b) the dependability ($\phi$) coefficient, which reflects the dependability of a score for making absolute decisions (such as whether a score meets or surpasses a prespecified cutoff). For example, researchers have used G and D Studies to explore the number of observation occasions are necessary to obtain dependable estimates of verbal behaviors of teachers (Hollo et al., 2020). Using G studies, Hollo et al. (2020) determined that occasion attributed relatively small amounts of error variance to measures of "teacher talk". For some measures (such as mean length of utterance in words, which indexes sentence complexity), a substantial portion of error variance was attributable to unmeasured facets of observation (e.g., time of day, content of the lesson), and a relatively small proportion of variance was attributable to true differences in teacher behavior (i.e., what the scores are purported to measure). For other measures (such as total words and total utterances, measures of quantity of talk), most variance was attributable to the person facet, meaning that scores from those measures were more dependable and representative of real differences in teacher behavior. Hollo and colleagues then conducted a D study to extrapolate from variance component estimates the number of observations needed to obtain dependable average scores of teacher talk. Results indicated that the total number of

words a teacher used in one 10-min observation would yield a dependable score of quantity of teacher talk. However, for measures which were less dependable, such as mean length of utterance in words, an average of 4 scores taken across 4 different days was needed to derive a dependable score of the complexity of teacher talk.

In the current study, we applied G theory to investigate sources of error variance and dependability of the WIOP for observing writing instruction in fourth and fifth-grade inclusive classrooms. Our research aims were to estimate the proportion of variance in WIOP scores attributable to true differences in the instructional practices of Grade 4 and 5 general education teachers (i.e., true variance) in comparison to the proportion of variance attributable to differences in occasion or other sources of error (i.e., error variance). In addition, we sought to examine the dependability of scores taken from a single observation and estimate the number of observations needed to obtain a dependable mean score for a given teacher. Thus, we applied G theory to investigate the following research questions:

1. What is the proportion of variance attributable to differences among teachers and occasion (e.g., time of year the observation was conducted), as well as interactions between these facets and other unmeasured sources of error, for the WIOP total instructional practices and quality of implementation scales?
2. What is the projected dependability of scores from the WIOP total instructional practices and quality of implementation scales across an average of increasing numbers of observations from 1 to 10?

Based on previous evidence that the WIOP observation tool afforded a reliable and valid observation tool of instructional measurements (Kotula et al., 2014), we expected a large portion of the variance would be attributable to true differences in individual teacher practices (i.e., persons). A contrasting finding of larger error variance in comparison to true person variance would suggest outside sources of error may introduce confounds into the measurement of classroom writing instruction. In the latter case, the number of observations required to obtain a dependable observation rating may preclude feasible and realistic research methods if a large number of observations are necessary to obtain stable measures.

## 3 Method

### *3.1 Participants*

We recruited the teacher sample from a central Texas school district comprised of six elementary schools. Principals from five of the six schools agreed to participate. District demographics during data collection represented a diverse population. Race and ethnicity across the district reported enrollment as: 4.8% African American, 72.2 % Hispanic, 21% White, 0.1% American Indian, 0.8% Asian, 0.1% Pacific Islander, and 1.0% two or more races. Approximately 71% of the student population received free or reduced lunch. Only 9% of students were identified as English

Learners (ELs), and 11% received special education services. Mobility rates for the district were slightly higher than the state average at 18.2%. Average class sizes for elementary Grades 4 and 5 were 21 and 30, respectively. Although the Grade 4 class sizes were comparable to the state means, the Grade 5 class sizes were substantially larger than the state average of 21.

Teachers who participated in the observations included 25 general education teachers who provide writing instruction to students in fourth and fifth grade. Specifically, 17 taught fourth grade and 8 were fifth grade. The large majority of teachers were female ($n = 20$). Of the 25 teachers, 8 identified as Hispanic, 15 were white, and 2 reported race/ethnicity as other groups. Teachers had taught on average 9.28 years ($SD = 7.28$), with years of experience ranging from 0 to 27. Twelver teachers (48% of the sample) held masters degrees. The average class size among teachers was reported as 21 students, ranging from 17 to 24. Approximately 3 students per class received special education services, with teachers reporting 0–10 students on a given roster. On average, 3 students per class were ELs, although one class reported the entire group of students were ELs ($n = 22$).

To contextualize writing lessons within English Language Arts (ELA) instruction, teachers reported additional information on a brief instructional survey about time dedicated to writing instruction. Average estimates indicated teacher devoted 209 min per week ($SD = 124.59$; range: 17.5–450) to teaching writing across a mean number of 3.76 lessons per week ($SD = 1.82$; range: 1–10). In these classrooms, teachers reported that students wrote on average 48.94 min a week ($SD = 52.03$; range: 5–275).

## *3.2 Measure of Writing Observation*

We documented writing practices of teachers in Grade 4 and 5 using the Writing Instruction Observation Protocol (WIOP; Kotula et al., 2014). The observation measure examines elements of effective writing instruction, including components of explicit instruction, student scaffolding, and integration of stages within the writing process. It includes two scales, one which documents the inclusion of specific instructional practices within a lesson and a second scale that estimates quality of implementation.

The first scale, referred to as *instructional practices,* includes a total of 22 items. Observers indicate the presence or absence of a given practice within the entire lesson with *observed* coded as 1 and *not observed* coded as 0. Individual items are categorized into four subgroups: (a) introduction; (b) skills, strategies instruction, and practice; (c) composing; and (d) closure. We provide general descriptions of observed teacher practices in each category in this section. A full copy of all items are available from Kotula et al. (2014).

Specifically, three items on the WIOP pertain to *Introduction* and capture teacher's verbally stating the lesson objective and activation of prior knowledge in relation to previous lessons and the current writing tasks. Items related to *Skills,*

*Strategies Instruction, and Practice* then examine how students engage in skills and strategy learning (e.g., "*Explicitly explains how previous lessons relate to today's writing session"*). A total of 11 items document whether teachers provide explicit examples, modeling, and practice to target a particular skill or strategy. *Instructional* items also explore the mode of student and teacher interactions, including the use of discussion, questioning, and feedback on lesson tasks. Example items include "*Provide direct skill/strategy instruction,*" and "*Engage the students in whole class discussion.*" Next, *Composing* items include 6 items with most comparable to the items embedded in *Skills, Strategies Instruction, and Practice.* The distinguishing difference between the two categories of items pertain to the purposes for which the practices are being utilized (i.e., to compose or learn a skill/strategy). Example *Composing* items include "*Provides an example of writing related to the lesson on composition,*" and "*Teacher gives feedback on student writing in different formats (may include individually, small group, whole class, written, verbal).*" For the purposes of our observations, we defined composing as instruction for the purpose of writing connected text with two or more sentences. This definition is comparable to other observation tools documenting student composing opportunities (e.g., generative writing; Coker et al., 2018a). In cases when lessons included practices for multiple purposes (i.e., instruction including strategies and composing), observers marked the items in each respective category. For example, during a lesson teaching a mnemonic for planning an essay, a teacher may provide students opportunities to practice using the strategy for planning. Teachers may also model how to use the strategy within the process of composing an essay. In this scenario, items were coded in each respective section of *Skills, Strategies Instruction, and Practice* (i.e., provides students with time, in class, to practice the focus skills/strategies) and *Composing* (i.e., models the relevant process of writing). The final *Closure* section includes 2 items to identify if the teacher provides closure to the lesson and if homework is assigned. Additional items descriptively capture if other adults were present during the lesson and open-ended comments/notes about the observed lesson.

The second WIOP scale, referred to as *quality of implementation*, measures implementation quality by requiring observers to rate teacher effectiveness of 10 different instructional categories. One general category measures overall lesson effectiveness on a 4-point scale, with lower ratings representing *not effective* (1) and higher scores as *very effective* (4). The remaining nine categories apply the same scale from 1 to 4, with an additional score to indicate when specific categories are *not observed* (0). These categories estimate the effectiveness of categories documented on the instructional practices including: (a) lesson objective, (b) activation/use of prior knowledge, (c) direct skill/strategy instruction, (d) models and/or modeling, (e) guided practice, and (f) feedback. The remaining 3 of the 10 categories relate to overall lesson components, all measured on the same 0 to 4 scale. One item evaluated the *academic challenge* of the lesson, in which raters considered the rigor of activities and topics within the writing instruction as well as the teacher's expectations for students. Another rated the *lesson integrity*, defined as following and staying true to the intended objective from the beginning to the end of the lesson to ensure instructional goals are met. Finally, observers appraised the *classroom*

*climate*, considering rapport between teachers and students, the enthusiasm and tone of the lesson, and the value placed on diverse perspectives.

An initial two-year pilot study examined inter-rater agreement among two and three observers within a given lesson. Reliability ranged from 89% to 96% for this protocol within the randomly selected classrooms (Kotula et al., 2014). In the initial pilot, inter-rater agreement was greater than 80% accuracy on the quality of implementation scale, and internal consistency of the individual quality ratings versus overall effectiveness was reported as 0.87 (Kotula et al., 2014). In the current study, inter-rater agreement was 91% among the four observers for the total instructional practice items and also 91% for the quality of implementation scale.

An initial two-year pilot study examined inter-rater agreement among two and three observers within a given lesson. Reliability ranged from 89% to 96% for this protocol within the randomly selected classrooms (Kotula et al., 2014). In the initial pilot, inter-rater agreement was greater than 80% inter-rater agreement on the quality of implementation scale, and internal consistency of the individual quality ratings versus overall effectiveness was reported as 0.87 (Kotula et al., 2014). In the current study, inter-rater agreement was 91% among the four observers for the total instructional practice items and also 91% for the quality of implementation scale.

### *3.3 Procedures*

Researchers observed three writing lessons for each of the 25 teachers across the school year (i.e., October to March). Observations were scheduled during the block designated as 'writing time' within each teacher's classroom schedule. Two observers independently observed 20% of the writing lessons to establish inter-rater agreement on the instructional practice items and reliability on the quality of implementation scale. Initially, inter-rater agreement for the two WIOP scales was established through observations of pre-recorded writing lessons. Training continued until a minimum of 80% inter-rater agreement was established.

For the two WIOP scales, we estimated inter-rater agreement as AGREEMENTS/(AGREEMENTS + DISAGREEMENTS). We defined an agreement on the instructional practices scale as the same coding decision for *observed* (1) and *not observed* (0) within a given writing lesson. Agreements on the quality of implementation scale were calculated as coding as the same quality rating (e.g., both observers coded the quality of implementation as a 3) or ratings ±1 within the scale (e.g., one observer coded quality as a 1 and the second observer coded it as a 2).

When two observers independently double-coded a writing lesson, observers debriefed after the observation to discuss the lesson codes and resolve any identified discrepancies. In addition, all writing lessons were audio recorded as a resource for resolving discrepancies. Inter-rater agreement estimates reported for the current study represent independent coding before discrepancies were resolved, whereas the final dataset used for the current analyses accounted for all discrepancy resolutions.

### *3.4 Data Analysis*

First, we calculated descriptive statistics for teachers on the WIOP observation scales by grade level and total sample. We ran two separate random effects multilevel models to estimate if the teacher observation scores differed by grade level for each of the two WIOP scales: (a) instructional practices, and (b) quality of implementation (Snijders & Bosker, 2011). The multilevel models account for nesting of observations (i.e., observation 1, 2, and 3) at Level 1 within teachers at Level 2. A dummy variable comparing Grade 4 (coded as 0) and Grade 5 (coded as 1) estimated differences between grades. The two models were run in Stata 14.2 using the 'mixed' command (StataCorp, 2015).

Next, we parsed variance in WIOP total instructional practice and quality of implementation scores using a fully crossed G studies with facets of person (i.e., teacher) and occasion (i.e., observation 1, 2, and 3) and their interaction. Our samples sizes of 25 teachers and 3 observation occasions align with Webb and colleagues' (1988) recommendations that G studies should include a minimum of 20 participants with at least two measurements per person. We treated teacher and observation facets as fully crossed because all levels were crossed with all other facets (i.e., we collected 3 observation occasions for each teacher; Webb et al., 2006). Furthermore, we treated all facets within the G studies as infinite random. This application assumes scores in each facet are sampled from an infinite universe of all possible teachers and observations and scores are allowed to vary randomly within the assumed universe. Applying this approach within the G Study analysis affords a more conservative estimates of dependability. Moreover, treating facets as infinite and random allows estimates to be generalized beyond the actual sample (Shavelson & Webb, 1991).

To answer our second research question, we conducted subsequent D studies for scores from the two WIOP scales (total instructional practices and quality of implementation). The D studies projected the dependability that could be achieved if scores were averaged across additional observation occasions. We estimated dependability for WIOP scores up to 10 occasions. All G and D Study analyses were conducted in R statistics program (R Core Team, 2022) using the package *gtheory* (Moore, 2016).

## 4 Results

Table 1 reports descriptive statistics for the WIOP scales by grade level as well as the total sample of general education teachers ($N = 25$) across the 75 writing lessons (i.e., 3 observations per teacher). Teachers' writing lessons averaged 52 mins across all observations ($SD = 13.08$), with the shortest lesson lasting 27 min and the longest duration as 85 min. The correlation between the instructional practices and quality of implementation scales was 0.75 ($p < 0.001$). Results of the multilevel

models indicated no statistically significant differences between teacher observations by grade for the instructional practices ($B = -1.04$; $p = 0.30$) and the quality of implementation scales ($B = -3.34$; $p = 0.18$).

Table 2 reports results for our first research question partitioning variance across teacher (i.e., person) and observation occasion (i.e., time point 1, 2, or 3). G Study results indicated 16.6% of variance in the WIOP total instructional practices scale was attributable to true differences in teacher practices (i.e., true person variance), and observation occasion accounted for only 7.8% of the variance. By contrast, residual accounted for 75.6% of the variance, suggesting the majority of variance was attributed to the interaction between the person and occasion facets, as well as

**Table 1** Descriptives for scales from the WIOP observation

| WIOP Scale | *M* | *SD* | Min | Max |
|---|---|---|---|---|
| Total nstructional practices | | | | |
| Grade 4[a] | 11.66 | 3.64 | 4 | 20 |
| Grade 5[b] | 10.62 | 3.30 | 4 | 17 |
| Total sample | 11.37 | 3.56 | 4 | 20 |
| Quality of implementation | | | | |
| Grade 4[a] | 22.62 | 7.72 | 6 | 37 |
| Grade 5[b] | 19.29 | 7.79 | 8 | 36 |
| Total sample | 21.69 | 7.84 | 6 | 37 |

*Note.* Total instructional practices included 22 items coded as 1 as *observed* and 0 as *not observed* for a maximum score of 22. Quality of implementation included 10 items on a 5-point for a maximum score of 40.
[a]Grade 4 $n = 18$
[b]Grade 5 $n = 7$

**Table 2** Mean square and percentage of variance for G studies of WIOP observation scales

| Source of variance | *df* | MS | Var (%) |
|---|---|---|---|
| Total instructional practices | | | |
| Teacher | 24 | 2.16 | 16.6 |
| Observation occasion | 2 | 1.02 | 7.8 |
| Residual | | 9.85 | 75.6 |
| Coef. | | | |
| Relative *G* Coef. | | | 0.40 |
| Absolute phi Coef. | | | 0.35 |
| Quality of implementation | | | |
| Teacher | 24 | 22..88 | 35.4 |
| Observation occasion | 2 | 8.00 | 12.4 |
| Residual | | 22.75 | 52.2 |
| Coef. | | | |
| Relative *G* Coef. | | | 0.18 |
| Absolute phi Coef. | | | 0.17 |

*Note. MS* mean square, *Var.* variance, *Coef.* coefficients

unexplored measurement facets (e.g., time of day, topic of instruction, length of lesson). For the WIOP quality of implementation scale, 35.4% of variance was accounted for by true teacher (i.e., person) variance. As such, one-third of the variance in scores was attributable to the quality of implementation among the writing lessons. Observation occasion contributed 12.4% of the error variance, with 52.2% residual variance remaining from unaccounted measurement error and the interaction between measured facets (i.e., person and occasion).

For our second research question, Figure 1 depicts the projected dependability (phi) coefficients obtained with an average of an increasing number of observation scores for the WIOP scale for: (a) total instructional practices, and (b) quality of implementation. D study results indicated the relative (i.e., G coefficient) and absolute estimates (i.e., ϕ coefficient) of dependability yielded from WIOP total instructional practice scores collected from 3 observations was 0.4 and 0.37, respectively. D study estimates indicated that even an average of WIOP total instructional practice scores from 10 observations would not reach acceptable levels of relative or absolute dependability (G coefficient = 0.69; ϕ coefficient = 0.66). Dependability of WIOP quality scores averaged across 3 observations was slightly higher, with estimated G coefficient of 0.67 and ϕ coefficient of 0.62. To obtain quality scores that had sufficient dependability for relative decisions, an average of scores from 6 observations were needed (G coefficient = 0.8) and to obtain quality scores that had sufficient dependability for absolute decisions, an average of scores from 8 observations were needed (ϕ coefficient = 0.81).

## 5 Discussion

The present study used generalizability theory to investigate the dependability of scores for fourth and fifth grade teachers' instructional practices and quality of implementation as measured by a writing observation tool (WIOP; Kotula et al., 2014). Interestingly, results revealed a moderate correlation between the two scales. Our findings, however, did not reveal significant differences between Grade 4 and 5 teacher observations on the two WIOP scales, suggesting grade level did not contributed to differences in observation scores. G studies parsed variance across teachers and observation occasions to estimate the proportion of variance attributable to each measurement facet. Findings indicated only a small proportion of variance was attributed to observation occasion, suggesting little variance was attributed to conducting observations across different time points of the school year. Relatively little variance was attributable to true differences in teacher instructional practices or quality. For both scales, the majority of variance was attributable to measurement error from unknown facets, suggesting examination of other sources of error variance are needed in future research to identify other important factors affecting scores from writing observation tools. Additional D studies examined the number of observations necessary to obtain dependable estimates on the two WIOP scales across an average of scores from 1 to 10. Results indicated at least 8 and more than

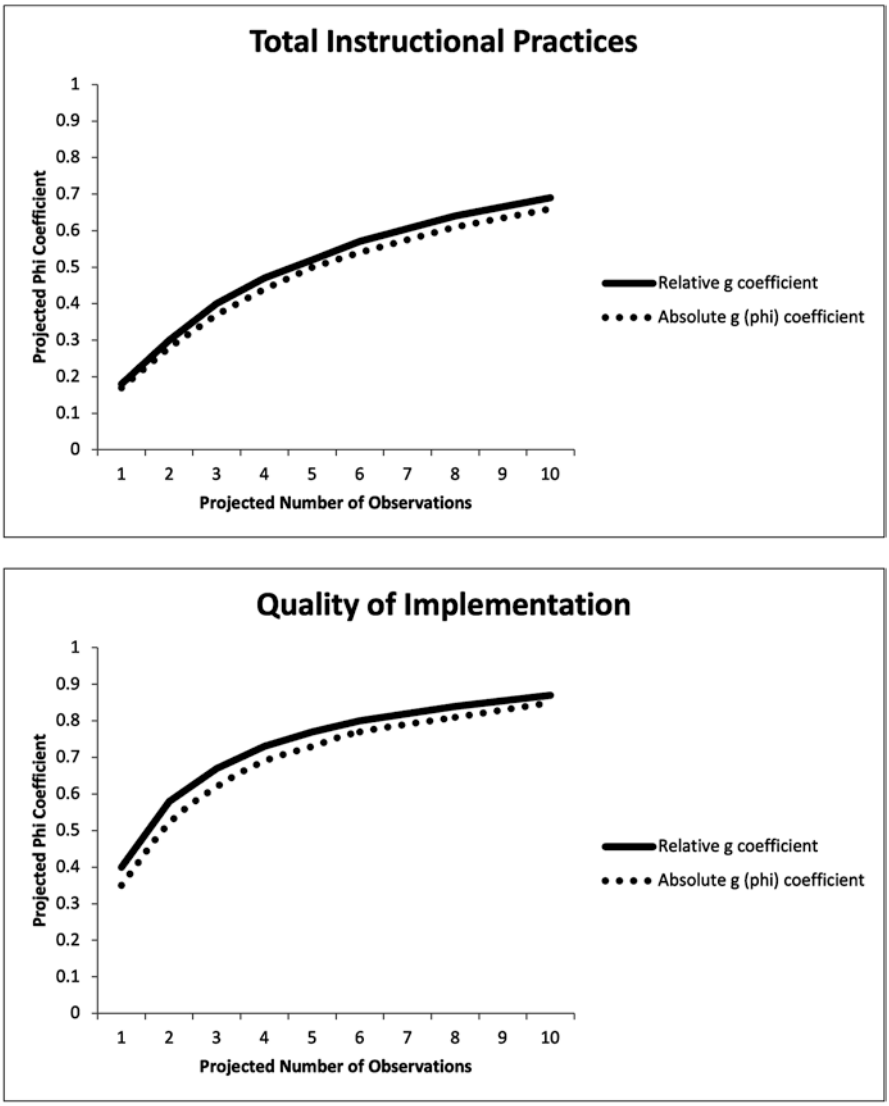



**Fig. 1** Projected dependability (phi) coefficients for the Writing Instruction Observation Protocol (WIOP) scale for Total Instructional Practices and Quality of Implementation. Graphs depict projected dependability (phi) coefficients obtained with an average of an increasing number of observation scores from each WIOP scale for teachers in Grades 4 and 5

10 observations may be needed to obtain dependable scores and ensure differences among teacher instructional practices and quality of implementation are reliably estimated when using the tool. In this section, we provide a deeper discussion of the practical implications of the present study findings and consider potential sources of error to examine in future research on writing observation tools.

An important difference in findings between the instructional practices and quality of implementation scales was the percentage of variance estimated among teachers. Given the WIOP tool aims to document differences in instructional practices and quality of implementation, we expected a reliable and consistent tool to detect subtle differences among teachers' writing lessons. That is, a high percentage of variance attributed to the facet of teachers would affirm that the observation tool was consistently detecting differences among teachers in their use of writing instructional practices and the quality of implementation. Our G Study findings, however, indicated only one-third of the variance on the quality of implementation scale and less than one-fifth of the variance on the instructional practices scale was attributed to differences in teachers. Moreover, differences in occasion accounted for negligible amounts of error variance, which is consistent with prior research suggesting teacher instructional behaviors tend to remain consistent when observed on multiple occasions (Smolkowski & Gunn, 2012). These findings prompt further questions as to what other facets may influence scores on writing instruction observation tools, and future research should consider how to refine current measurement methods.

One consideration in this investigation is none of the WIOP items in the current analysis captured variation in lesson topics or types of writing activities, and this unmeasured facet may potentially explain the large amount of unaccounted for residual variance. It is plausible that different topics of writing lessons may favor use of certain instructional practices over others. For example, components of a lesson on capitalization and punctuation may be quite different than one focused on the process of writing an informational essay. Given prior observation studies suggest the types of writing may predict student progress over time (Coker et al., 2018a, b) and inconsistencies may exist in the writing instruction provide within U.S. classrooms, the topic and focus of the lesson cannot be overlooked when drawing conclusions on if classroom practice is adequate and effective. On the other hand, effective teachers may integrate key components of explicit instruction (e.g., modeling, opportunities for practice) regardless of the topic or activity, as suggested by prior research examining stability of observed teacher behaviors (Smolkowski & Gunn, 2012). Thus, iterative development of writing observation tools is needed to determine the most reliable methods for documenting lesson topics alongside use of instructional practices and quality of implementation. Observation tools capturing all three features of instruction may afford a more comprehensive measure for depicting typical classroom practice, and more importantly for making decisions about teacher professional development.

Another reason for the lack of variance attributable to teachers may relate to the measurement procedures for observing a single writing lesson. Although our G Study revealed the occasion facet (representing different time points of observation) did not contribute to large portions of variance on the two observation scales, our observations were conducted during the time frame teachers indicated they provided writing instruction to students. Consequently, other factors such as the length of the lesson (i.e., time in minutes) and the timing of the lesson during the school day (i.e., morning or afternoon) varied greatly across teachers. More time observing a teacher may provide opportunities to observe more instructional practices and could also influence the perceptions of the quality of implementation. Relations

among these potential facets, however, were not investigated in the present study. It is possible differences in these unmeasured facets may contribute to the large amount of unexplained residual variance in the observation data for both the instructional practices and quality of implementation scales.

Moreover, unmeasured facets such as the length of the lesson may explain the D studies findings indicating approximately 8 observations are needed to establish consistent quality of implementation, and greater than 10 observations may be necessary for a dependable measure of instructional practices. Because time constraints may not allow teachers to engage in every instructional practice in a given lesson and the time observed varied from teacher to teacher, our findings may suggest a certain number of observations are needed to capture all of the possible practices teachers utilize during their writing instruction. For example, if a teacher was observed consistently as having opportunities for student practice without accompanying modeling or examples, the teacher may only receive a score of 1 out of 3 for utilizing only one of the possible practices. Drawing conclusions that a lower score is associated with less effective instruction, however, may inadvertently overlook the possibility that a teacher modeled and provided practices in earlier lessons. Consequently, future research on writing observation tools may consider comparing and contrasting methodological sampling procedures, such as comparisons between observing lessons at random across the school year versus observations conducted on consecutive school days. Future research should also triangulate multiple sources of data, such as teacher surveys and interviews, to gain additional insight on practices teachers utilize during writing instruction. Exploration of the different procedures may afford a more comprehensive summary of teacher instructional practices and may strengthen conclusions on the consistency of implementation.

Our present study findings lead us to consider the validity of the WIOP observation tool because interpretation of data collected rests on the assumption that a higher score on the WIOP tool relates to better writing instruction. Specifically, on the instructional practices scale, we interpreted higher totals as teachers utilizing more effective practices within a given lesson, and in turn a lower score implying an absence of critical instructional practices needed to ensure students improve in their writing development. Our G study findings, however, may bring light to several pitfalls in this assumption. This perspective (i.e., assuming higher scores represent better lessons) fails to consider if teachers must engage in all practices in a single lesson for it to be deemed effective, and artificially low scores may compromise validity of the observation tool. Indeed, research indicates certain practices, such as explicit instruction, are associated with improved student writing outcomes (e.g., De Smedt & Van Keer, 2018; Graham et al., 2012). Nonetheless, there is also evidence to suggest certain practices, such as opportunities to compose connected text, may predict higher student achievement (Coker et al., 2018a). An alternative observation procedure to enhance reliability and validity of the WIOP observation tool may be to identify a specified proportion of the identified practices and necessary threshold that may still yield effective writing instruction. Likewise, another measurement approach may be to estimate the frequency of the instructional practices within a given lesson and perhaps to consider weighting instructional elements that may be deemed more critical to student learning. With these other observation

methods, writing lessons may be considered effective even when teachers do not engage in all practices within a single lesson.

Together, our findings suggest a need for further research on writing observation tools to strengthen how constructs are defined and interpreted from the scales. Comprehensive observation tools with clearly identified ranges of acceptable scores may enhance the dependability when using these measures to identify differences among teacher practices. Moreover, given the large portion of variance attributed to unmeasured facets, future research is needed to establish observation tools as meaningful instruments for use in everyday school settings where the measurement context varies substantially. As such, our findings suggest future generalizability studies in which other facets are measured and systematically varied could help to identify and then minimize sources of error variance in WIOP scores. Specifically, future observation studies should document variation in lesson topics, types of writing activities, lesson length, and the time of delivery in the school day to determine if the residual variance in the current study is attributable to these other facets. More studies of the WIOP observation tool that investigate alternative measurement procedures will increase the utility of this tool as a mechanism for guiding changes to everyday teacher practice.

### *5.1 Limitations and Future Directions for Research*

This study is novel given few studies have examined reliability of observation tools available in writing; however, interpretations of findings are tempered by several limitations. First, the sample included 25 fourth and fifth grade teachers. Although these sample sizes exceed Webb and colleagues' (1988) recommendation that generalizability studies include minimum of 20 participants with at least two measurements per person, future research is needed to determine if findings hold with larger samples of teachers. Second, to our knowledge, this is only the second study that has utilized the WIOP tool, and our conceptualization of the two scales was largely exploratory. Therefore, findings should not dissuade use of the tool in future studies aiming to document typical practice in the area of writing. Results merely suggest scores should be interpreted with caution, with acknowledgement that many sources of measurement error have yet to be identified. Future experiments examining other measurement facets, such as variance in observers, may identify ways to enhance the dependability of the observation tool.

## 6 Conclusion

This study extends present knowledge on the dependability of an observation tool in documenting writing instruction in upper elementary classrooms. Observation tools offer useful tools for identifying strengths in instruction and areas for improvement.

Yet, conclusions derived from such tools are only useful when they provide dependable data for detecting differences among classroom instruction. The present study suggests that despite recent development of observation tools in writing, more work is needed to establish these tools as reliable measurement instruments. More research is needed to ensure teachers receive timely professional development targeted at areas that will enhance instruction and ultimately advance student progress in writing.

**Conflicts of Interest** We have no known conflicts of interest to disclose.

## References

Biancarosa, G., Afflerbach, P., & Pearson, P. D. (2020). Teaching reading for understanding: Synthesis and reflections on the curriculum and instruction portfolio. *Reaping the rewards of the reading for understanding initiative*. National Academy of Education.

Coker, D. L., Jennings, A. S., Farley-Ripple, E., & MacArthur, C. A. (2018a). When the type of practice matters: The relationship between typical writing instruction, student practice, and writing achievement in first grade. *Contemporary Educational Psychology, 54*, 235–246. https://doi.org/10.1016/j.cedpsych.2018.06.013

Coker, D. L., Jennings, A. S., Farley-Ripple, E., & MacArthur, C. A. (2018b). The type of writing instruction and practice matters: The direct and indirect effects of writing instruction and student practice on reading achievement. *Journal of Educational Psychology, 110*(4), 502–517. https://doi.org/10.1037/edu0000232

Connor, C. M., Morrison, F. J., & Petrella, J. N. (2004). Effective Reading comprehension instruction: Examining child x instruction interactions. *Journal of Educational Psychology, 96*(4), 682–698. https://doi.org/10.1037/0022-0663.96.4.682

Cronbach, L. J., Rajaratnam, N., & Gleser, G. C. (1963). Theory of generalizability: A liberalization of reliability theory. *British Journal of Statistical Psychology, 16*(2), 137–163.

De Smedt, F., & Van Keer, H. (2018). Fostering writing in upper primary grades: A study into the distinct and combined impact of explicit instruction and peer assistance. *Reading and Writing: An Interdisciplinary Journal, 31*(2), 325–354. https://doi.org/10.1007/s11145-017-9787-4

Doabler, C. T., Clarke, B., Kosty, D., Fien, H., Smolkowski, K., Liu, M., & Baker, S. K. (2021). Measuring the quantity and quality of explicit instructional interactions in an empirically validated tier 2 kindergarten mathematics intervention. *Learning Disability Quarterly, 44*(1), 50–62.

Durkin, D. (1978). What classroom observations reveal about reading comprehension instruction. *Reading Research Quarterly, 14*(4), 481–533.

Durkin, D. (1984). Is there a match between what elementary teachers do and what basal reader manuals recommend? *The Reading Teacher, 37*(8), 734–744.

Edmonds, M., & Briggs, K. L. (2003). The instructional content emphasis instrument: Observations of reading instruction. In S. Vaughn & K. L. Briggs (Eds.), *Reading in the classroom: Systems for the observation of teaching and learning* (pp. 31–52). Brookes.

Graham, S., & Harris, K. R. (2015). Common Core State Standards and writing: An introduction. *Elementary School Journal, 115*, 458–463.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879–896. https://doi.org/10.1037/a0029185

Graham, S., Capizzi, A., Harris, K. R., Hebert, M., & Morphy, P. (2014). Teaching writing to middle school students: A national survey. *Reading and Writing, 27*(6), 1015–1042.

Hollo, A., Staubitz, J. L., & Chow, J. C. (2020). Applying generalizability theory to optimize analysis of spontaneous teacher talk in elementary classrooms. *Journal of Speech, Language, and Hearing Research, 63*(6), 1947–1957.

Johnson, E. S., Zheng, Y., Crawford, A. R., & Moylan, L. A. (2019). Developing an explicit instruction special education teacher observation rubric. *The Journal of Special Education, 53*(1), 28–40.

Johnson, E. S., Zheng, Y., Crawford, A. R., & Moylan, L. A. (2020). Examining rater accuracy and consistency with a special education observation protocol. *Studies in Educational Evaluation, 64*, 100827.

Kelcey, B., & Carlisle, J. F. (2013). Learning about teachers' literacy instruction from classroom observations. *Reading Research Quarterly, 48*(3), 301–317. https://doi.org/10.1002/rrq.51

Kotula, A. W., Aguilar, C. M., & Tivnan, T. (2014). *Developing a writing instruction observation protocol: Implications for practice, research, and policy*. Education Development Center.

Moore, C. T. (2016). *gtheory: Apply generalizability theory with R*. R package version 0.1.2. Retrieved from https://CRAN.R-project.org/package=gtheory

National Center for Education Statistics (NCES). (2012). *The nation's report card: Writing. 2011*. (NCES 2012-4700) Institute of Education Sciences, U.S. Department of Education. http://nces.ed.gov/nationsreportcard/pdf/main2011/2012470.pdf

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*(2), 109–119.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. URL http://www.R-project.org/

Semmelroth, C. L., & Johnson, E. (2014). Measuring rater reliability on a special education observation tool. *Assessment for Effective Intervention, 39*(3), 131–145.

Shavelson, R. J., & Webb, N. M. (1991). *A primer on generalizability theory*. Sage Publications.

Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the Classroom Observations of Student–Teacher Interactions (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly, 27*(2), 316–328. https://doi.org/10.1016/j.ecresq.2011.09.004

Snijders, T. A. B., & Bosker, R. J. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.

StataCorp. (2015). *Stata statistical software: Release 14*. StataCorp LP.

Swanson, E. A. (2008). Observing reading instruction for students with learning disabilities: A synthesis. *Learning Disability Quarterly, 31*(3), 115–133.

Troia, G. A., & Olinghouse, N. G. (2013). The common core state standards and evidence-based educational practices: The case of writing. *School Psychology Review, 42*(3), 343–357. https://doi.org/10.1080/02796015.2013.12087478

Webb, N. M., Rowley, G. L., & Shavelson, R. J. (1988). Using generalizability theory in counseling and development. *Measurement and Evaluation in Counseling and Development, 21*(2), 81–90.

Webb, N. M., Shavelson, R. J., & Haertel, E. H. (2006). Reliability coefficients and generalizability theory. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26, pp. 81–124). Elsevier.

# Relations Among Teachers' Efficacy Beliefs, Knowledge, Preparation, Abilities, and Practices: Expanding Our Understanding of Teacher Characteristics That Impact Writing Instruction


**Gary A. Troia**

**Abstract** Malleable factors believed to affect classroom writing instructional practices such as teachers' preparation for writing instruction, their perceived competence to teach writing, and their knowledge about writing-related concepts and instructional practices have been investigated in a small number of studies (and indeed found to be related to instruction), though typically not all simultaneously, and no study has evaluated the impact of teachers' own writing ability on their instruction or how ability is related to these other factors. This chapter presents data from a small scale study of 41 fourth- and fifth-grade teachers that examines the relationships between all four of these factors and their associations with instructional practices, specifically frequency of teaching writing in general, specific components of writing, and particular genres of writing, and with observed writing lesson quality. Ability and knowledge were found to be significantly correlated, as were efficacy beliefs and preparation. Professional preparation and perceived competence for teaching writing were related to writing lesson quality and how often each week teachers spent on writing generally, though not to how frequently they spent teaching specific aspects of writing or writing genres. Conversely, teacher knowledge and ability had little relationship to writing instructional practices. These findings highlight the importance of quality professional learning opportunities for teachers of writing, as these opportunities likely positively impact their efficacy beliefs and, ultimately, their writing instruction.

**Keywords** Writing · Teacher knowledge · Teacher preparation · Teacher motivation · Lesson quality

G. A. Troia (✉)
Michigan State University, East Lansing, MI, USA
e-mail: gtroia@msu.edu



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_16

## 1 Introduction

Over the past four decades or so, scholars have increasingly focused investigations on the malleable factors believed to affect classroom instructional practices, namely (a) teachers' preparation to teach both generally (e.g., Darling-Hammond et al., 2005) and to teach specific domains of content (e.g., Ball & McDiarmid, 1989), (b) teachers' knowledge of instructional content and associated pedagogical approaches (e.g., Ball, 2000; Shulman, 1987), and (c) teachers beliefs and attitudes regarding their students, themselves as educators, their instruction, and so forth (e.g., Fang, 1996; Woods, 1996). The logic model that undergirds these efforts presumes there is a positive relationship between each of these factors and teachers' instructional practices (i.e., more preparation, broader and deeper knowledge, and positive beliefs and attitudes are associated with more desirable instructional practices reflective of clear, consistent, and convincing scientific evidence), that each of these factors can indeed be enhanced through directed interventions such as targeted high-quality professional development, and that subsequent changes in one or more of these factors result in meaningful transformation of teachers' practices, which then imparts greater achievement gains for students in their classrooms. Generally, the results of investigations have borne out positive relationships between each putatively malleable factor and teachers' instructional practices. The research evidence regarding just how malleable each of these factors are and the degree to which changes in them affect educational practice is somewhat less robust and thus should be considered emergent evidence (e.g., Cook et al., 2012; Dillard, 2004; Gallagher et al., 2017; Oh, 2011; Wolbers et al., 2017).

In this chapter, I summarize the findings from investigations related to associations between teacher beliefs (specifically, teaching efficacy beliefs or perceived competence for teaching), preparation, and knowledge, both with each other and with classroom practices in the instructional domain of written expression. I then present findings from a funded research study that evaluated in a sample of 41 fourth- and fifth-grade educators the relations between their writing-related professional preparation, knowledge, and efficacy beliefs, as well as their own writing abilities, and how these factors were associated with the teachers' reported time spent on teaching writing and observed writing lesson quality. Teachers' writing abilities are included as a variable in this study because there is evidence that teachers who do not like to write or perceive themselves as weak writers (both of which may indicate actual weaker writing skills) tend to adopt less beneficial instructional practices: they ask their students to write less (Claypool, 1980), avoid conferencing with students about writing (Bizaro & Toler, 1986), and use their own writing and experiences for modeling infrequently (Lane, 1993). Finally, I discuss the implications of this study for the larger body of research on malleable factors thought to influence instruction.

### *1.1 Teacher Efficacy*

Self-efficacy, an individual's assessment of their competence to perform a future task, is perhaps the most well established and well researched aspect of human motivation (Bandura, 1997). Generally speaking, measures of self-efficacy are positively related to the amount of effort expended to perform a task, persistence with a difficult task, the recruitment of strategies to accomplish a task, and actual task performance, regardless of one's age, gender, or ethnicity (e.g., Bandura, 1997; Pajares, 1996; Pintrich & DeGroot, 1990). Self-efficacy also is a key variable in teachers' achievement, that is, their teaching success. When teachers believe they can execute courses of action to teach well (see Tschannen-Moran et al., 1998, for a theoretical model of teaching self-efficacy), they frequently exhibit desirable teaching behaviors that yield positive effects on their students' motivation to learn (Ashton & Webb, 1986; Midgley et al., 1989) and their students' academic achievement (e.g., Ashton & Webb, 1986). In fact, a systematic review of 165 teacher efficacy studies (none of which focused on writing) by Zee and Koomen (2016) demonstrated that teacher efficacy has direct and indirect influences on classroom instruction and student success. Teachers with strong teaching self-efficacy tend to (1) be more likely to adopt innovative teaching practices (Guskey, 1982, 1988; Smylie, 1988); (2) use more student-centered learning activities and proactive management techniques (Rose & Medway, 1981); (3) provide more support and feedback and generally persist more when students struggle (Bender & Ukeje, 1989; Gibson & Dembo, 1984; Shaw et al., 2007); and (4) refer students for special education less often, presumably because they feel confident they can address struggling learners' needs (Soodak & Podell, 1993). Thus, much in the same way that positive self-efficacy beliefs are the hallmark of motivated learners, they are an equally powerful force in the motivation for teaching.

Teachers usually report slight to moderate levels of perceived competence in teaching writing across the globe, including North America (e.g., Gilbert & Graham, 2010), South America (e.g., Bañales et al., 2020), Europe (e.g., De Smedt et al., 2016; Rietdijk et al., 2018), and East Asia (e.g., Hsiang & Graham, 2016). A number of investigations, using surveys and/or observations, have found that, compared with teachers who are less confident regarding their capacity to teach writing effectively, more confident teachers spend a greater amount of time teaching writing (and their students spend more time writing) and tend to provide instruction aligned with evidence-based practices (Brindle et al., 2016; De Smedt et al., 2016; Gilbert & Graham, 2010; Gillespie et al., 2014; Rietdijk et al., 2018; Tschannen-Moran & Barr, 2004; Veiga Simão et al., 2016). For instance, in Gilbert and Graham's (2010) survey study of grade 4–6 teachers in the United States, efficacy beliefs for teaching writing plus preparation to teach writing predicted a significant amount of variance (nearly 10%) in teachers' use of evidence-based writing instruction practices, controlling for time devoted to writing and students' personal characteristics, with teacher efficacy providing the statistically unique contribution to variance in practice. Similarly, an observational study of grade 4–6 teachers in the Netherlands

(Rietdijk et al., 2018) found that teacher efficacy was statistically related to time allocated for teaching writing (10% of variance). A survey study conducted by Hsiang and Graham (2016) of grade 4–6 teachers in the Greater China region (Taipei City, Macao, and Beijing) found that teachers' attitudes toward their own writing, attitudes toward teaching writing, and efficacy for teaching writing collectively accounted for a statistically significant 7%, 17%, and 7% of variance in teaching writing, providing additional writing support, and facilitating the writing process, respectively, after time devoted to teaching writing, teacher certification, and classroom characteristics were controlled. Efficacy beliefs accounted for statistically unique variance for all three of these teaching practices, while attitudes toward teaching writing accounted for statistically unique variance in teaching writing and providing additional support. Similar findings were obtained in a subsequent survey study by the authors of grade 7–9 teachers in the Greater China region (Hsiang et al., 2018). Thus, research evidence shows a fairly consistent positive predictive relationship between efficacy beliefs for teaching writing and teaching practices, though there are exceptions when examining teachers' adaptations for struggling learners (cf. Graham et al., 2003, 2016).

## *1.2 Teacher Knowledge and Professional Preparation*

Both teachers' subject-matter content knowledge (e.g., knowledge of English orthographic conventions, how written language represents morphophonemic information in oral language) and pedagogical content knowledge (e.g., how to teach students orthographic conventions and relations between oral and written language at multiple levels of language to improve their spelling abilities) are important for promoting high quality instruction. In the domain of English language arts, most research has focused on teachers' content knowledge, especially knowledge related to the structures and functions of language necessary to help children learn to read or spell. A consistent finding is that, though there is great variability in practicing (and preservice) teachers' and teacher educators' content knowledge about language, the majority lacks a sufficient level of knowledge thought to be necessary for designing meaningful lessons and providing effective feedback to students (e.g., Joshi et al., 2009; McCutchen & Berninger, 1999; Parr et al., 2007; Washburn et al., 2011). Piasta et al. (2009) found that more instructional time with a first grade teacher knowledgeable about literacy concepts yielded better word reading gains than more instructional time with a less knowledgeable teacher. Although there is little work on teacher writing-related content and pedagogical content knowledge, save for research on spelling and grammar-related knowledge (e.g., Cajkler & Hislam, 2002; Carreker et al., 2010; Harper & Rennie, 2009; Myhill et al., 2013), the available research suggests that those with greater knowledge are at least more likely to be inclined to incorporate more writing activities in their classrooms (Chambless & Bass, 1995; Street, 2003; Vaisman & Kahn-Horwitz, 2020).

Furthermore, there is evidence that content knowledge alone is insufficient for effective teaching practice—teachers must possess an integrated schema for subject-matter knowledge and the pedagogical principles used to apply that knowledge in the classroom, that is, pedagogical content knowledge—to become better teachers of writing (Bartels, 2005; Borg, 2003; Burns & Knox, 2005).

Professional preparation, both prior to employment (pre-service) and during (in-service) employment as an educator is the most obvious mechanism for increasing teachers' content and pedagogical content knowledge about literacy broadly and writing more specifically. There is a growing body of research that suggests strong professional preparation and ongoing professional development (see Graham, 2019, for a list of key attributes of in-service writing-related professional development) can positively impact current and future teachers' writing knowledge, instructional practices, and their students' writing outcomes (e.g., Gallagher et al., 2017; Harris et al., 2012; McCutchen et al., 2002, 2009; Purvis et al., 2016; Wolbers et al., 2017). However, most teachers report that the coverage of writing content and pedagogical content knowledge in their teacher education programs was at best adequate but more often inadequate (cf. Brindle et al., 2016; Hochstetler, 2007; Kiuhara et al., 2009). This may be due to a variety of reasons, including limited pre-service coursework dedicated to writing instruction (Totten, 2005). In a nationally representative sample of third through eighth grade teachers in the United States, Troia and Graham (2016) found one in five teachers reported having no exposure to writing instruction pedagogy in any of their preservice courses, while about a third of the teachers had taken at least one course devoted to writing instruction. These findings were partially replicated in a later survey study (Troia & Graham, 2017) with a similar group of teachers, where one-fifth of the respondents had no coursework with content devoted to writing, but about 60% had at least one course devoted to writing. A survey of 63 literacy teacher educators in the U.S. revealed that only about a quarter (28%) of these college and university faculty taught a course devoted solely to writing instruction (Myers et al., 2016). This is distressing because teacher preparation uniquely predicts significant variance in teachers' use of effective instructional practices (e.g., Bañales et al., 2020; De Smedt et al., 2016; Gilbert & Graham, 2010; Veiga Simão et al., 2016).

The relationships between teacher efficacy, knowledge, and preparation in the domain of writing have not been thoroughly investigated. There is evidence for a modest positive zero-order correlation between the degree of preparation to teach writing and teaching writing efficacy beliefs, with correlations ranging from about 0.25 (Troia & Graham, 2017) to about 0.35 (Gilbert & Graham, 2010; Hsiang et al., 2020). How well teacher knowledge about writing-related concepts and pedagogy is related to teaching efficacy or professional preparation has rarely been addressed in the extant research. Carreker et al.'s (2010) intervention study found that special education teachers provided with 0 minutes professional development that focused on phonological, morphological, and syllable structure of words versus some amount of professional development (30, 60, and 120 minutes) displayed significantly less knowledge about these topics and associated instructional activities (i.e.,

pedagogical content knowledge), and teachers who were provided the maximum amount of professional learning opportunities (which included two 60-minute sessions bookending a year of teaching with observation and feedback) displayed significantly greater knowledge than those with less preparation. Overall, differences between the four groups accounted for 28% of variance in spelling content and pedagogical content knowledge.

Reported below is a study in which the relations between teachers' writing-related professional preparation, content and pedagogical knowledge, efficacy beliefs, and writing abilities are examined, and the degree to which these variables are associated with their self-reported time spent teaching writing (generally, with respect to specific aspects of writing instruction, and with respect to individual genres) and observed writing lesson quality are evaluated. Because relations between teachers' knowledge and writing ability with the other variables are not well understood, there were no a priori expectations involving these; however, significant relations between preparation to teach writing and efficacy beliefs with the instructional variables were expected based on the available research. The specific research questions addressed in this study include: (1) Are there significant zero-order correlations between efficacy beliefs regarding teaching writing, preparation to teach writing, writing content and pedagogical content knowledge, writing ability (spelling and essay quality), teacher-reported characteristics of writing instruction, and observed writing lesson quality and (2) How much variance in writing lesson quality is predicted by preparation to teach writing, efficacy beliefs, knowledge, and writing ability, controlling for how often writing is taught?

## 2 Method

### *2.1 Participants*

Forty-one teachers (39 females; 37 European Americans) of fourth- and fifth-grade students (ages 9–11) from 24 different schools representing urban, suburban, and rural areas in the mid-Michigan region of the United States participated in data collection for this study. The teachers were on average 42 years old ($SD$ = 9.76) with 15.24 years of overall teaching experience ($SD$ = 8.42) and 6.76 years of teaching experience for grade four or five ($SD$ = 5.64). The majority of teachers (80.5%) had earned a master's degree in education. These teachers taught classes with anywhere between 15 and 30 students ($M$ = 24.90, $SD$ = 2.96). Their classes comprised students ranging from 3.85% to 100% considered low income based on eligibility for free or reduced cost school meals ($M$ = 43.41%, $SD$ = 35.86%), 3.85–53.33% eligible for special education services ($M$ = 12.55%, $SD$ = 9.76%), and 3.45% to 69.57% classified as non-native English learners ($M$ = 11.86%, $SD$ = 16.79%).

## 2.2 Instrumentation

We administered the following instruments to the teachers in this study either through an online questionnaire (via Qualtrics survey platform) or, for the tasks to assess teacher writing ability, in person at their school. The online questionnaire first asked teachers to provide sociodemographic information about themselves and their classrooms (as reported above), then information regarding their preparation to teach writing, followed by information regarding their writing instruction, their efficacy beliefs, and finally, their knowledge of writing and writing pedagogy. When the teachers had completed the questionnaire, a time was arranged for the author to administer the norm-referenced writing ability test individually within 1 month of completing the questionnaire.

### Preparation to Teach Writing

Three questionnaire items asked teachers to report the nature of their pre-service and in-service professional development opportunities to learn about teaching writing. They were asked: (1) how many pre-service courses they had that included some information about writing instruction (0 = none, 1 = one, 2 = two or more) or were devoted fully to writing instruction (3 = one, 4 = two or more); (2) how many in-service activities focused on writing instruction they had participated in over the prior 5 years (0 = none, 1 = 1–2, 2 = 3–4, 3 = 5–6, 4 = more than 6) such as live or online workshops or formal or informal coaching/mentoring activities; and (3) how many unique independent (i.e., beyond pre-service in-service opportunities) learning activities in which they had engaged to learn about writing instruction, including spending more time writing, reading about effective writing instruction, observing other teachers' writing instruction, soliciting feedback on their writing instruction, and taking additional courses or workshops (thus, a maximum of five independent efforts was possible). These three items were summed to yield a professional development composite score.

### Characteristics of Writing Instruction

Three questionnaire items asked teachers about how often they taught different aspects of writing: (1) how many days per week writing is taught; (2) how many minutes per week (up to 100 minutes) were devoted to teaching spelling, handwriting, keyboarding/typing, grammar, capitalization and punctuation, planning, drafting, and revising and editing; and (3) how many weeks per year (up to 20 weeks) were devoted to teaching the macro genres of narration, persuasion, information, and poetry.

### Efficacy Beliefs

Eight questionnaire items asked teachers about their perceived competence for teaching writing using a scale of 1 (strongly disagree) to 6 (strongly agree), with negatively worded items reverse scored such that a higher mean score across items represented greater teaching efficacy. The items asked teachers if they had effective ways to teach writing, knew how to increase student retention of introduced concepts, knew the steps for teaching a writing concept or skill for quick mastery, could try to help students with their most difficult writing problems, could exert additional effort to help a student write better, could adjust the difficulty of a writing assignment for a student experiencing trouble, could accurately assess the reason a student was experiencing difficulty and make an appropriate accommodation, and knew how to redirect disruptive behavior during writing time. These same eight items have been used in prior studies and have been often found to yield a single factor with strong internal consistency reliability ($\alpha > 0.80$, e.g., Brindle et al., 2016; Graham et al., 2001).

### Knowledge of Writing and Writing Pedagogy

A 32-item instrument with 116 unique multiple-choice or fill-in responses scored as correct or incorrect—the Teaching Writing Knowledge Test (TWKT)—was used to assess teachers' writing content knowledge and pedagogical content knowledge. It includes items from research-based spelling and grammar knowledge tests for teachers (e.g., Cajkler & Hislam, 2002; Harper & Rennie, 2009; Myhill et al., 2013; Spear-Swerling et al., 2005), released items from other available tests used to assess basic language skills and pedagogical content knowledge in teachers (e.g., Cambridge English Teaching Knowledge Test), and unique items developed by the author. The TWKT assesses teachers' knowledge of morphemes, phonemes, syllables, consonant and vowel digraphs and consonant blends, root words and suffixes (both derivational and inflectional), regular and irregular spelling patterns, parts of speech, sentence structure (e.g., simple versus compound versus complex), writing mechanics (capitalization, punctuation, and spelling), genre traits, evidence-based writing instruction practices, and targeted instructional activities to address varied aspects of writing. Correct answers are summed to arrive at a total score. Based on the data from all teachers in this study and a group of five other teachers on which it was piloted, the TWKT has $\alpha = 0.72$ and a test-retest reliability of 0.99.

### Writing Ability

The Spelling and Written Expression subtests of the Wechsler Individual Achievement Test-Second Edition (WIAT-II), a test with normative data for individuals as old as 85 years (the more recent version does not have normative data for adults older than 50 years of age), was administered to measure teachers' writing abilities. Raw scores on each subtest were converted to standard scores with a mean

of 100 and standard deviation of 15. The Spelling and Written Expression subtests have average split-half internal consistency reliabilities of 0.94 and 0.77, respectively, and average test-retest reliabilities of 0.96 and 0.85, respectively. Average scorer agreement based on intraclass correlations for the Written Expression subtest is 0.85.

### *2.3 Observation*

After administering the online questionnaire and norm-referenced writing ability test, live observations of teachers' classroom writing instruction were conducted. Observations lasted the duration of the writing lesson, usually about an hour. For the purposes of the study reported here, lesson quality ratings are used that were assigned based on four dimensions: (1) student engagement—the degree of student attention, interest, persistence, and excitement about the lesson; (2) classroom climate—the degree to which the teacher displayed sensitivity to diverse perspectives and backgrounds, approachability, and high expectations; (3) classroom organization—the degree of positive behavior management and effective use of space and materials to promote safety, efficiency, and productivity; and (4) instruction—the degree of thought provoking questions, immediate and concrete feedback, high level discussions, rigorous tasks, and effective learning scaffolds. Each dimension was rated by the author on a scale of 1 (minimal) to 5 (extensive), and the mean of the ratings across dimensions served as the lesson quality score. If more than one observation was conducted for a teacher (2–3 months apart), the ratings across lessons also were averaged. Eight teachers were observed four times, 13 were observed three times, five were observed twice, and 15 were observed only once. The goal was to observe all teachers 3–4 times during the academic year, but the COVID-19 pandemic prevented this from occurring for the most recent cohort of teachers participating. Twenty-two percent of the live observations were rated by a second trained doctoral student observer present in the classroom to calculate interobserver reliability. A two-way random effects ICC for single measures using absolute agreement was calculated for the mean quality ratings assigned by the two raters and found to be 0.46, which is considered fair agreement (Cicchetti, 1994).

## 3 Results

### *3.1 Basic Descriptive Statistics*

All variables were normally distributed except for how much time teachers reported spending teaching handwriting or keyboarding, which were positively skewed. Thus, a square root transformation was applied for these two variables for any subsequent analyses. Basic descriptive statistics for each variable are provided in Table 1.

**Table 1** Means, standard deviations, and ranges for study variables

| Variable | Mean | SD | Range |
|---|---|---|---|
| Professional development composite score | 5.66 | 2.91 | 1–13 |
| Days/week teach writing | 4.17 | 1.02 | 1–5 |
| Minutes/week teaching spelling | 20.34 | 24.58 | 0–100 |
| Minutes/week teaching handwriting | 1.44 | 5.11 | 0–25 |
| Minutes/week teaching keyboarding | 5.90 | 13.20 | 0–60 |
| Minutes/week teaching grammar | 26.32 | 20.45 | 0–100 |
| Minutes/week teaching capital/punctuation | 17.51 | 15.91 | 0–60 |
| Minutes/week teaching planning | 40.34 | 28.06 | 0–100 |
| Minutes/week teaching drafting | 48.98 | 25.57 | 0–100 |
| Minutes/week teaching revising/editing | 41.27 | 23.91 | 0–100 |
| Weeks/year teaching narrative | 7.68 | 3.98 | 0–18 |
| Weeks/year teaching persuasive | 7.73 | 3.96 | 0–18 |
| Weeks/year teaching informative | 9.05 | 4.24 | 0–19 |
| Weeks/year teaching poetic | 2.37 | 2.35 | 0–10 |
| Mean efficacy beliefs | 4.16 | 0.59 | 2.75–5.00 |
| TWKT total score | 89.02 | 10.48 | 62–107 |
| WIAT-II spelling standard score | 110.73 | 7.31 | 96–123 |
| WIAT-II written expression standard score | 120.37 | 10.45 | 98–141 |
| Mean observed lesson quality | 3.85 | 0.85 | 1.50–5.00 |

## *3.2 Relationships Between Variables*

Teacher knowledge about writing (TWKT total score) was significantly correlated with standard scores on the WIAT-II Spelling ($r = 0.41$, $p = 0.01$) and Written Expression ($r = 0.40$, $p = 0.01$) subtests, together referred to as teacher writing ability, which were significantly correlated with each other ($r = 0.35$, $p = 0.03$). Preparation to teach writing was not significantly associated with any of these, but was significantly related to teacher efficacy beliefs ($r = 0.36$, $p = 0.02$). Teacher efficacy beliefs and WIAT-II Spelling subtest scores were significantly related ($r = -0.37$, $p = 0.02$). Preparation to teach writing, teaching efficacy beliefs, teaching knowledge, and writing ability were not significantly correlated with reported time spent teaching any particular instructional practice, aside from inverse relationships between the WIAT-II Spelling and Written Expression subtest scores and the number of minutes spent teaching typing each week ($r = -0.33$, $p = 0.04$ and $r = -0.41$, $p = 0.01$, respectively). Preparation to teach writing and teaching efficacy beliefs were, however, significantly positively correlated with the number of days each week writing was reportedly taught ($r = 0.32$, $p = 0.04$ and $r = 0.35$, $p = 0.03$, respectively). There were significant positive correlations between observed lesson quality and teacher preparation to teach writing ($r = 0.41$, $p = 0.01$) and teacher efficacy ($r = 0.32$, $p = 0.04$), as well as the number of days per week writing was taught ($r = 0.36$, $p = 0.02$) and the number of weeks each year the informative writing genre was reportedly taught ($r = 0.34$, $p = 0.03$).

### 3.3 Prediction of Writing Lesson Quality

To determine the variance in mean lesson quality attributable to preparation to teach writing, writing-related teacher efficacy beliefs, teacher writing ability, and teacher knowledge about writing, we performed regression analysis using simultaneous entry, following entry of the number of days per week writing was taught. Number of days each week spent teaching writing accounted for 12.7% of variance in lesson quality, $F(1, 39) = 5.70$, $MSE = 0.65$, $p = 0.02$. The remaining predictors accounted for an additional 17.1% of variance in writing lesson quality, $F(6, 34) = 2.41$, $MSE = 0.60$, $p = 0.05$. However, none of the variables were significant unique predictors.

## 4 Discussion

### 4.1 Summary of Findings

As a group, the fourth- and fifth-grade teachers in this study demonstrated stronger writing skills than the general population of U.S. adults, which would be expected given that teachers are more highly educated than the general population. Nevertheless, the teachers did not demonstrate a high degree of knowledge about writing-related concepts and pedagogy—they on average scored 77% accuracy on the TWKT and no teacher achieved a score over 92% on this measure. Though there was a modest relationship between teachers' writing ability and knowledge, it does not appear to be the case that high average to above average writing ability directly translates to superior knowledge about writing and teaching writing—good writing skills are likely necessary but insufficient for developing the metalinguistic competence (e.g., correctly segmenting written words into phonemes, morphemes, and syllables, correctly identifying various syntactic and discourse structures) and pedagogical knowledge evaluated with the TWKT. Other scholars have noted facility with language does not necessarily confer explicit insight about language and how to use that insight to inform teaching (e.g., Borg, 2003).

The teachers in this study did not have a large number of professional learning experiences devoted to writing instruction, with a mean score of 5.7 out of 13 maximum points on the professional development composite. They expressed slightly positive (around a 4 on a 6-point scale) teaching writing efficacy beliefs, a finding that tracks closely with other studies that have examined this construct (e.g., Gilbert & Graham, 2010; Troia & Graham, 2017), meaning that they were only somewhat confident they could address their students' writing needs in the classroom. The significant relationship observed between degree of preparation to teach writing and perceived competence for teaching writing perhaps highlights how merely adequate professional preparation is tied to only fair teaching confidence, though it is unclear from these data the directionality of this relationship, or if the relationship is

mediated by another factor, issues unresolved in other studies that have observed this relationship (Carreker et al., 2010; Gilbert & Graham, 2010; Hsiang et al., 2020; Troia & Graham, 2017).

The group of teachers typically spent 4 days per week teaching writing. During each week, on average, they devoted between 40–50 minutes to composing (planning, drafting, revising, and editing), between 18–26 minutes to conventions (spelling, grammar, capitalization, and punctuation), and less than 6 minutes to keyboarding and handwriting. The limited time allocated for text production skills perhaps is warranted considering students in grades 4 and 5 probably already have had sufficient instruction in these skills, though this is not a certainty and may not adequately consider the needs of students who may continue to struggle with handwriting and keyboarding. Teachers' instruction across the school year focused mostly on informative writing (about 9 weeks of instruction on average), followed closely by persuasive (almost 8 weeks of instruction) and narrative (about seven and a half weeks of instruction) writing, and then poetry (only 2 weeks of instruction). It is disappointing that so little time was devoted to poetry in light of the potential for this genre to highlight semantic, pragmatic, and metalinguistic aspects of language, bolster creativity and reflection, enhance motivation to write, etcetera (e.g., Certo et al., 2010), but, of course, this genre is not usually assessed in annual accountability measures for writing achievement.

Although preparation to teach writing and efficacy beliefs were positively associated with the number of days per week writing was taught, they were not associated with how frequently teachers taught specific genres, aspects of the writing process, or different writing skills. Moreover, teacher knowledge and writing ability were mostly unrelated to teachers' allocated time for teaching writing each week or any particular aspect of writing. Thus, these data suggest that professional development and perceived competence for teaching writing have measurable effects at a macroscopic level—how often each week teachers spend on writing generally—but little effect at more granular levels of writing instruction, and that teacher knowledge and ability have a negligible impact. These results only partially comport with extant findings from other research studies as described earlier (cf. Bañales et al., 2020; Gilbert & Graham, 2010; Rietdijk et al., 2018; Vaisman & Kahn-Horwitz, 2020), perhaps because of how writing was defined in this study (areas and genres of writing rather than unique instructional practices), how time spent teaching was measured (minutes per week or months per year rather than a 6- or 7-point frequency scale ranging from never to daily), and the use of a novel teacher writing-related knowledge assessment. Importantly, there were significant though modest correlations between teacher preparation, efficacy beliefs, and writing lesson quality, which also was related to number of days each week writing was taught and number of weeks each year informative writing was taught (perhaps because a majority of observed lessons were of teachers working with students on informative writing). Thus these zero-order correlations indicate professional development and teaching writing efficacy are indeed important contributors to variance in writing instruction quality, confirming results from the few other observation studies available (e.g., Rietdijk et al., 2018; Troia et al., 2011). When examining the combined

and unique contributions to lesson quality made by teacher preparation, efficacy beliefs, knowledge, and ability, controlling for how many days each week writing was taught, these predictors explained 17.1% of variance, but none made a significant unique contribution.

## *4.2 Implications*

Because the structure of teacher education programs and the kinds and quality of on-the-job professional development teachers are provided are most amenable to change, and these experiences likely have a potentiating effect on teachers' content and pedagogical content knowledge and perceived competence for teaching, the nature of professionally-related learning experiences available to educators is perhaps the most worthy teacher background variable for intervention. A recent meta-analysis (Darling-Hammond et al., 2017) identified seven traits of effective professional learning experiences insofar as they positively impact teaching behaviors. Effective professional development: (1) is content focused, including a focus on related teaching actions, and is informed by outside experts for relevance, currency, and evidentiary support, (2) uses active learning (e.g., problem solving activities, role playing, examining and reflecting on student artifacts and teacher demonstrations), (3) involves meaningful collaboration (e.g., teachers develop rapport with trusted colleagues and have ample opportunities to work together with partners or groups), (4) employs modeling of effective teaching practices through video-recorded or live demonstrations, case studies, and peer observation, (5) uses contextualized and personalized coaching, (6) provides substantive and constructive feedback and opportunities for reflection at regular and frequent intervals, and (7) is sustained and, when possible, embedded in actual classrooms and schools. These qualities are important not just for in-service professional development but also pre-service teacher education. Moreover, these qualities directly reflect Graham's (2018) writers-within-communities model of writing development. Though primarily focused on writing development and the intrapersonal (e.g., neuropsychological, cognitive, and physical attributes of individuals), interpersonal (e.g., communicative goals, reader response, community supports and expectations), and societal (e.g., cultural, institutional, historical) forces that shape said development as writers participate in concentrically organized communities of writing practice, this model also can be applied to teacher learning about writing and writing instruction as teachers navigate their communities of professional practice. Ensuring that all these professional development qualities are reflected in learning opportunities for future and current teachers of writing should help elevate their confidence in teaching writing and their actual writing instruction to the benefit of their students.

Although teacher knowledge and ability were not overtly important for predicting writing instruction practices in this study, this does not foreclose the need for further investigation of these variables. This study used a newly developed test of teacher content and pedagogical content knowledge and is the first known study to

include teacher writing ability as a predictor. These considerations, combined with the small convenience sample of upper elementary teachers, warrant the continued exploration of teacher knowledge and ability in the domain of writing. At any rate, findings from the original research presented in this chapter reinforce the importance of professional development and teacher efficacy beliefs for strong classroom writing instruction (beyond how often writing is taught) and extend similar findings from the small number of observation studies and the larger group of survey studies noted previously.

**Funding** The research described in this chapter was supported in part by grant #R305A160049 from the U.S. Department of Education, Institute of Education Sciences, to Michigan State University. Statements do not necessarily reflect the positions or policies of this agency, and no official endorsement by it should be inferred.

## References

Ashton, P. T., & Webb, R. B. (1986). *Making a difference: Teachers' sense of efficacy and student achievement*. Longman.

Ball, D. L. (2000). Bridging practices: intertwining content and pedagogy in teaching and learning to teach. *Journal of Teacher Education, 51*(2), 241–247. https://doi.org/10.1177/0022487100051003013

Ball, D. L., & McDiarmid, G. W. (1989). *The subject matter preparation of teachers*. National Center for Research on Teacher Education.

Bañales, G., Ahumada, S., Graham, S., Puente, A., Guajardo, M., & Muñoz, I. (2020). Teaching writing in grades 4–6 in urban schools in Chile: A national survey. *Reading and Writing: An Interdisciplinary Journal, 33*, 2661–2696. https://doi.org/10.1007/s11145-020-10055-z

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/Henry Holt & Co..

Bartels, N. (Ed.). (2005). *Applied linguistics and language teacher education*. Springer. https://doi.org/10.1007/1-4020-2954-3

Bender, W. N., & Ukeje, I. C. (1989). Instructional strategies in mainstream class-rooms: Prediction of the strategies teachers select. *Remedial and Special Education, 10*(2), 23–30. https://doi.org/10.1177/074193258901000206

Bizaro, P., & Toler, H. (1986). The effects of writing apprehension on the teaching behaviors of writing center tutors. *The Writing Center Journal, 7*, 37–43.

Borg, S. (2003). Teacher cognition in grammar teaching: A literature review. *Language Awareness, 12*(2), 96–108. https://doi.org/10.1080/09658410308667069

Brindle, M., Graham, S., Harris, K. R., & Hebert, M. (2016). Third and fourth grade teachers' classroom practices in writing: A national survey. *Reading and Writing: An Interdisciplinary Journal, 29*, 929–954. https://doi.org/10.1007/s11145-015-9604-x

Burns, A., & Knox, J. (2005). Realisation(s): Systemic functional linguistics and the language classroom. In N. Bartels (Ed.), *Applied linguistics and language teacher education* (pp. 235–246). Springer. https://doi.org/10.1007/1-4020-2954-3_14

Cajkler, W., & Hislam, J. (2002). Trainee teachers' grammatical knowledge: The tension between public expectations and individual competence. *Language Awareness, 11*(3), 161–177. https://doi.org/10.1080/09658410208667054

Carreker, S., Joshi, R. M., & Boulware-Gooden, R. (2010). Spelling-related teacher knowledge: The impact of professional development on identifying appropriate instructional activities. *Learning Disability Quarterly, 33*, 148–158. https://doi.org/10.1177/073194871003300304

Certo, J. L., Apol, L., Wibbens, E., & Yoon, S. (2010). Teaching poetry writing PreK-12: A review of research and suggested directions for teacher practice and development. In G. A. Troia, R. K. Shankland, & A. Heintz (Eds.), *Putting writing research into practice: Applications for teacher professional development* (pp. 94–114). Guilford Press.

Chambless, M. S., & Bass, J. A. (1995). Effecting changes in student teachers' attitudes toward writing. *Reading Research and Instruction, 35*(2), 153–160. https://doi.org/10.1080/19388079509558204

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*, 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Claypool, S. H. (1980). *Teacher writing apprehension: Does it affect writing assignments across the curriculum*? Washington, DC: Research in Education Report. (ERIC Document Reproduction Service No. ED216387).

Cook, B. G., Shepherd, K. G., Cook, S. C., & Cook, L. (2012). Facilitating the effective implementation of evidence-based practices through teacher-parent collaboration. *Teaching Exceptional Children, 44*(3), 22–30. https://doi.org/10.1177/004005991204400303

Darling-Hammond, L., Holtzman, D. J., Gatlin, S. J., & Heilig, J. V. (2005). Does teacher preparation matter? Evidence about teacher certification, Teach for America, and teacher effectiveness. *Education Policy Analysis Archives, 13*, 1–48. https://doi.org/10.14507/epaa.v13n42.2005

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (with Espinoza, D.). (2017). *Effective teacher professional development*. Learning Policy Institute.

De Smedt, F., Van Keer, H., & Merchie, E. (2016). Student, teacher, and class-level correlates of Flemish late elementary school children's writing performance. *Reading and Writing: An Interdisciplinary Journal, 29*, 833–868. https://doi.org/10.1007/s11145-015-9590-z

Dillard, S. G. (2004). *A study of the impact of the Mississippi Writing Project Summer Institute on teacher efficacy in writing*. Mississippi State University.

Fang, Z. (1996). A review of research on teacher beliefs and practices. *Educational Research, 38*(1), 47–65. https://doi.org/10.1080/0013188960380104

Gallagher, A., Arshan, N., & Woodworth, K. (2017). Impact of the National Writing Project's college-ready writers program in high-need rural districts. *Journal of Research on Educational Effectiveness, 10*, 570–595. https://doi.org/10.1080/19345747.2017.1300361

Gibson, S., & Dembo, M. (1984). Teacher efficacy: A construct validation. *Journal of Educational Psychology, 76*, 569–592. https://doi.org/10.1037/0022-0663.76.4.569

Gilbert, J., & Graham, S. (2010). Teaching writing to elementary students in grades 4–6: A national survey. *The Elementary School Journal, 110*, 494–518. https://doi.org/10.1086/651193

Gillespie, A., Graham, S., Kiuhara, S., & Hebert, M. (2014). High school teachers' use of writing to support students' learning: A national survey. *Reading and Writing: An Interdisciplinary Journal, 27*, 1043–1072. https://doi.org/10.1007/s11145-013-9494-8

Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educational Psychologist, 53*, 258–279. https://doi.org/10.1080/00461520.2018.1481406

Graham, S. (2019). Changing how writing is taught. *Review of Research in Education, 43*, 277–303. https://doi.org/10.3102/0091732x18821125

Graham, S., Harris, K. R., Fink, B., & MacArthur, C. A. (2001). Teacher efficacy in writing: A construct validation with primary grade teachers. *Scientific Studies of Reading, 5*, 177–202. https://doi.org/10.1207/s1532799xssr0502_3

Graham, S., Harris, K. R., MacArthur, C., & Fink-Chorzempa, B. (2003). Primary grade teachers' instructional adaptations for weaker writers: A national survey. *Journal of Educational Psychology, 95*, 279–293. https://doi.org/10.1037/0022-0663.95.2.279

Graham, S., Harris, K. R., Bartlett, B. J., Popadopoulou, E., & Santoro, J. (2016). Acceptability of adaptations for struggling writers: A national survey with primary-grade teachers. *Learning Disability Quarterly, 39*, 5–16. https://doi.org/10.1177/0731948714554038

Guskey, T. R. (1982). Differences in teachers' perceptions of personal control of positive versus negative student learning outcomes. *Contemporary Educational Psychology, 7*, 70–80. https://doi.org/10.1016/0361-476x(82)90009-1

Guskey, T. R. (1988). Teacher efficacy, self-concept, and attitudes towards the implementation of instructional innovation. *Teaching and Teacher Education, 4*, 63–69. https://doi.org/10.1016/0742-051x(88)90025-x

Harper, H., & Rennie, J. (2009). "I had to go out and get myself a book on grammar": A study of pre-service teachers' knowledge about language. *Australian Journal of Language and Literacy, 32*(1), 22–37.

Harris, K. R., Lane, K., Graham, S., Driscoll, S., Sandmel, K., Brindle, M., & Schatschneider, C. (2012). Practice-based professional development for self-regulated strategy development in writing: A randomized control study. *Journal of Teacher Education, 63*, 103–119. https://doi.org/10.1177/0022487111429005

Hochstetler, S. (2007). The preparation of preservice secondary English teachers in writing instruction: A case study of three California colleges' education programs. *Action in Teacher Education, 29*(2), 70–79. https://doi.org/10.1080/01626620.2007.10463450

Hsiang, T. P., & Graham, S. (2016). Teaching writing in grades 4–6 in urban schools in the Greater China Region. *Reading and Writing, 29*(5), 869–902. https://doi.org/10.1007/s11145-015-9597-5

Hsiang, T. P., Graham, S., & Wong, P. M. (2018). Teaching writing in grades 7-9 in urban schools in Chinese societies in Asia. *Reading Research Quarterly, 53*(4), 473–507. https://doi.org/10.1002/rrq.2018.53.issue-410.1002/rrq.213

Hsiang, T. P., Graham, S., & Yang, Y. (2020). Teachers' practices and beliefs about teaching writing: A comprehensive survey of grades 1 to 3 teachers. *Reading and Writing: An Interdisciplinary Journal, 33*, 2511–2548. https://doi.org/10.1007/s11145-020-10050-4

Joshi, R. M., Binks, E., Hougen, M., Dahlgren, M., Ocker-Dean, E., & Smith, D. (2009). Why elementary teachers might be inadequately prepared to teach reading. *Journal of Learning Disabilities, 42*, 392–402. https://doi.org/10.1177/0022219409338736

Kiuhara, S., Graham, S., & Hawken, L. (2009). Teaching writing to high school students: A national survey. *Journal of Educational Psychology, 101*, 136–160. https://doi.org/10.1037/a0013097

Lane, B. (1993). *After the end: Teaching and learning creative revision*. Heinemann.

McCutchen, D., & Berninger, V. W. (1999). Those who *know*, teach well: Helping teachers master literacy-related subject-matter knowledge. *Learning Disabilities Research & Practice, 14*, 215–226. https://doi.org/10.1207/sldrp1404_3

McCutchen, D., Abbott, R. D., Green, L. B., Beretvas, S. N., Cox, S., Potter, N. S., Quiroga, T., & Gray, A. L. (2002). Beginning literacy: Links among teacher knowledge, teacher practice, and student learning. *Journal of Learning Disabilities, 35*, 69–86. https://doi.org/10.1177/002221940203500106

McCutchen, D., Green, L., Abbott, R. D., & Sanders, E. A. (2009). Further evidence for teacher knowledge: supporting struggling readers in grades three through five. *Reading and Writing: An Interdisciplinary Journal, 22*, 401–423. https://doi.org/10.1007/s11145-009-9163-0

Midgley, C., Feldlaufer, H., & Eccles, J. S. (1989). Change in teacher efficacy and student self- and task-related beliefs in mathematics during the transition to junior high school. *Journal of Educational Psychology, 81*, 247–258. https://doi.org/10.1037/0022-0663.81.2.247

Myers, J., Scales, R. Q., Grisham, D. L., Wolsey, T. D., Dismuke, S., Smetana, L., Yoder, K. K., Ikpeze, C., Ganske, K., & Martin, S. (2016). What about writing: A national exploratory study of writing instruction in teacher preparation programs. *Literacy Research & Instruction, 55*, 309–330. https://doi.org/10.1080/19388071.2016.1198442

Myhill, D., Jones, S., & Watson, A. (2013). Grammar matters: How teachers' grammatical knowledge impacts on the teaching of writing. *Teaching and Teacher Education, 36*, 77–91. https://doi.org/10.1016/j.tate.2013.07.005

Oh, S. (2011). Preservice teachers' sense of efficacy and its sources. *Psychology, 2*(3), 235–240. https://doi.org/10.4236/psych.2011.23037

Pajares, R. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*, 543–578. https://doi.org/10.3102/00346543066004543

Parr, J. M., Glasswell, K., & Aikman, M. (2007). Supporting teacher learning and informed practice in writing through assessment tools for teaching and learning. *Asia-Pacific Journal of Teacher Education, 35*(1), 69–87. https://doi.org/10.1080/13598660601111281

Piasta, S. B., Connor, C. M., Fishman, B. J., & Morrison, F. J. (2009). Teachers' knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading, 13*(3), 224–248. https://doi.org/10.1080/10888430902851364

Pintrich, P. R., & DeGroot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Purvis, C. J., McNeill, B. C., & Everatt, J. (2016). Enhancing the metalinguistic abilities of preservice teachers via coursework targeting language structure knowledge. *Annals of Dyslexia, 66*, 55–70. https://doi.org/10.1007/s11881-015-0108-9

Rietdijk, S., van Weijen, D., Janssen, T., van den Bergh, H., & Rijlaarsdam, G. (2018). Teaching writing in primary education: Classroom practice time teachers' beliefs and skills. *Journal of Educational Psychology, 110*(5), 640–663. https://doi.org/10.1037/edu0000237

Rose, J. S., & Medway, F. J. (1981). Measurement of teachers' beliefs in their control over student outcomes. *Journal of Educational Research, 74*, 185–190. https://doi.org/10.1080/00220671.1981.10885308

Shaw, D. M., Dvorak, M. J., & Bates, K. (2007). Promise and possibility—hope for teacher education: Pre-service literacy instruction can have an impact. *Literacy Research and Instruction, 46*(3), 223–254. https://doi.org/10.1080/19388070709558469

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22. https://doi.org/10.17763/haer.57.1.j463w79r56455411

Smylie, M. A. (1988). The enhancement function of staff development: Organizational and psychological antecedents to individual teacher change. *American Educational Research Journal, 25*, 1–30. https://doi.org/10.3102/00028312025001001

Soodak, L. C., & Podell, D. M. (1993). Teacher efficacy and student problem as factors in special education referral. *Journal of Special Education, 27*, 66–81. https://doi.org/10.1177/002246699302700105

Spear-Swerling, L., Brucker, P. O., & Alfano, M. P. (2005). Teacher literacy-related knowledge and self-perception in relation to preparation and experience. *Annals of Dyslexia, 55*, 266–296. https://doi.org/10.1007/s11881-005-0014-7

Street, C. (2003). Pre-service teachers' attitudes about writing and learning to teach writing: Implications for teacher educators. *Teacher Education Quarterly, 30*, 33–50.

Totten, S. (2005). Writing to learn for preservice teachers. *Quarterly-National Writing Project, 27*(2), 17.

Troia, G. A., & Graham, S. (2016). Common core writing and language standards and aligned state assessments: A national survey of teacher beliefs and attitudes. *Reading and Writing: An Interdisciplinary Journal, 29*, 1719–1743. https://doi.org/10.1007/s11145-016-9650-z

Troia, G. A., & Graham, S. (2017). Use and acceptability of writing adaptations for students with disabilities: Survey of grade 3-8 teachers. *Learning Disabilities Research & Practice, 32*, 257–269. https://doi.org/10.1111/ldrp.12135

Troia, G. A., Lin, S. C., Cohen, S., & Monroe, B. W. (2011). A year in the writing workshop: Linking writing instruction practices and teachers' epistemologies and beliefs about writing instruction. *The Elementary School Journal, 112*, 155–182. https://doi.org/10.1086/660688

Tschannen-Moran, M., & Barr, M. (2004). Fostering student learning: The relationship of collective teacher efficacy and student achievement. *Leadership and Policy in Schools, 3*(3), 189–209. https://doi.org/10.1080/15700760490503706

Tschannen-Moran, M., Woolfolk-Hoy, A., & Hoy, W. K. (1998). Teacher efficacy: Its meaning and measure. *Review of Educational Research, 68*, 202–248. https://doi.org/10.3102/00346543068002202

Vaisman, E. E., & Kahn-Horwitz, J. (2020). English foreign language teachers' linguistic knowledge, beliefs, and reported practices regarding reading and spelling instruction. *Dyslexia, 26*(3), 305–322. https://doi.org/10.1002/dys.1608

Veiga Simão, A. M., Malpique, A. A., Frison, L. M. B., & Marques, A. (2016). Teaching writing to middle school students in Portugal and Brazil: An exploratory study. *Reading and Writing: An Interdisciplinary Journal, 29*, 955–979. https://doi.org/10.1007/s11145-015-9606-8

Washburn, E. K., Joshi, R., & Binks-Cantrell, E. S. (2011). Teacher knowledge of basic language concepts and dyslexia. *Dyslexia, 17*(2), 165–183. https://doi.org/10.1002/dys.426

Wolbers, K. A., Dostal, H. M., Skerrit, P., & Stephenson, B. (2017). The impact of three years of professional development on knowledge and implementation. *Journal of Educational Research, 110*, 61–71. https://doi.org/10.1080/00220671.2015.1039112

Woods, D. (1996). *Teacher cognition in language teaching*. Cambridge University Press.

Zee, M., & Koomen, H. M. (2016). Teacher self-efficacy and its effects on classroom processes, student academic adjustment, and teacher well-being: A synthesis of 40 years of research. *Review of Educational Research, 86*, 981–1015. https://doi.org/10.3102/0034654315626801

# Teaching Reading and Writing in Primary Grades in Macao: A Qualitative Study


**Tien Ping Hsiang**

**Abstract** In order to capture the dynamic nature of teaching reading and writing in primary grades in Macao, China, an interview-based qualitative approach was used in this study. Fifteen grades 1–3 Macao Chinese language teachers from nine private schools which included Chinese, English, and international schools were interviewed. The study revealed that the participants adopted a balanced approach to teaching Chinese reading and writing, although what they focused on varied; PIRLS affected reading instruction; limited word recognition reduced reading and writing performance; limited time was arranged for teaching writing compositions; and reading and writing were influenced by institutions, society, culture, and policies. Because school regulations shape instructional practices, it is urgent to do research on how to balance reading and writing instruction through school-based curriculum development.

**Keywords** Reading/writing instruction · Chinese language arts · Handwriting · Textbooks · Literary books · Teaching Chinese as a second language (TCSL)

## 1 Introduction

The purpose of this study was to understand how Chinese reading and writing were taught in grades 1–3 in Macao schools and what shaped instructional practices. The study was essential for six reasons.

First, taking a balanced approach to teaching reading and writing is a universal challenge for teachers worldwide. Although reading and writing are key factors to children's success (Graham et al., 2018b), schools/teachers need to decide how to use the limited time scheduled for the teaching of language arts to help students succeed. It is necessary to study how teachers resolve the dilemmas of teaching

T. P. Hsiang (✉)
University of Macau, Macao, China
e-mail: tphsiang@um.edu.mo



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_17

reading and writing in different cultural contexts (i.e., teach to the test or to inspire love for reading and writing; teach skills or knowledge; teach narrative or informative texts; focus on word recognition or comprehension instruction; whether authentic literacy activities and explicit instruction are offered; balance the time spent on teaching reading and writing; which kinds of instruction organization are used; integrate Internet technology with instruction or not; see Cutler & Graham, 2008; Graham et al., 2018b; Fisher et al., 2020; Roe & Smith, 2012).

Second, Chinese is a logographic and morpho-syllabic writing system (Hsiang et al., 2018). Each Chinese character is constructed of at least one component consisting of stroke(s) with its corresponding established stroke sequence, structure, one syllable, and meaning(s). There are a large number of homophones and polyphones (Tse et al., 2007). Because most modern Chinese words are composed of two characters, word recognition and handwriting with an emphasis on the practice of morphological and orthographic awareness is necessary (Liu & Liu, 2020; Wang & Leland, 2011). Due to the difficulties of learning Chinese vocabulary, studying how teachers teach Chinese reading and writing can expand the knowledge of alphabetic based reading and writing instruction (Hsiang et al., 2021).

Third, Chinese students need to recognize at least 1000–1200 Chinese characters before they can do the most basic reading and writing (Wang et al., 2008). It is important to know how teachers help students learn characters and how teachers help primary grade students to read and write when the number of characters for recognition and handwriting is limited (Graham et al., 2018a).

Fourth, there have been only a few empirical studies on how Chinese reading and writing are taught in elementary schools in China. Those studies included textbook analysis (Hsiang, 2012; Hsiang et al., 2021), and how teachers teach Chinese characters (Ruan et al., 2018; Wong, 2013; Yang et al., 2016), reading (Fan et al., 2016; Quan et al., 2019; Xie et al., 2016), and writing (Chan, 2020; Wei et al., 2020; Zhang et al., 2018) in their own classrooms. Mind maps, e-learning, picture books, process writing, and Self-Regulated Strategy Development were the elements which have been integrated with reading or writing instruction in China during the past 5 years. Two studies focused on how writing (L1) or word recognition (L2) was taught among schools (Hsiang & Graham, 2016; Hsiang et al., 2020b). There is no comprehensive research on Chinese reading and writing instructional practices together across schools in China. As writing can improve reading and reading can enhance writing (Graham & Hebert, 2011; Graham et al., 2018c), studies of how both reading and writing are taught in primary grade classrooms are needed.

Fifth, Macao is a multi-lingual society which raises the complexity and diversity of reading and writing instruction due to political, historical, and economic reasons. In Macao, a Special Administrative Region of China, both local and international school students are required to write traditional Chinese characters while students are also suggested to recognize simplified characters (Education and Youth Development Bureau, 2016a, b). As Macao was governed by Portugal, both Chinese and Portuguese are the official languages while Cantonese is the most widely spoken language. English is widely used in education because of the gaming and tourism industries. Schools in Macao can choose Cantonese, Mandarin, English, or

Portuguese as their main medium of instruction (Macao Government Tourism Office, 2020; Macao SAR Government, 2020a). Studying how Chinese reading and writing is taught in different types of schools in Macao can contribute to national and international literacy studies (Li, 2020).

Sixth, according to the results of the 2016 Progress in International Reading Literacy Study (PIRLS; Warner-Griffin et al., 2018), Macao's overall average reading score was higher than the PIRLS scale center point but lower than the averages for 18 regions including the two Chinese cities, Hong Kong and Taipei. Has the PIRLS performance stimulated any changes in Macao primary grade reading and writing instruction? If yes, which part has been influenced by PIRLS? A comprehensive study of reading and writing instructional practices in Macao is necessary to understand how PIRLS has stimulated changes in Macao primary reading and writing instruction.

In order to capture the dynamic nature of teaching reading and writing, an interview-based qualitative approach was used in this study. Because the time arranged for teaching reading or writing affects how it is taught (Hsiang et al., 2018, 2020a; Hsiang & Graham, 2016; Graham et al., 2008), I asked the following four questions:

Q1: How many Chinese language arts classes are there in a regular week?
Q2: Is there extra time devoted to teaching reading or writing?
Q3: How are reading and writing taught?
Q4: What challenges do teachers meet when teaching reading and writing?

During the interview, the interviewer offered keywords such as teaching methods, materials/textbooks, learning activities, homework, and evaluation to stimulate more discussion.

## 2 Methodology

### *2.1 Context*

There are four kinds of primary schools in Macao:

- Chinese schools (N = 48), in which most subjects are taught in Cantonese or Mandarin.
- English schools (N = 9), which offer both primary and secondary education and aim to use English as the language of instruction in all subjects in their secondary section (except Chinese language arts). English receives the most attention by students/parents from primary grades, and some English schools have a Chinese section for students to study in.
- International schools (N = 3) which use English as the language of instruction except Chinese language arts taught in Mandarin. Local students who don't have foreign passports can study in international schools.

- Portuguese schools (N = 1), in which Portuguese is the main language of instruction.

For additional context, 89% (N = 54) of the Macao primary schools are private schools (Education and Youth Development Bureau, 2021). The majority of the schools offer a 15-year free education (K-12). Each school can choose Cantonese or Mandarin as the medium of instruction when teaching Chinese language arts (Macao SAR Government, 2020b).

This study focused on teaching Chinese reading and writing in Macao's Chinese, English, and international schools which included teaching Chinese as a first or second language. It didn't include teaching Chinese to non-native Chinese speakers.

## *2.2 Participants*

The fifteen participants (one male) came from nine private schools which included four Chinese schools (7 teachers), two English schools (2 teachers), and three international schools (6 teachers). Teachers taught grades 1–3 or were program coordinators (the leaders of lesson preparation in teaching Chinese language arts). All participants were certified elementary grade teachers, while six teachers were also qualified to teach secondary schools. One third of the participants had a Master's degree. Teaching experience ranged from 3–34 years, averaging 15.

## *2.3 Procedures and Data Collection*

Program coordinators of Chinese language arts in three international schools, three English schools, and six Chinese schools (three Mandarin and three Cantonese medium) received an introductory letter explaining the nature and purpose of the study, inviting them and their experienced colleagues to participate in the study. The schools were selected because their teachers expressed that they were willing to engage in educational research. Twenty teachers responded. The teachers received a semi-structured interview outline and a participant consent form before the interviews. Finally, two focus group interviews (150 minutes each) and four individual formal interviews (70–120 minutes each) were conducted for the fifteen teachers who agreed to participate.

Interviews were recorded and typed into anonymous transcripts. During the processes of analysis, for member checking, I interviewed each participant several times whenever questions needed to be confirmed. School-based documents (i.e., reading book lists for grade level, guided reading questions to improve comprehension, and test papers), websites, and teaching materials were also analyzed to increase credibility. Several informal interviews with teachers and parents were conducted as another way for triangulation. The analysis results were

sent to participants to collect feedback to follow ethical practice and contribute to the credibility of the findings (Cronin, 2019).

### *2.4 Data Coding and Analysis*

I employed a grounded theory approach to analyzing the data. The analysis involved an interactive process of generating, developing, and verifying concepts following the coding procedures described by Corbin and Strauss (2008). I used methods of constant comparison, which included strategies for continually asking questions and making comparisons. During the process of beginning coding, microanalysis was used to break apart data and to delineate concepts to stand for ideas contained in raw data as a more detailed type of open coding. During the process of axial coding, I related concepts/categories to each other. During the process of integration, I linked categories around a core category to get a comprehensive explanation for the instructional practices. The findings of the previous analysis inspired me to add specific questions into the later interviews. New codes were added when they emerged and previous transcripts were reviewed and coded again until no new codes or categories could be identified (Cronin, 2019). School-based documents/materials were coded corresponding to the interview analysis.

For purposes of reflexivity, it is important to situate the researcher within the research to acknowledge their role in interpreting the findings. As a researcher, I have taught the teaching of Chinese language arts for 12 years and previously was a teacher and coordinator of Chinese language arts in an elementary school.

## 3 Findings

According to the participants' responses, each school has its own schedule for teaching Chinese language arts (Q1), extra arrangements for teaching reading (Q2), and regulations for curriculum and instruction (Q3) which couldn't be changed by individual teachers. Therefore, the data for answering questions 1, 2 and 3 are based on the schools (n = 9).

### *3.1 Time Scheduled for Teaching Chinese Language Arts*

During a normal week, students in all nine schools had their Chinese language arts classes every day (from Monday to Friday). However, the number of classes scheduled for teaching Chinese language arts in the Chinese, English, and international schools differed.

In the three international schools, students attended six Chinese language arts classes (40–45 minutes each) weekly.

In the two English schools, students attended six classes (35 minutes each) and seven classes (40 minutes each), respectively.

In the four Chinese schools, the number of Chinese language arts classes (40 minutes each) varied from 5–7 (i.e., 5, 6, 6, and 7).

In the four schools teaching Chinese in Cantonese including two Chinese schools and two English schools, there was also one more class for teaching Putonghua (i.e., teaching Mandarin and Pinyin spelling which was evaluated by oral tests).

## *3.2 Extra Time Devoted to Reading Instruction or Book Reading*

Three of the four Chinese schools arranged 1–2 (i.e., 1, 1.5, and 2) extra reading classes (40, 60, and 80 minutes) weekly. Their instructional practices are illustrated in section "Extra Reading Classes".

Additionally, each school arranged extra time or activities for students to read books in class or at home, although teachers did not teach students how to read during the extra reading time.

Both Chinese and English schools had extra time for reading with three different schedules: (1) one morning period every day (e.g., in-class reading at 10:50–11:05 a.m. in one English school); (2) one morning and one afternoon period: in one Chinese school, students read a self-selected book in the morning and read the same book in the afternoon accompanied by their teachers 25 minutes a day, 4 days a week; in the other Chinese school, students read 15 minutes in the morning for different subjects and read recommended Chinese books 30 minutes after lunch every day; (3) one morning or one afternoon period twice a week: three schools had students read books written in Chinese in the morning or afternoon (10, 15, and 20 minutes) twice a week.

In one Chinese school, students were required to read 30 books (written in traditional Chinese characters) from the recommended book list and write reports based on 20 books (i.e., summarize the story, introduce the author, record reflections, or use a picture to express feelings or opinions). The books were selected based on Macao students' life experience, the suggested reading materials listed in China's curriculum standards (Ministry of Education of the People's Republic of China, 2012, p. 41), and the preparation for PIRLS (more translated classical English literature was chosen). Picture books published in Taiwan were included because of the written language.

In one international school, a book week was scheduled once a year. During this week, many activities were designed to encourage reading (i.e., having students read one paragraph from a chosen book to their teachers; giving a book as an award

to students who could name a book they were reading when called at home by their teacher; requesting students to read at home with parents; inviting parents to hold a book talk or tell a story).

In all nine schools, students were encouraged to borrow books weekly/monthly. A reading corner or bookshelf was arranged for each classroom. The books were exchanged between classrooms or from the library.

Generally, more classes were scheduled for teaching reading and writing in Chinese schools (7–8.5 classes per week; this included teaching Chinese language arts, extra reading classes, and Putonghua).

## *3.3 Reading and Writing Instruction*

### The Medium of Instruction Chosen

Mandarin was used as the language of instruction by two Chinese schools and the three international schools because of its multi-functionality: as a mission to attract students/parents, to help students who cannot understand Cantonese, for improving students' writing (because Mandarin is the written language), and to benefit students' social life and admission to higher education in mainland China and Taiwan. Mandarin was taught from K1 in the five schools, and students were suggested not to speak Cantonese on campus from first grade.

Alternately, Cantonese was chosen as the medium of instruction by one Chinese schools and the two English schools because of circumstances. Namely, it was part of the school tradition, it was the students' mother language, and/or the teachers in the school lacked sufficient proficiency in Mandarin.

Diversity was another consideration. In one Chinese school, "*because teaching is a kind of art,*" each teacher was allowed to decide their medium of instruction (Mandarin, Cantonese or both) as long as the use of the language "*can bring the best instruction to children.*" In the international schools, speaking English in Chinese language arts class was not encouraged. However, English is not prohibited because students with parents who did not speak Mandarin may have needed more support in English. Teachers were also suggested to repeat or simplify sentences in Mandarin to enhance understanding. Students were also allowed to speak English when they couldn't express themselves in Mandarin.

In Mandarin medium schools, Pinyin spelling (recognition and handwriting) was taught from the first semester of first grade, which enables students to consult dictionaries with the Pinyin/radical/number of strokes system from the second grade. In Cantonese medium schools, which don't teach Pinyin from the first grade, consulting dictionaries with the radical/number of strokes system was taught in/after the third grade.

## Three Approaches Implemented in Chinese Language Arts Classes

### Textbook-Based Approach

In the eight schools where textbooks were adopted, including the three international schools, the progression of teaching Chinese followed the textbook series chosen. These textbook series were all developed for L1 learners.

Each lesson in grades 1–3 textbooks has at least one text (i.e., a poem, story, letter, narrative/descriptive/expository essay, diary, or script) followed by a vocabulary list, exercises and learning activities. Text reading was the core of the curriculum. Word recognition and handwriting were taught in coordination with each text to help students learn the meaning/usage of characters/words in the context of sentences. Some lessons might not be selected to be taught. Teachers used 3–5 class periods to teach a lesson. The time spent on teaching reading and writing (handwriting and any writing activities) for a lesson was about 50% each, whereas the teaching of speaking and listening was integrated with reading and writing instruction. Grades 1–2 teachers indicated that one third to one half of the time spent on teaching each lesson was focused on teaching single and compound character word recognition and handwriting. One teacher reported that she spent half of the time on teaching word recognition and handwriting in each class.

Generally, the procedure for teaching text reading in the Chinese language arts classes was: (1) reading text aloud by teachers or using recordings as demonstration; (2) reading text aloud by students; (3) having students roughly summarize the main ideas of the whole text; (4) teaching word recognition and vocabulary usage (i.e., discussing the meaning and pronunciation(s) of each introduced word or character, asking students to use a character to compose words or use a word to make sentences); (5) teaching handwriting: demonstrating the writing of strokes, stroke sequence, and the structure of components in a character; pointing out the radical of the character and the common errors in handwriting; having students trace the stroke sequence and write the character; (6) teaching reading comprehension: discussing the text in detail (including words, sentences, and paragraphs) and lead students to do the reading comprehension exercises such as identifying the main characters in the story, writing the main point of each paragraph within a supporting sentence, sharing the lesson which can be learnt from the story, finding alternative solutions to the problems encountered by the story's characters, and explaining the meaning of the figurative language in text; (7) introducing usage and language knowledge of punctuation, sentence and text structures, text types, grammar, rhetorical devices, dictionaries, authors and Chinese culture by leading students to do exercises/activities following the text.

A school-based reading comprehension book consisting of short texts and exercises was used in two Chinese schools from grade one and in one international school from grade two. Students were required to respond to multiple-choice questions, or answer orally, or write answers in words or sentences. Teachers read each question aloud for first grade students to control for word reading ability.

### Teaching Book Reading

Besides textbooks, literary works were used several times a semester in one Cantonese medium Chinese school and in the three international schools. Shared reading and assisted reading were used in class.

In the Chinese school, ten picture books and two chapter books were selected for first grade students to read during and after school. More chapter books were assigned to second grade and above students. Students were encouraged to predict what would happen next in the story read and to guess the meaning of unknown characters in the sentences (or skip them while reading). A reading award was given to the students who spent the most effort on reading books and writing the corresponding reports.

Due to the limits of word recognition, teachers in the international schools would read aloud books to primary grade students and then ask them to tell what the story was about to stimulate their interest in reading books. Students were encouraged to read a book they liked and then share the story in class by storytelling or role playing.

### Hybrid Approach

Apart from the other eight schools, in one English school, teaching character/word recognition and handwriting was the core of the first grade curriculum after the textbook-based approach was abandoned 6 years ago.

In the English school, two books containing collections of nursery rhymes and children's poetry were used for first and second grade teachers to intensively teach a group of characters/words (character-centered approach) in meaningful text reading (meaning-centered approach) as the implementation of a hybrid approach for teaching Chinese characters. Teachers were trained to teach more characters based on one simple or compound character in the text. It included teaching the characters by combining simple characters/components (日 [sun] + 一 [one; whole; once] = 旦 [morning; one day]), teaching the character by dividing a compound character (燒 [to burn; fever] – 堯 [Emperor Yao] = 火 [fire]), and teaching a group of characters which share the same radical (鳥 [bird], 鵝 [goose], and 鴨 [duck]). For the characters highlighted in texts for handwriting, the stroke sequence, and structure of the components was introduced to students with exercises in the books. Pictograph characters were introduced to students by using pictures to show the historical origin of how the characters were created. Word meanings (synonyms/antonyms) and parts of speech were covered in the books. Exercises for students to use the words were also included, (e.g., word or character association - write any words or characters which come to mind given a specific character; using the words/characters in the context of given sentences; using the given sentence structures to finish a poem).

Two chapter books were also used by first and second grade teachers to teach word recognition and handwriting; teach the differences between written language (Mandarin) and spoken language (Cantonese); teach story grammar and story lines;

teach other language knowledge (e.g., punctuation, sentence and text structures, rhetorical devices, dictionaries, and authors); teach reading comprehension and reading strategies. Teachers asked different levels of questions to facilitate reading comprehension and taught students how to find the answers (e.g., the answers can be found directly in lines; the answers can be worked out after summarizing the text or through inductive reasoning; students need to evaluate written materials with facts or reasons as support; students have to use their imagination to produce new ideas based on the materials read). Students in grade 2 and above were taught to ask themselves different levels of questions with teacher support. Students in grades 2–3 were assigned to read five chapter books per school year. For example, *The Journey to the West* [i.e., 西遊記, a Chinese classical novel] was for grade 3.

Commercial materials were used in grades 2–3 for teaching language and culture knowledge. Third graders also used the materials for learning reading comprehension (narrative and descriptive essays, poetry, and letters) and reading strategies. The reading materials were published in Hong Kong, except for some chapter books for Grade 3 and above students published in Taiwan.

## Writing Instruction

Writing classes (40–80 minutes each) were scheduled three to five times a semester within the allotted teaching hours of Chinese language arts in all nine schools. The writing tasks were: (1) grades 1–3: writing sentences (with prompts, i.e., words and one picture; given model sentences; given conjunctive words such as first…then…) and reorganizing given words and punctuation marks into a sentence; (2) grades 2–3: writing paragraphs (students chose three or more words from a word bank to write a paragraph; given a model paragraph; with prompts, i.e., words and four pictures); (3) third grade: writing narrative essays with an assigned topic (about 100 characters). In one Mandarin medium Chinese school, students had to write children's poems following model poems in textbooks from the second grade.

When teaching writing, teachers explained the requirements for the writing task, discussed with students how to finish the tasks, had students share their sentence/paragraph or describe a picture(s) orally and offered advice on modifying before translating. Sometimes students were suggested to discuss how to write a story or essay on the assigned topic with parents before the writing class. Discussion and peer review were not allowed during transcription in Chinese and English schools due to the consideration of fairness in scoring. Teachers would model the writing of unknown characters for each student. Revising the wrongly chosen or incorrectly written characters and sentence structure mistakes was required. One English school teacher had her students copy their essays with her revisions every time. Teachers would read good works written by students in class if time permitted.

Giving praise, playing games, and allowing to experience were often done in writing classes in the international schools. For example, students were commended after sharing the sentences they made; games such as cutting sentences into words and putting the words into a box for students to pull out to make new sentences were often used; students were allowed to make funny sentences (e.g., the boy is singing

while on the toilet); and students were invited to taste a slice of lemon and share their feelings while other students observed their facial expression before writing. One international school teacher led her students to experience the process of creating their own small book by reading a picture book they liked, choosing papers and materials, imitating the story read or creating a new story, drawing pictures, and sharing the book with classmates.

### Extra Reading Classes

Three of the four Chinese schools arranged extra reading classes. Their reading instructional practices were influenced by PIRLS.

Of these, one Mandarin medium school had two reading classes a week and one third of the total class time was used to teach calligraphy. Reading materials included: literary works (i.e., 15 picture books for grades 1–2 students to read in a year; short story collections and chapter books for third graders) and a reading comprehension book with texts (i.e., narrative, descriptive, and expository essays, diaries, letters, and notes) and exercises. When teaching picture book reading, a book was presented by using a projector with the title covered to invite students to predict the title from the picture on the cover; teachers asked students to predict what would happen next in the story and vividly role play the story orally. Teachers asked students four levels of questions developed by the Comprehension Processes of PIRLS (see Mullis & Martin, 2015, pp. 18–22). Students were required to write reflections such as what you have learnt from this book, which character/word you appreciate the most and why, or draw a picture of the favorite part of the book.

Another Mandarin medium school had 1.5 reading classes a week. The reading materials included: Chinese classical literature (i.e., ancient poetry for each grade; "Standards for Being a Good Student and Child" for grades 2–3; and "Three-Character Canon" for grades 3–4) and released assessment questions of PIRLS (for third grade). Informative writings such as airplane tickets were added into reading materials. Modern Chinese was used to explain the classical literature. Videos or pictures were presented to facilitate comprehension. Students were required to read aloud, chant, recite, and draw a picture to show the meaning of the poem. For third graders, the meaning of key words in each poem would be tested. Storytelling and role play were used to teach"Standards for Being a Good Student and Child," if time permitted.

The third school allowing either Mandarin or Cantonese to be used had one extra reading class per week. A series of reading comprehension books were the main reading materials while picture books were also used. The reading comprehension books consisted of short texts (narrative and descriptive writings for grades 1 and above and expository writings for grades 2 and above) and questions to facilitate retrieving information, making inferences, evaluating and thinking creatively, and associating and writing imaginatively. The questions teachers asked in the reading classes and the tests were developed based on the Comprehension Processes of PIRLS, although it was not required.

## Assignments

Four types of assignments were reported. First, previewing the text. Students were encouraged to read the text aloud, circle unknown characters or words in the text, and answer the reading comprehension questions related to the text at home. In one international school, grades 2–3 students were suggested to consult dictionaries to add Pinyin to each introduced word. Second, in-class assignments and correspondent homework. For example, copying words and characters, using a word(s) to make a sentence, using a sentence structure type or one kind of rhetorical device to make a sentence, putting the words in the correct order to complete the sentence, and selecting and writing the appropriate words in a sentence's context were common assignments in different schools. In Mandarin medium Chinese schools, recognizing, reciting, and writing the initial consonants and simple or compound vowels, and the Pinyin spelling were also assigned as homework. Third, reviewing and preparing for a test. For example, reading texts aloud, copying questions and their answers, and self-studying the writing of characters and words in the text. Fourth, reading extracurricular books along or with parents, which included online reading. Sometimes students were requested to write in response to reading. Most of the written homework was checked by teachers and students were required to make corrections.

## Evaluation

Chinese and English schools had their students take tests from the first semester of grade 1, but teachers could read aloud the reading comprehension text and the correspondent questions in semester one only. In the international schools, students don't have tests until the second semester.

Three kinds of tests were arranged: (1) weekly or monthly quizzes: e.g., dictation of words/paragraph(s), writing from memory (sentence/paragraph/poem), or reading aloud from memory; (2) written exams twice a semester: students were required to write the characters corresponding to the given Pinyin (Mandarin as medium of instruction), write the radical for the character, count the number of strokes in a character, use words appropriately in the sentence's context (e.g., noun, verb, adjective, quantifier, or idiom), correct the wrongly chosen characters in the sentence, put the words in the correct order to complete the sentence, make sentences, use a type of sentence structure or rhetorical device to rewrite or expand a sentence, punctuate sentences, answer questions from the textbook, do reading comprehension, and take writing tests (write sentences/an essay from prompts; writing tests were only assigned in the Chinese and English schools); (3) oral tests: in one Chinese and two English schools, students had to read aloud a text with expression, tell a story, or describe a picture(s) each semester.

### Group Learning

Some of the participants indicated that group learning was used for the discussions of the meaning of words/characters, pronunciation/Pinyin spelling, role play, and learning written and spoken language knowledge. However, one teacher reported that her students never learn in groups due to lack of space.

### Using Internet-Based Information Technology

The frequency of internet-based information technology usage differed among schools. In some schools, internet-based IT was used regularly to teach word recognition and handwriting (i.e., using videos to explain the meaning of a character/word, demonstrating the stroke and stroke sequence, and recording each student's stroke sequence for evaluation), offer individual picture book reading, upload audios/videos to demonstrate the reading of texts, and post online reading comprehension tests to record students' performance and calculate their scores. In some schools, internet-based IT was not frequently used because of a lack of equipment, slow internet connection, time-consuming borrowing procedures, and students' underdeveloped IT skills.

### What Shaped Instructional Practices

***Curriculum Standards*** The three international schools designed their reading and writing curriculum based on Macao's "Basic Academic Attainments (BAA) (second language)", while the BAA (primary/first language) was followed by the four Chinese schools and the two English schools (Education and Youth Development Bureau, 2016a, b). The BAAs are the basic and minimum standards. In one of the international schools, teachers were required to teach students who have different abilities or learning goals in the same class with different evaluation standards.

In one Chinese school, the curriculum standards of mainland China (Ministry of Education of the People's Republic of China, 2012) were also considered (i.e., recognize more and write fewer characters; a book reading list).

***Preparation for College Entrance Examinations*** In one international school, the curricula were also designed based on the International General Certificate of Secondary Education (IGCSE, Chinese as First, Second, and Foreign Language).

***Preparation for School Tests*** The tests' content and the scoring rubrics decided the instructional practices. For example, in one English school, peer review was not applied in reading or writing classes for fairness in scoring; texts which wouldn't be included in tests might not be taught; listening, speaking, and reading aloud wouldn't be evaluated individually because those items weren't scored.

***Love for Reading and Writing*** Teachers from Chinese, English, and international schools mentioned: (1) focusing on helping students love reading and writing through teaching strategies (e.g., use imagination to memorize how to write a character), forming habits, and helping students succeed (e.g., select interesting texts or interactive games to facilitate learning characters) is more important than what knowledge was taught because teachers cannot inculcate knowledge into students; (2) only when a teacher truly loves reading and writing, can the teacher better inspire students.

***Others*** The reading and writing instruction of the participants has been changing over the years for the following reasons: (1) parents' feedback: teachers reduced the amount of homework, slowed the speed of progress but high-achieving students were assigned extra worksheets or book reading, and low-achieving students were offered optional free remedial instruction (e.g., a 40-minute class twice a week after school to reteach reading and writing skills); (2) principal's decisions: for example, a principal requested teachers to change teaching methods and materials to raise students' interest in learning Chinese characters and ability to self-study; (3) school-based in-service training (i.e., textbook selection, teaching methods, and integration of Internet with instruction): with experts' support, teachers' knowledge and skills in teaching reading and writing expanded; (4) external circumstances: PIRLS affected the materials read (text type and length) and reading strategies taught in Chinese and English schools. The Macao Education and Youth Development Bureau first published its own textbook series in the 2020/2021 school year, and three of the eight schools where participants taught selected this series while textbooks published in Hong Kong were selected by all eight schools in the previous year.

## *3.4 Challenges in Teaching Reading and Writing*

***Teaching Characters Following Textbooks*** Characters were not taught in a systematic way in the chosen textbooks. The commonly used characters and the single-component characters were not introduced first increasing the difficulties in teaching reading and writing. The total number of characters which textbooks highlighted for word recognition/handwriting didn't meet the curriculum standards. Teachers had to develop additional materials for word recognition and handwriting to make up for the gaps.

***Limited Character Recognition*** Students couldn't read independently or express their feelings or opinions. Students often had to ask teachers for help in writing characters and were unable to use synonyms with fine nuance of meaning.

***Resistance to Handwriting*** Lots of handwriting was assigned to primary grade students because they need to recognize and use hundreds of characters each year.

Self-motivated handwriting practice was expected for test preparation. However, the heavy workload made students resist handwriting even if their grades would suffer as a result. The international school students have even lower motivation.

***Limited Time for Teaching Writing Compositions*** The time scheduled for narrative writing in third grade was not enough to teach how to plan, translate, or review their writings. One English school teacher reported that questions were usually posted by teachers to remind students what to write in each paragraph; however, students wrote to answer questions instead of writing a coherent composition.

***Influence of Other Languages Used*** Adding Pinyin above each character couldn't help students who don't speak Mandarin as their first language to read the text in Mandarin independently. More demonstrations of character pronunciation and more text reading aloud were necessary. Teachers also had to spend more time on teaching vocabulary and sentence structure due to the grammatical and lexical differences between the spoken language and written Mandarin. Teachers who taught at the Chinese school where Mandarin and English were both emphasized from the first of 3 years of kindergarten stated that Mandarin sentence structures of students with high proficiency in English were influenced by English grammar.

***The Differences in Parents' Involvement and Expectations*** While teachers expected parents to guide children to do homework, prepare for tests, read printed books, and do online reading, many parents were unable to do so due to their lack of knowledge or their working schedules. English and international school teachers also had to communicate with parents with different expectations regarding their children's literacy in Chinese. Some parents expected their children to write many characters in the first month of grade one, while some foreign parents didn't want their children assigned any homework because they couldn't help them. Some parents expected their children to read more English books.

## 4 Discussions

Hsiang and Graham (2016) reported that textbooks and school guidelines played an important role in developing Macao grades 4–6 writing programs. Participants in the current study indicated that grades 1–3 reading and writing instruction in eight of nine Macao schools followed the textbooks. These findings show that selection and use of textbooks is a critical topic in teacher training, especially if the total number of introduced characters in textbooks does not meet the number recommended by the curriculum standards (Hsiang et al., 2021; Smart & Jagannathan, 2018).

This study extended Hsiang and Graham's survey (2016) to explain how school guidelines shaped reading and writing instruction. The curriculum and instruction

of reading and writing were decided by schools' aims (i.e., Chinese, English, or international school; the medium of instruction; curriculum standards; college entrance examinations and school tests; teaching materials and scoring rubrics) and physical environment (i.e., classroom size; support system for online instruction). Moreover, principals' decisions, parents' feedback, school-based training, international assessment (i.e., PIRLS), and textbook policy (a new textbook series published by the local government) could make changes.

This study provided evidence that a balanced literacy approach was adopted by the participants.

First, in order to prepare students well for meeting national and international standards (including PIRLS), teachers decided what to teach from a cross-cultural perspective: (1) striking a balance between reading narratives and informative writings although poetry and narratives are the traditional texts for Chinese primary grade students (Lam, 2011); (2) translated classical English literature was added into reading lists (Ministry of Education of the People's Republic of China, 2012), and a large number of modern translated English books published in Taiwan (written in traditional Chinese characters) were read in class; (3) self-selected reading was encouraged whether the book is written in traditional or simplified Chinese characters, and writing simplified Chinese characters is acceptable as long as they are written correctly. A majority of the participants not only "taught to the test", but also used different ways to inspire students' love for reading and writing Chinese.

Second, whether textbooks were used or not, authentic literary activities were offered in this study. It shows that literary works were used to develop reading/writing abilities in rural/urban schools in China (Chan, 2020; Yang et al., 2016).

Third, instruction also maintained a balance among character/word recognition, handwriting, vocabulary usage, and reading comprehension from grade one in each school. To increase word recognition, both character-centered approach and meaning-centered approaches were adopted by participants (Lam, 2011).

Fourth, literary knowledge and strategies were both taught to help students succeed. School-based curriculum developments and PIRLS promoted the reform of teaching reading and writing (i.e., from asking students questions to teaching students how to find answers to questions of different comprehension processes and self-questioning).

Fifth, teachers tried to balance teacher-centered and student-centered approaches. Personalized reading (i.e., read self-selected books and online reading) and differentiated instruction (i.e., worksheets or test papers were designed for different level students) were implemented in some classrooms to meet the needs of individual students (International Literacy Association, 2018).

However, concerns about writing instruction are raised. This study and previous surveys (Hsiang & Graham, 2016; Hsiang et al., 2018, 2020a) indicated that Chinese writing was taught infrequently in many grades 1–9 classrooms. Writing strategies were also seldom mentioned by participants in the current study. Obviously, more time scheduled for teaching writing compositions through process writing is necessary (Cutler & Graham, 2008; Graham & Hebert, 2011).

Another issue is the difficulties in teaching Chinese characters. Both reading and writing performance were reported to be influenced by the limited character recognition. However, some students resisted handwriting even though it is an effective way for both native and non-Chinese-speaking children to recognize characters (Liu & Liu, 2020; Wang et al., 2018). Research on teaching handwriting and how to empower young children to recognize and write characters is necessary (Kong, 2020).

The last concern is the needs of young leaners in local international schools who are native Chinese speakers as they are taught Chinese as a second language. In Mainland China, Hong Kong, and Macao, many international schools accept Chinese students who have local passports. However, most studies on TCSL/TCFL have focused on teaching international students/adults (Li, 2020). Further studies can focus on primary grade CSL instructional practices and the development of CSL textbook series for Chinese students in local international schools to increase motivation and performance of reading and writing.

## 5 Conclusion

In this study, students normally had at least one Chinese language arts class every day from Monday to Friday. The number of Chinese language arts classes scheduled in different schools varied from 5–7 (35–45 minutes each). Whether textbooks were used or not, grades 1–2 teachers spent a lot of time on teaching word recognition and handwriting to facilitate reading and writing in Chinese languages classes. Compared to the English and international schools, the Chinese schools scheduled more time for teaching Chinese reading and writing (this included teaching Chinese language arts, extra reading classes, Putonghua, and free reading in school). The medium of instruction chosen (Mandarin/Cantonese) influenced how reading and writing were taught. A balanced approach to teaching Chinese reading and writing was adopted by the 15 grades 1–3 teachers in all schools, although what they focused on varied. PIRLS affected reading instruction in Chinese and English schools. The difficulties in teaching reading and writing were: teaching characters following textbooks, limited word recognition, resistance to handwriting, limited time for teaching writing compositions, influence of other languages used, and the differences in parents' involvement and expectations. It reveals that reading and writing are influenced by institutions, society, culture, and politics (Graham, 2018). Research on how to balance reading and writing instruction through school-based curriculum development is needed since instructional practices are influenced by school regulations.

**Funding** This research was supported by the Higher Education Fund of the Macao SAR Government [14/DSESHSS-UM/2019].

## References


Chan, U. M. (2020). 運用繪本與自我調整策略發展培養寫作能力 [Using picture books with SRSD to develop writing ability: A case study in a second grade school leisure activities course]. [Unpublished master's thesis]. University of Macau.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing Grounded theory* (3rd ed.). Sage.

Cronin, C. (2019). Openness and praxis: Exploring the use of open educational practices in higher education. In S. B. Merriam & R. S. Grenier (Eds.), *Qualitative research in practice: Examples for discussion and analysis* (2nd ed., pp. 215–234). Jossey-Bass.

Cutler, L., & Graham, S. (2008). Primary grade writing instruction: A national survey. *Journal of Educational Psychology, 100*(4), 907–919. https://doi.org/10.1037/a0012656

Education and Youth Development Bureau. (2016a). 小學教育階段中文基本學力要求 (第一語文即教學語文) [The requirements of basic academic attainments in Chinese language for primary education (primary language, that is, teaching language)]. *Boletim Oficial da Região Administrativa Especial de Macau* [Official Bulletin of the Special Administrative Region of Macao], *9*, 126–137.

Education and Youth Development Bureau. (2016b). 小學教育階段中文基本學力要求 (第二語文) [The requirements of basic academic attainments in Chinese language for primary education (second language)]. *Boletim Oficial da Região Administrativa Especial de Macau* [Official Bulletin of the Special Administrative Region of Macao], *9*, 137–144.

Education and Youth Development Bureau. (2021). *Educational statistic: Primary education*. Retrieved February 14, 2021, from https://portal.dsedj.gov.mo/webdsejspace/internet/Inter_main_page.jsp?id=8525

Fan, M.-S., Wu, F.-T., & Wang, Yu (2016). 数字阅读: 电子书对小学生语文阅读能力的影响 [E-reading: Influence of e-books on reading ability of students in primary school]. 电化教育研究 [e-Education Research], *12*, 106–128.

Fisher, D., Frey, N., & Akhavan, N. (2020). *This is balanced literacy: Grades K-6*. Corwin.

Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educational Psychologist, 53*(4), 258–279. https://doi.org/10.1080/00461520.2018.1481406

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744.

Graham, S., Harris, K. R., Mason, L., Fink-Chorzempa, B., Moran, S., & Saddler, B. (2008). How do primary grade teachers teach handwriting? A national survey. *Reading and Writing, 21*(1), 49–69. https://doi.org/10.1007/s11145-007-9064-z

Graham, S., Harris, K. R., & Adkins, M. (2018a). The impact of supplemental handwriting and spelling instruction with first grade students who do not acquire transcription skills as rapidly as peers: A randomized control trial. *Reading and Writing, 31*(6), 1273–1294. https://doi-org.libezproxy.um.edu.mo/10.1007/s11145-018-9822-0

Graham, S., Liu, X., Aitken, A., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2018b). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly, 53*(3), 279–304. https://doi.org/10.1002/rrq.194

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018c). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research, 88*(2), 243–284. https://doi.org/10.3102/0034654317746927

Hsiang, T. P. (2012). 小學語文科詩歌教材的比較研究:以中國內地與澳門版教科書為例 [A comparative study on primary school poetry teaching materials with examples of Chinese language textbooks edited for Mainland China and Macao]. *Journal of Macau Studies, 3*, 145–160.

Hsiang, T. P., & Graham, S. (2016). Teaching writing in grades 4 to 6 in urban schools in the Greater China region. *Reading and Writing, 29*(5), 869–902. https://doi.org/10.1007/s11145-015-9597-5

Hsiang, T. P., Graham, S., & Wong, P. M. (2018). Teaching writing in grades 7–9 in urban schools in Chinese societies in Asia. *Reading Research Quarterly, 53*(4), 473–507. https://doi.org/10.1002/rrq.213

Hsiang, T., Graham, S., & Yang, Y. (2020a). Teachers' practices and beliefs about teaching writing: A comprehensive survey of grades 1 to 3 teachers. *Reading and Writing, 33*(10), 2511–2548. https://doi.org/10.1007/s11145-020-10050-4

Hsiang, T. P., Wang, Z., & Gong, Y. (2020b). 雙語教育下小學漢字教學的挑戰與因應: 澳門的研究發現與啟示 [Facing the challenges of teaching Chinese characters in bilingual primary schools: A study of Macao and its implications]. 臺灣教育評論月刊 [Taiwan Educational Review Monthly], *9*(10), 60–70.

Hsiang, T. P., Graham, S., Wang, Z., & Gong, Y. (2021). How Chinese characters are taught: An analysis of three popular textbooks used in Macao. *Technology, Knowledge and Learning, 26*(2), 355–381. https://doi.org/10.1007/s10758-021-09492-8

International Literacy Association. (2018). *Standards for the preparation of literacy professionals 2017*. Retrieved December 2, 2020, from https://literacyworldwide.org/get-resources/standards/standards-2017

Kong, M. Y. (2020). The association between children's common Chinese stroke errors and spelling ability. *Reading and Writing, 33*(1), 635–670.

Lam, H. C. (2011). A critical analysis of the various ways of teaching Chinese characters. *Electronic Journal of Foreign Language Teaching, 8*(1), 57–70.

Li, M. (2020). A systematic review of the research on Chinese character teaching and learning. *Frontiers of Education in China, 15*(1), 39–72. https://doi-org.libezproxy.um.edu.mo/10.1007/s11516-020-0003-y

Liu, Y., & Liu, D. (2020). Morphological awareness and orthographic awareness link Chinese writing to reading comprehension. *Reading and Writing, 33*, 1701–1720. https://doi.org/10.1007/s11145-019-10009-0

Macao Government Tourism Office. (2020). *About Macao*. Retrieved July 5, 2021, from https://www.gastronomy.gov.mo/#about-macao

Macao SAR Government. (2020a). *2017 Macao in figures*. Retrieved July 5, 2021, from http://www.dsec.gov.mo/CMSPages/e_mn_indicator.aspx?IACM=1

Macao SAR Government. (2020b). *About Macao SAR: Fact sheet-education*. Retrieved July 5, 2021, from https://www.gcs.gov.mo/files/factsheet/Education_EN.pdf

Ministry of Education of the People's Republic of China. (2012). 义务教育语文课程标准2011年版) [Compulsory Education Chinese Language Curriculum Standards (2011 ed.)]. Beijing Normal University.

Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 assessment framework* (2nd ed.). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/pirls2016/framework.html

Quan, G., Dong, N., Chen, J., & Chen, M. (2019). 基于思维导图的情感效用阅读实证研究 [Reading with mind map: Empirical research on emotional utility]. 中国电化教育 [China Educational Technology], *3*, 55–63.

Roe, B. D., & Smith, S. H. (2012). *Teaching reading in todays' elementary schools* (11th ed.). Wadsworth.

Ruan, Y., Qi, C., & Liu, X. (2018). 多重感觉通道训练对听写落后儿童的教学 [Multisensory training as a teaching method on Chinese children with spelling difficulties]. 中国临床心理学杂志 [Chinese Journal of Clinical Psychology], *26*(4), 821–825.

Smart, A., & Jagannathan, S. (2018). *Textbook policies in Asia: Development, publishing, printing, distribution, and future implications*. Asian Development Bank. Retrieved from https://www.adb.org/sites/default/files/publication/478946/textbook-policies-asia.pdf

Tse, S. K., Marton, F., Ki, W. W., & Loh, E. (2007). An integrative perceptual approach for teaching Chinese characters. *Instructional Science, 35*, 375–406. https://doi.org/10.1007/s11251-006-9011-4

Wang, J., & Leland, C. H. (2011). Beginning students' perceptions of effective activities for Chinese character recognition. *Reading in a Foreign Language, 23*(2), 208–224.

Wang, C.-C., Hung, L.-Y., Chang, Y.-W., & Chen, H.-F. (2008). 一到九年級學生國字識字量發展 [Number of characters school students know from grade 1 to G9]. 教育心理學報 [Bulletion of Education Psychology], *39*(4), 555–568.

Wang, Y., McBride, C., Zhou, Y., Joshi, R. M., & Farver, J. A. M. (2018). Becoming literate in Chinese: A comparison of native-speaking and non-native-speaking children. *Journal of Research in Reading, 41*(3), 511–524. https://doi.org/10.1111/1467-9817.12122

Warner-Griffin, C., Liu, H., Tadler, C., Herget, D., & Dalton, B. (2018). *Reading achievement of U.S. fourth-grade students in an international context*. National Center for Educational Statistics (NCE). https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2018017

Wei, X.-F., Yang, F., Shi, X., & Qiao, K.-D. (2020). 协作思维导图策略促进小学生习作的行动研究 [Action research on the collaborative mind mapping strategy to improve primary school students' writing]. 现代教育技术[Modern Educational Technology], *30*(6), 47–54.

Wong, L. (2013). 識字教學行動研究：以融合教育模式中的讀寫障礙生為例 [Literacy instruction by action research : Integrated education model for dyslexic students] [Unpublished master's thesis]. University of Macau.

Xie, Y., Wu, L., Li, H., Kwok, L., Huang, Y., Xiao, L., & Yang, Y. (2016). 智慧学习环境下小学语文阅读课生成性教学路径的探究 [Exploration of generative teaching approach for primary Chinese reading class in the smart learning environment]. 中国电化教育 [China Educational Technology], *6*, 36–42.

Yang, B., Zhang, L., & Li, X. (2016). 故事书阅读中考核与注音对农村一年级儿童识字量的影响 [The effect of examination and Pinyin notation during storybook reading on first-grade children's Chinese characters literacy in rural areas]. 中国特殊教育 [Chinese Journal of Special Education], *6*, 84–90.

Zhang, Y., Sun, J., & Mutlf. (2018). 基于知识建构的同伴互评教学实践研究 [Research on Peer Review Based on Knowledge Building]. 电化教育研究 [e-Education Research], *7*, 108–113.

# Part VI
# Writing Practices by Adult Learners

# Postsecondary Developmental Education in Writing: Issues and Research


**Charles A. MacArthur**

**Abstract** In the United States, two-year public community colleges with open admissions and low tuition offer the opportunity for students to attend college who would otherwise not be able to afford it or to be accepted. However, students who are evaluated as underprepared for college are required to take noncredit courses in writing, reading, and or math, termed developmental courses. These courses expand opportunity, but completion rates are low, and required developmental courses may be a barrier to progress for many students rather than an opportunity. This chapter reviews research relevant to recent efforts to reform community college developmental writing programs, including research on the concerns driving the reforms and research on the effects of various reform efforts. Major reform efforts have included changes in criteria and assessments for placement into developmental writing; accelerated progress through developmental courses with compressed courses, summer bridge programs, or combined reading and writing courses; and corequisite courses that integrate developmental and regular first-year courses. In contrast to these changes in placement and the structure of courses, little research has focused on improved pedagogical methods. As part of the discussion of pedagogical research, I will discuss our own research on strategy instruction with self-regulation. The chapter will conclude with recommendations for further research.

**Keywords** College developmental education · College success · Policy reforms · Writing strategy instruction

C. A. MacArthur (✉)
University of Delaware, Newark, DE, USA



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_18

# 1 Postsecondary Developmental Education in Writing: Issues and Research

In the United States there are over 1000 community colleges serving six million students, about 39% of all undergraduate students (Snyder et al., 2019). In fall of 2016, there were nearly 1,000,000 first-time community college students, including about 600,000 full-time students and 360,000 part time students. Completion rates for community college students are not high overall. A longitudinal study of students beginning college in 2012 (Chen et al., 2019), found that 62% of the full-time and 45% of the part-time community college students continued to take courses a year later in fall 2017. After 6 years, 39% had earned some credential: 18% earned an associate degree, 13% earned a bachelors, and 8% earned a certificate. These statistics have been fairly consistent over time. For example, 36% of students starting in 1995 earned some credential. In comparison, completion rates after 6 years for four-year colleges are 59% of students earning a bachelor's degree and 8% earning an associate degree or certificate. It's important to keep this general context for community colleges in mind in thinking about developmental education because the problem of college completion is not limited to developmental education students. Community colleges face many challenges. By design, they serve students who might not otherwise be able to go to college including low-income students and minorities (Perin, 2006). The colleges face economic challenges themselves and, partly as a result of those challenges, rely on adjunct faculty, with two-thirds of courses taught by adjuncts (Goldrick-Rab, 2010).

Substantial proportions of community college students are required to take developmental courses in reading, writing, or math, although statistics vary across studies. A study based on analysis of transcripts from 2003 to 2009 (Chen, 2016) found that 68% of students took at least one developmental course and 48% took two or more; on average these students took 2.9 developmental courses. The largest proportion, 59%, took at least one developmental math course, while 28% took developmental English. Similar results were found by Bailey et al. (2010) using a large database from the Achieving the Dream project; 59% of students were referred to developmental math courses and 33% to developmental reading courses. Both studies found that assignment to developmental education courses was more common for low-income, Black and Hispanic, and first-generation college students.

## *1.1 Outcomes of Developmental Education*

Overall, students who are referred to developmental education are less likely to graduate (Bailey et al., 2010; Chen, 2016). However, students who finish the developmental education sequence of courses and enroll in credit courses do as well as non-remedial students (Bahr, 2010; Bailey et al., 2010; Chen, 2016). They are equally likely to pass a gatekeeper course in math or English, and they are equally

likely to get a degree or to transfer. Some would say that is the goal of developmental education, and the results indicate that it is successful (Goudas & Boylan, 2012). However, most students do not finish the sequence of required developmental education courses. According to data from the Achieving the Dream project (Bailey et al., 2010), 30% of students who were referred to developmental reading never enrolled in such a course. Of those who did enroll in developmental education, 16% failed the course, another 8% passed but never enrolled in the credit course overall, and only 37% of the students originally referred to developmental reading passed a related college credit course. Another study at a community college with a high proportion of minority students (Nastal-Dema, 2019), found that only a small proportion of students who took a first developmental writing course ever passed first-year composition; 56% passed their developmental education course; of those, 2/3 took the next level; of those, 3/4 passed that level; of those, about 3/4 enrolled in first year composition (FYC); and of those, 63% passed for an overall pass rate in FYC of just 12%.

How do we interpret these descriptive findings? One interpretation is to consider that students who are referred to developmental education have lower achievement, lower economic status, and more barriers in their lives, so they don't do as well. Under that interpretation, every student who makes it is evidence of the success of developmental education. The other interpretation is to consider that developmental education itself may be a barrier that discourages students and takes time and money without leading to college credit. The descriptive studies mentioned above don't control for differences adequately to decide whether developmental education has positive or negative effects and for whom.

Regression discontinuity studies control for background variables in students by comparing students who fall right below and above the cut point for developmental education placement. Most community colleges have used standardized test scores to place students in developmental education (Rutschow et al., 2019). The students right above and below the cut point are likely to be very similar in background variables. A meta-analysis (Valentine et al., 2017) of 11 studies that used regression discontinuity to examine the effect of placement in developmental education found that students who fell below the cut point and were assigned to developmental education were less likely to pass the first college credit course in that area: 78% of students who did not take developmental education, versus 68% of students who did, passed that first credit course. They also found small differences in the number of college credits earned and whether or not students earn some sort of credential.

One of the limitations of this research design is it only compares students who are just above and below the cutoff score; students who are further below the cutoff may still benefit from developmental education. One study (Boatman & Long, 2018) used regression discontinuity to examine potential benefits for students with lower skills placed in lower-level developmental reading and writing courses. They compared outcomes for students just above and below the cutoff between the lower- and higher-level developmental courses. The study found positive effects of placement in lower-level courses on persistence to a second year and on number of college credits earned.

Additional evidence on the overall effects of developmental education placement comes from states that have attempted reforms in placement. Florida made developmental education courses optional for students beginning in 2013. Fortunately for researchers looking at the outcomes (Woods et al., 2019), students had taken standardized achievement tests in high school, so it was possible to know which students would have been required to take developmental education prior to the policy making them optional. Of the students eligible for developmental writing, one half opted to skip it and take FYC, one quarter took developmental writing, and one quarter did not take any English course in the first year. Interestingly, over 60% of those who took FYC passed the course. This indicated that many students who would have been placed in developmental education could be and, in fact, were successful in FYC.

To summarize, the descriptive data show, first of all, that a majority of community college students are referred to some developmental education course. About 30% are required to take developmental education in writing and or reading. In addition, students who persist through the developmental education courses do relatively well in FYC, but relatively few developmental education students ever take and pass FYC. Furthermore, as shown by the regression discontinuity studies and the results of the policy change in Florida, many students assigned to developmental education could be successful in college credit courses without taking the developmental education courses. On the other hand, many students probably do need extra support, but there may be ways to provide support other than assignment to developmental education. We turn our attention now to efforts by community colleges and state agencies to reform developmental education practices and policies.

## 2 Reform Efforts

In response to concerns about low graduation rates and, in particular, to dropout rates among students assigned to developmental education, community colleges and state agencies have initiated a variety of reforms to policies and practices (Hodges et al., 2020). One reform is to switch from reliance on a single placement test, like Accuplacer (College Board, 2016) or Compass (ACT, 2006), to the use of multiple measures (MM), usually including high school grade point average (GPA) and often other measures. This reform directly addresses findings from regression discontinuity studies that many students just below the assessment cutoff could be successful in FYC. Another reform aims to accelerate progress through developmental courses by offering short courses, combining levels of developmental writing, or integrating developmental reading and writing. Another popular reform option is corequisite courses, in which students take FYC simultaneously with a support course with the same instructor, or sometimes other forms of extra support. A practice guide from the What Works Clearinghouse (Bailey et al., 2016) found sufficient evidence to recommend use of multiple measures for placement and accelerating progress by combining developmental courses or using corequisite courses.

Most reform efforts are structural changes to placement and the organization of courses. Less research has focused on improvements in instructional design to increase student success. In fact, the practice guide (Bailey et al., 2016) made only one pedagogical recommendation, to teach self-regulation, and it was supported by a single study in mathematics. In the next few sections, we discuss research on each of these directions for reform: multiple measures for placement, accelerated progress through developmental courses, and corequisite courses. After that, we will discuss our own pedagogical research on strategy instruction in developmental writing courses.

## *2.1 Multiple Measures for Placement*

Nearly all community colleges use placement tests to evaluate whether students are prepared for college math and writing (Rutschow et al., 2019). Typically, they have accepted SAT or ACT scores if students have taken them, or required students to take placement tests designed for the purpose like Accuplacer (College Board, 2016). However, these tests are relatively poor predictors of success in key first year courses, and they are substantially worse predictors than simple high school grade-point average (HS GPA). Two reviews (Belfield & Crosta, 2012; Scott-Clayton, 2012) found correlations of Accuplacer and Compass (ACT, 2006) with course grades of .08 to .18, compared to correlations of FYC grades with HS GPA of .24 to .35. Adding information from high school transcripts about courses taken, and even non-cognitive measures, may enhance correlations further. Both community colleges' practices and statewide policies have begun to require use of multiple measures for placement. A recent report from the Education Commission of States (Whinnery & Pompelia, 2018) found that 19 states encourage or mandate colleges to consider multiple measures in placement. A 2016 survey of 2-year and 4-year colleges (Rutschow et al., 2019) found that although 99% of community colleges continue to use placement tests, 40% also use HS GPA, 22–30% consider students' high school courses, and 15% use non-cognitive measures of motivation.

Seven community colleges in the State University of New York (SUNY) system participated in a randomized control trial (RCT) of a multiple measures (MM) placement system (Barnett et al., 2020). Using data on former students' placement tests, HS GPA, other measures that varied by college, and their success in college courses (e.g., FYC), colleges developed algorithms for placement. Students were randomly assigned to be placed by the placement test alone (business as usual, BAU) or by MM. For writing, 80% of MM students and 46% of BAU students were placed in FYC. Although differences between the groups in passing FYC declined over time, the differences in enrollment and passing FYC remained significant after 3 semesters; MM students were 5.3% more likely to enroll in and 2.9% more likely to pass FYC than BAU students. However, although 80% of students were placed in FYC and 71% enrolled, only 47% had passed it after 3 semesters. Note that for

math, initial increases in enrollment and passing a college math course during the first semester were no longer significant by the second semester.

Another RCT of MM placement was conducted in Minnesota (Cullinan et al., 2019). Only results from the first semester have been reported to date. The MM placement system first exempted students with adequate SAT or ACT scores; others took the Accuplacer. If they fell below the Accuplacer cutoff score, they took the Learning and Study Strategies Inventory (LASSI, Weinstein et al., 1987), which was used with their HS GPA to determine whether to move them up to FYC. Thus, in this study, students could only be bumped up to FYC. Only students who took the Accuplacer were included in the study. Comparing the MM group and the BAU (Accuplacer only) group, 14.5% more MM students were placed in FYC (61.8% vs. 47.3%), 4.7% more enrolled, but only a non-significant 1.6% more passed (20.7 vs. 19.2%). Follow-up results are needed to see results after students have had time to complete developmental courses and attempt FYC. However, it is somewhat promising that of the students bumped up to FYC by the MM system, 54.5% enrolled in and 34.5% passed FYC (a pass rate of 63%). This provides further evidence for the need for better placement systems.

## *2.2 Accelerated Progress Through Developmental Writing*

One approach to reducing dropout from developmental programs is to accelerate progress through developmental courses by compressing courses into shorter time periods. For example, two levels of developmental writing might be compressed into a single course, or a developmental course might be compressed into a few weeks followed by FYC in the same semester. Another way to reduce the number of courses is to integrate separate reading and writing courses into a single course. A survey of a nationally representative sample of public 2-year and 4-year colleges (Rutschow et al., 2019) found that among community colleges, 67% offered traditional multi-semester sequences of reading and writing developmental courses, 64% offered integrated reading and writing courses, and 54% offered compressed courses.

Only a few studies have evaluated the effects of compression and integration. Hodara and Jaggars (2014) studied outcomes at five community colleges in the CUNY system for students who entered between 2001 and 2007. All colleges used a common cut-off score for placement into developmental writing. However, 3 colleges had two sequential levels of developmental writing while 2 others had only a single level prior to FYC. The study examined records over 3 years for all students whose scores would have placed them in the lower-level course, using propensity score matching to form equivalent groups. Students in colleges with a single level were 9.7% more likely to enroll in, and 6% more likely to pass FYC; they also earned 2 more college credits and were 2% (10% vs. 8%) more likely to earn an associate degree in 3 years. Analysis of syllabi indicated that courses differed in content. The 2-course sequence generally had grammar and paragraphs in the first

course and grammar review and paragraph/essay writing in the second. The single course focused more on writing in response to readings, which is better aligned with FYC. Thus, it is not clear whether the differences were due to structural differences in time or curricular differences in content.

Two studies have compared students who took separate developmental reading and writing courses with students who took a single integrated reading and writing course (Edgecombe et al., 2014; Paulson & Van Overschelde, 2019). Edgecombe et al. (2014) studied the program at Chabot College, a community college that had for a long time offered students the option of taking separate developmental reading and writing courses or an integrated course. They examined transcripts of all students entering from 1999 to 2010 who completed the separate or integrated option, using propensity score matching to compare outcomes over 5 years. Students who chose the integrated course were more likely to pass FYC (60 vs. 40%), earn more credits (36 vs. 30), and graduate (25% vs. 18% after 5 years); all comparisons were significant after controlling using propensity score matching. Based on qualitative analysis using interviews with faculty and administrators, student focus groups, and class observations, the authors concluded that some of the difference was due to the content of the integrated course, which was better aligned with FYC, which required students to integrate reading and writing to write essays with sources. It is also possible that the propensity score matching did not fully compensate for differences between the students who chose one option or the other.

In a contrasting setting, Paulson and Van Overschelde (2019) took the opportunity created by a new policy in Texas requiring all community colleges to offer integrated reading and writing (IRW) developmental courses. Using a large database available from the state, the researchers used data on students who began community college during a ten-year spanning the implementation of the new policy. Using passing grades on FYC or a reading-intensive credit course, the study compared students taking the IRW course with students who took separate reading and writing courses. They ran separate analyses comparing students who took only IRW (a) with students who took only developmental reading, (b) with students who took only developmental writing, and (c) with those taking both. They used propensity matching based on test scores, high school GPA, and demographics. For all three analyses, students who took the separate courses were more likely to pass the subsequent college-credit class. The authors interpret the negative results of the IRW reform as the result of decreased time on task, which is the essence of acceleration. The IRW course was half the credits and time required for both separate courses. However, the same could be said about the study of compressed developmental writing courses at CUNY (Hodara & Jaggars, 2014) and the IRW course at Chabot College (Edgecombe et al., 2014), both of which produced positive results. Another reason for the better results at CUNY and Chabot College may be that faculty in both of those settings had worked for some time to develop courses that would prepare students for FYC, whereas in Texas, faculty were mandated to combine courses without guidance or time for thoughtful curriculum development. Structural change demands thoughtful curricular change.

Another form of compressed course is offered in summer bridge programs (SBP), which have a long history (Sablan, 2014). SBPs focus on college readiness for students who may need extra support, such as low-income, minority, first-generation students, and students needing remedial work. The content typically includes developmental courses, college orientation, and study skills. Most published reports have described programs and outcomes at individual institutions (Sablan, 2014). Recently, a few studies have used rigorous quantitative methods to examine outcomes. Note that both of these studies combined results for students assigned to writing and math developmental courses.

Barnett et al. (2012) used a randomized experimental design to study SBPs at 8 community colleges in Texas. Students needing remedial education who planned on full-time study and who consented to participate in the study were randomly assigned to the SBR program or a control condition in which they were permitted to use other services offered by the colleges. Although the SBP had a short-term effect on completion of first-year college credit courses (writing or math depending on need), this difference disappeared after 2 years, and no differences were found in courses taken, or in persistence over 2 years.

More positive results of SBPs were found by Douglas and Attewell (2014) in a pair of complementary studies, both using propensity score matching to compare students who did and did not participate in SBPs. The first study used nationally representative transcript data from the US Department of Education. They found that students who had taken SBPs were 9.6% more likely to earn an associate or bachelor's degree within 6 years. The second study used data from six community colleges in one state; the sample was smaller but included placement test scores and included only remedial students. SBP students were more likely to remain enrolled after 2 years and earned more college credits.

Differences in the outcomes might be explained by motivation and life-circumstance factors not captured by the research designs. In the Barnett et al. study (2012), both treatment and control students had applied to participate in the SBP and intended to study full time. In contrast, in the Douglas and Attewell study (2014), all students placed in remedial education were included. Thus, the samples may have differed in motivation for college success and availability for summer study.

## *2.3 Corequisite Courses*

### Accelerated Learning Program

The original corequisite course program in writing was the Accelerated Learning Program (ALP), developed at Baltimore Community College by Peter Adams et al. (2009). In this program, students who place into the highest level of developmental writing have the option of taking the traditional developmental course or participating in ALP. In ALP, 8 developmental students participate in sections of FYC together with 12 regular students; at the same time, they take a support class taught by the

same instructor. Internal evaluation of the program (Adams et al., 2009) found positive effects on developmental students' completion of FYC and a second English course, so they invited the Community College Research Center (CCRC) to study the program more rigorously. CCRC researchers conducted two studies of ALP (Cho et al., 2012; Jenkins et al., 2010). The first CCRC study (Jenkins et al., 2010) used data on students enrolled in ALP or standard developmental writing during the first three semesters of ALP and their progress over 3 semesters, using regression analysis to control for student variables and full-time vs. part-time faculty. The second study (Cho et al., 2012) used an expanded data set covering 10 semesters of ALP again with follow-up for 3 semesters; it added propensity score matching to control for differences in students' demographics and prior achievement measures. The matching was important because only about 10% of students participated in ALP, and the ALP and non-ALP groups differed significantly on gender, race, SES, and full-time status (Cho et al., 2012). Despite differences, the results of the studies are similar, so data from the larger study are discussed here.

Comparing students who enrolled in ALP with those who chose traditional developmental writing over their first three semesters of college, Cho et al. (2012) found that ALP students were more likely to pass FYC (74.7% to 38.5%). This difference was due to the fact that only 53% of developmental students attempted FYC; pass rates and grades in FYC did not differ. ALP students were also more likely to enroll in (54% vs. 22%) and pass (37.5% vs. 16.8%) a second required English course. ALP students also were more likely to persist in college for a second semester and earned more college credits (3 more). The study also investigated the effects on the non-developmental students included in the ALP sections of FYC, comparing them to similar students in non-ALP sections of FYC. Those students did equally well in FYC but were 3% less likely to take and pass the second required English course.

### Corequisite Courses Mandated by State Policy

Several states have set policies encouraging or mandating development of corequisite courses in community colleges. These new policies offered opportunities for rigorous evaluation of the effects of corequisite courses. Tennessee (TN) was the first state to require colleges to offer corequisite courses (Ran & Lin, 2019); by 2015, 11 of 13 community colleges in TN were regularly offering corequisite courses. Using data from statewide high school testing and college transcripts for students entering college from 2010 to 2016, Ran and Lin (2019) used regression discontinuity methods to compare students placed into FYC with those just below the cutoff who were placed in traditional developmental writing or in corequisite courses. Most of the colleges (82%) used a 3-credit support course combined with FYC. Similar to ALP, the support course was specifically intended to help students pass the FYC course, but no information was provided about course content. Students placed in developmental writing compared to corequisite courses were 13% less likely to pass FYC in year 1 and still 9% less likely after 3 years. However,

no significant effects were found for total college credits earned, enrollment persistence, earning a credential, or transferring to a 4-year college. No significant differences were found between those placed in FYC or in corequisite courses for any outcomes. Thus, most of the effect seems to be due to placement in FYC rather than specific effects of the corequisite course; as for other regression discontinuity studies, the results apply to students who are close to the level required for FYC.

Not all college programs interpret corequisite options as a full 3-credit support course. In some state policies and community college practices, corequisite course programs have been defined broadly as programs that enroll underprepared students in FYC together with simultaneous supports in a variety of forms that are intended to help students pass FYC. For example, Texas mandated that all colleges offer corequisite programs but left decisions about the support component up to individual community colleges. A survey of 36 community colleges (CC) in Texas in the early years of the new policy found five models for corequisite courses (Daugherty et al., 2018). The most common model (27% of CCs) was to offer FYC and a developmental writing (DW) course in the same semester but without the same instructor or much coordination. Nearly as many CCs (23%) offered a combined FYC/DW course with extended time that enrolled just developmental students. Models similar to ALP with an extra support course taught by the same instructor, but with credits ranging from 1 to 3, were offered by 18% of CCs. Another model (14% of CCs) placed developmental and non- developmental students in FYC with mandatory weekly academic support for the developmental students (office hours or writing center). The fifth model (11%) offered separate sections of FYC for developmental students with technology support. Interviews with faculty and administrators found substantial concerns about limited buy-in among faculty, issues with scheduling and advising logistics, limited preparation for model design and instruction, and concerns about the rapid speed of change in state policy.

In this Texas context, a RCT (Miller et al., 2020) was conducted at five community colleges using three of the models just outlined. Two colleges used a modified ALP model with developmental and non-developmental students in FYC classes with a 1-credit support course for the developmental students, compared to the 3-credit support course in ALP. Two colleges provided even less support; they placed developmental and non-developmental students together in FYC classes with extra academic support for developmental students in the form of mandatory attendance at the instructor's office hours or at the writing center. Students who met the criteria for the highest level of development writing based on the Texas assessment were randomly assigned to (a) a treatment group assigned to the corequisite course option or (b) a control group assigned to an IRW developmental course. (The colleges had replaced developmental writing with IRW courses.) Results were assessed at the end of the first and second year of college. Developmental students in the treatment group were 21% more likely than controls to pass FYC by the end of year one and still 16% more likely after 2 years. Treatment students were also 6% more likely to pass English composition II after 2 years. To assess impact on reading, since the treatment students missed the reading instruction in the IRW course, success in a college-level reading intensive course (e.g., history) was assessed;

treatment students were 6% more likely to pass such a course. Treatment students also completed 1.5 more college credits by the end of the second year. Persistence in college was not affected by the treatment. Interestingly, no significant differences were found among the three approaches to corequisite support despite substantially different levels of required resources. There were, however, differences in outcomes between institutions, possibly reflecting differences in how supports were implemented. The gains may be due primarily to direct placement into FYC, as in the regression discontinuity studies.

## *2.4 Improved Pedagogical Methods*

Most of the research on reform efforts in developmental education has evaluated reforms in placement of students into developmental education courses or structural changes in the organization of courses, as discussed above. Research has found significant effects of such changes on successful completion of FYC and sometimes on outcomes such as persistence in college or credits earned. Even without those improvements, reducing the overall time that students spend in developmental education can save students time and money. However, limited research has investigated pedagogical methods for improving outcomes.

It is important to note that many of the structural reforms in courses discussed above have included significant work on course content and teaching methods. For example, research on compressed courses, integrated reading and writing courses, and corequisite courses seemed to produce better effects when faculty developed new courses at their own initiative than when reform was mandated by policy. For example, as noted above, the successful compressed developmental writing courses at CUNY (Hodara & Jaggars, 2014) and the IRW course at Chabot College (Edgecombe et al., 2014) were designed and refined over time by faculty with the specific intention to integrate the contents with the requirements of FYC. In contrast, the less successful effort to implement IRW courses in Texas (Paulson and Van Overschelde (2019) was initiated by policy mandates from the state. ALP is an example of a successful corequisite program developed by faculty (Adams et al., 2009), and Peter Adams has a textbook available for corequisite courses (Adams, 2019). Qualitative research by Edgecombe et al. (2013) found that the process of implementing significant innovations in course structures requires faculty involvement over time in curriculum design to match course content and instructional methods to the expectations of later courses.

Basic writing and college composition are active fields of scholarship with several respected journals, and books are available for instructors on effective teaching strategies for developmental reading and writing (e.g., Flippo & Bean, 2018). However, few studies have systematically evaluated the effects of instructional methods on student outcomes. The What Works Clearinghouse (WWC) produced a practice guide on developmental education (Bailey et al., 2016) which includes six recommendations, mostly focused on the reforms in placement and structural

changes in courses; only one recommendation is about instruction, and it is supported by only one study in math.

A review of research on instruction for students in developmental reading and writing (Perin & Holschuh, 2019) included over 30 studies, but nearly all were descriptive reports of teaching practices or pretest-posttest-only studies; only two experimental studies with control groups were found. Perin et al. (2013) conducted an experimental study of instruction in reading and writing strategies using supplementary instructional manuals that students used independently in the content of credit-bearing content area courses. The instructional materials guided students in a series of 12 steps for reading a text and writing a summary and an opinion article on a controversy related to the topic. Compared to a test-only control group, treatment students wrote better summaries on a posttest. The other experimental study included in the review was our work on strategy instruction, the Supporting Strategic Writing project (MacArthur et al., 2015; Traga Philippakos & MacArthur, 2020) which we summarize in the next section.

## Strategy Instruction in Writing

The central idea of strategy instruction is that it is possible to teach students to use strategies based on the cognitive and metacognitive strategies used by proficient readers or writers (MacArthur, 2011; MacArthur & Graham, 2016). A large body of research shows strong effects of writing strategy instruction on quality of writing with elementary and secondary students, including students with disabilities, especially when combined with self-regulation strategies; a review that integrated multiple meta-analyses of writing instruction (Graham et al., 2016) reported an average effect size of strategy instruction on writing quality of 1.26 across 42 studies. One of the most effective models of strategy instruction is Self-Regulated Strategy Development (SRSD, Harris & Graham, 2009), which integrates domain-specific strategies (e.g., planning and revising for writing) with self-regulation strategies, such as goal setting, progress monitoring, and self-reinforcement. The effect size for SRSD studies of writing instruction was 1.59 (Graham et al., 2016).

A few studies have investigated strategy instruction with adult learners and college students with disabilities. Two studies have investigated SRSD with Adult Secondary Education learners (MacArthur & Lembo, 2009; Berry & Mason, 2012); both studies used single-case designs to evaluate instruction in strategies for writing argumentative essays, and both found positive effects on writing quality. In a study with college students with developmental disabilities (Woods-Groves et al., 2014), students learned a strategy for taking essay tests; the experimental study found a significant effect on writing quality (ES = .95). Nicholas and Menchetti (2005) conducted an experimental study with African American college students with learning disabilities. Both treatment and control students participated in a workshop on test-taking strategies, and the treatment group received a second workshop on a strategy for planning and drafting; treatment students included more details in their

essays, but there was no difference in overall quality. The lack of effects compared to other research on strategy instruction may be due to the very short instructional time.

### Supporting Strategic Writers

Supporting Strategic Writers (SSW, MacArthur et al., 2015; Traga Philippakos & MacArthur, 2020) is an instructional approach and curriculum for developmental writing courses based on strategy instruction with self-regulation. The goals of the SSW approach are widely shared – that students will develop knowledge of academic writing genres; strategies for critical reading, planning and revising; and the motivational beliefs that support continued critical reading and writing in the future (Rose, 1989; CWPA et al., 2011). The instructional approach is based on strategy instruction with self-regulation (Harris & Graham, 2009; MacArthur, 2011) integrated with practices common in college composition and with an emphasis on learning self-evaluation. Students learn genre-based strategies based on the rhetorical purposes and text structures of genres often taught in college, especially argumentation. Genre features integrate the strategies for planning and revising, as well as critical reading and note-taking. The strategies provide an initial map for students unsure about how to engage in the writing process. Equally important, students learn metacognitive strategies for goal-setting, task management, progress monitoring, and reflection. Journaling and class discussions engage students in reflecting on how they can take control of their own learning through setting goals, selecting strategies, and monitoring progress. Self-evaluation and reflection on one's progress are critical to developing a growth mindset (Yeager & Dweck, 2012) that learning is possible with effort and strategic choices. Pedagogical methods include discussion of good and weak essays, think-aloud modeling of strategies, collaborative writing, peer review and self-evaluation, reflective journaling, and gradual release of responsibility as students master the strategies. The writing and reading strategies, self-regulation strategies, and pedagogical methods will be explained further after a summary of the research.

Research on SSW began with three cycles of design research in collaboration with community college faculty teaching two levels of developmental writing (MacArthur & Philippakos, 2012, 2013). The resulting curriculum included units of instruction on writing essays without sources in several genres (personal narrative, procedural, causal explanation, comparison, and argument), as well as procedures for professional development (workshops and coaching). Subsequently, a quasi-experimental study (MacArthur et al., 2015) was conducted involving two colleges, 13 instructors, and 276 students (48% minority, 10% non-native English speakers); classes at two levels of developmental writing were included. The SSW curriculum was compared to control classes that received typical instruction for a full semester. SSW had a large effect on quality of argumentative writing (ES = 1.22). It also had a large effect on self-efficacy (confidence) and a moderate effect on mastery motivation.

Next, a rigorous experimental study (MacArthur et al., 2019) was conducted at two community colleges with 19 instructors of developmental writing courses randomly assigned to treatment and control and 207 students (62% female; 57% minority, 12% non-native English speakers). The SSW approach had a large effect on quality of argumentative writing on a final exam (ES = 1.18). It also had positive effects on a standardized writing assessment (NAEP) (ES = 0.67) and on motivation outcomes – self-efficacy for writing and affect.

Having accumulated solid evidence of the effects of SSW in courses focused on teaching writing without sources, we returned to a semester of design research to extend the approach to include writing using sources, which required integration of reading and writing strategies. For success in college, students need to read source materials critically and synthesize information across sources to write their own critical essays. The change was consistent with movement in developmental education toward integrated reading and writing and corequisite courses focused on tasks typical of FYC. Strategies were added to support critical reading and note-taking, writing summary-response papers, and integrating source information into essays. Like the writing strategies, these new strategies were based on genre analysis and supported students in considering author/source, audience, and purpose and in using genre elements in taking notes, summarizing, and responding to source articles. Another experimental study (MacArthur et al., 2021) was conducted in developmental writing courses at two community colleges with 23 instructors randomly assigned to treatment and control and 243 students. The primary outcome measure was an argumentative essay using two source articles. A moderate positive effect was found on quality of those essays (ES = .58, $p < .01$). However, no significant effects were found on the standardized essay or on the motivational variables.

Most recently, a quasi-experimental study (Nefferdorf, 2020) evaluated an adapted version of the course in a compressed developmental writing course that met 4 days a week for 4 weeks at the start of the semester, leaving time for an 11-week FYC class. The curriculum included writing of summary-response papers and essays with sources. Five instructors (2 T, 3 C) and 65 students participated. The SSW approach had a large effect on the quality of argumentative essays with sources (ES = 0.97). No significant effects were found on grades in the subsequent FYC course.

**Components of SSW** The research results do not tell us which of the components of SSW were responsible for the positive effects on writing performance and motivation. However, it is important to consider which components are critical and even how components might be implemented and adapted in a variety of settings, for example, in IRW or corequisite courses, or even in FYC. Teaching writing strategies is hardly a new idea in college writing; rhetorical analysis, brainstorming, idea mapping, outlining, and using rubrics for evaluation are often mentioned in college writing texts. Thus, it is important to understand what components or instructional features make the approach work. We think about this in terms of three

questions: What strategies to teach? What pedagogical strategies are important? How to develop independent self-regulation?

The writing and reading strategies in SSW integrate rhetorical knowledge of genres (Rose & Martin, 2012) with cognitive processes used by proficient writers (Hayes, 1996). In particular, the strategies draw on research on text-structure based strategies for writing (Englert et al., 1991) and reading (Bogaerds-Hazenberg et al., 2020). The purposes and organizational structures of genres integrate the strategies for planning, revising, and critical reading. For planning, students analyze author and purpose, choose an appropriate genre, and use knowledge of the elements of that genre to guide brainstorming and organizing ideas graphically during planning. Evaluation and revision is guided by rubrics based on the same genre elements. The strategies for critical reading likewise involve rhetorical analysis and identification of main ideas based on the genre elements. Notes from reading are used to write summary-response papers, which prepare them to integrate sources in their own essays.

The pedagogical methods are based on extensive research in strategy instruction (Graham et al., 2016; Harris & Graham, 2009). Students who have limited experience with academic writing need explicit instruction to develop knowledge about genres and effective writing processes. Instruction in a new genre begins with discussion of the purpose of the genre and analysis of good and weak examples. Then the instructor models use of the strategy, thinking aloud to make the cognitive and metacognitive processes visible. Students and teachers engage in collaborative practice followed by guided practice. Students are prepared for peer review via more modeling and collaborative practice of applying a genre-specific rubric to evaluate and make revision suggestions. The instructor also provides feedback and an editing lesson before students revise their papers. To support independence, students write a second paper with less support.

The importance of metacognitive, self-regulation strategies has been demonstrated by research on SRSD (Harris & Graham, 2009). Following that model, students learn self-regulation strategies for goal setting, task management, progress monitoring, and reflection. Drawing on terminology from community colleges (Bickerstaff et al., 2017), these strategies are called Strategies for Academic Success. As in SRSD, instructors discuss the importance of taking control of the writing process and include self-regulation comments in their think-aloud modeling. However, the primary means for developing self-regulation is regular reflection in journals about goals, strategies, and progress together with class discussion based on the journal reflections. The class discussions provide group support as students learn that they can take control of the writing process and begin to see themselves as writers (Oyserman, 2007).

Finally, a key aspect of self-regulation is the development of skill in self-evaluation of writing. Two meta-analytic reviews (Graham et al., 2015; Hillocks, 1986) found positive effects on writing quality from instruction in evaluation criteria. Furthermore, research has demonstrated the value of giving feedback during peer review (Cho & MacArthur, 2011; Philippakos & MacArthur, 2016), which

provides practice in evaluation. In SSW, evaluation criteria and revision processes are introduced at the start of each genre-based unit of instruction through discussion and evaluation of good and weak essays. Also, students are prepared for peer review by collaborative practice evaluating essays by unknown peers and making specific suggestions for improvement.

## 3 Implications for Research and Practice

Research on developmental writing and alternatives for underprepared students has made substantial progress over the past decade. Some research has shown benefits of developmental writing for students placed in lower-level courses (Boatman & Long, 2018), and research shows that students who persist in completing developmental courses do well (Bailey et al., 2010; Chen, 2016). However, research has also shown consistently that many students who would have been placed in development writing or reading courses based on placement tests can be successful in FYC and make better progress in subsequent classes (Valentine et al., 2017; Woods et al., 2019). Thus, developmental courses can serve as a barrier to college success for some students, particularly those right below the cutoff for placement in FYC. Placement in developmental education delays college work and may be discouraging. Following logically from these findings, many states and colleges have implemented reforms involving the use multiple measures for placement, including high school GPA or courses taken in addition to a placement test. These reforms to placement have shown some success in improving outcomes (Bailey et al., 2016; Barnett et al., 2020; Cullinan et al., 2019).

Other reform efforts have involved restructuring courses. One alternative that has shown success is corequisite courses that allow students to take FYC without delay but still receive extra instructional support (Adams et al., 2009; Cho et al., 2012; Daugherty et al., 2018; Jenkins et al., 2010; Miller et al., 2020; Ran & Lin, 2019). These programs avoid the potential problems of developmental courses while still recognizing the needs of some students who are underprepared for college writing. Short of placing students directly in FYC, efforts to accelerate progress by combining reading and writing courses or reducing the levels of developmental courses have also shown promise (Edgecombe et al., 2014; Hodara & Jaggars, 2014).

However, the research on restructuring courses has also provided indications that such restructured courses are more likely to be successful when the reforms are initiated by faculty and new courses are carefully designed to prepare students for FYC. These findings suggest the need for more research on pedagogical methods. In K-12 education, substantial bodies of research provide evidence-based instructional practices for writing and reading instruction (Graham et al., 2016). However, the same cannot be said about instruction for underprepared college students (Perin & Holschuh, 2019). As shown by the example of the SSW project, improvements in instruction can have large effects on students' writing performance. Further research on instruction in postsecondary settings is much needed.

One promising area for such research is investigation of strategy instruction. A large body of research with elementary and secondary students (Graham et al., 2016) has shown substantial positive effects on writing quality and reading comprehension, both for students with disabilities and for typical students. One possible direction for that research is to adapt components of the SSW instruction to varied settings. For example, the revising strategies using genre-specific rubrics and careful preparation for peer review could be studied in a variety of classes from developmental writing to corequisite classes to FYC. Research shows positive effects of peer review, but only when students receive some instruction in how to evaluate writing and give helpful feedback (Graham & Perin, 2007; MacArthur, 2016). Another possibility is to add instruction in metacognitive, self-regulation strategies to writing courses using journal writing and class discussion. Many community colleges offer separate courses in study strategies, but those strategies might be learned more effectively in challenging writing courses where the strategies are tied directly to specific academic tasks; students might be more likely to see the benefits of such strategies on their work and grades. Another line of research could explore which students benefit from strategy instruction and whether students at different achievement levels need different amounts of systematic instruction. For example, in the context of a corequisite course, the typical students might only need explanations or brief demonstrations of strategies, whereas underprepared students might need more extensive think-aloud modeling, collaborative practice, and guidance in applying them. One final suggestion is for more research on strategies for critical reading of sources and integration of ideas from sources into students own writing. Research on such 'synthesis writing' (van Ockenburg et al., 2019) has raised interesting questions about the cognitive processes required for success in such tasks. The approach in the SSW curriculum could doubtless be improved. Writing using sources is substantially harder than writing without that requirement.

It is time for research on structural changes in developmental courses to be integrated with research on instructional methods. Although there is a wealth of professional knowledge in the field about effective instruction, systematic outcomes-based research provides a valuable service in understanding which methods work well and which students benefit from them.

## References

ACT, Inc. (2006). *COMPASS/ESL reference manual*. Author.

Adams, P. (2019). *The Hub: A place for reading and writing*. Bedford/St. Martin's.

Adams, P., Gearhart, S., Miller, R., & Roberts, A. (2009). The accelerated learning program: Throwing open the gates. *Journal of Basic Writing, 28*(2), 50–69.

Bahr, P. R. (2010). Revisiting the efficacy of postsecondary remediation: The moderating effects of depth/breadth of deficiency. *The Review of Higher Education, 33*(2), 177–205. https://doi.org/10.1353/rhe.0.0128

Bailey, T., Jeong, D. W., & Cho, S.-W. (2010). Referral, enrollment, and completion in developmental education sequences in community colleges. *Economics of Education Review, 29*, 255–270. https://doi.org/10.1016/j.econedurev.2009.09.002

Bailey, T., Bashford, J., Boatman, A., Squires, J., Weiss, M., Doyle, W., Valentine, J. C., LaSota, R., Polanin, J. R., Spinney, E., Wilson, W., Yeide, M., & Young, S. H. (2016). *Strategies for postsecondary students in developmental education: A practice guide for college and university administrators, advisors, and faculty*. Institute of Education Sciences, What Works Clearinghouse. https://eric.ed.gov/?id=ED570881

Barnett, E. A., Bork, R. H., Mayer, A. K., Pretlow, J., Wathington, H. D., & Weiss, M. J. (2012). *Bridging the gap: An impact study of eight developmental summer bridge programs in Texas*. National Center for Postsecondary Research. https://eric.ed.gov/?id=ED533824

Barnett, E. A., Kopko, E., Cullinan, D., & Belfield, C. R. (2020, October). *Who should take college-level courses? Impact findings from an evaluation of a multiple measures assessment strategy*. Center for the Analysis of Postsecondary Readiness. https://academiccommons.columbia.edu/doi/10.7916/d8-262b-wq33

Belfield, C. R., & Crosta, P. M. (2012). *Predicting success in college: The importance of placement tests and high school transcripts* (CCRC Working Paper No. 42). Community College Research Center. https://eric.ed.gov/?id=ED529827

Berry, A., & Mason, L. H. (2012). The effects of self-regulated strategy development on the writing of expository essays for adults with written expression difficulties: Preparing for the GED. *Remedial and Special Education, 33*, 124–136. https://doi.org/10.1177/0741932510375469

Bickerstaff, S., Barragan, M., & Rucks-Ahidiana, Z. (2017). Experiences of earned success: Community college students' shifts in college confidence. *International Journal of Teaching and Learning in Higher Education, 29*, 501–510. https://eric.ed.gov/?id=EJ1151074

Boatman, A., & Long, B. T. (2018). Does remediation work for all students? How the effects of postsecondary remedial and developmental courses vary by level of academic preparation. *Educational Evaluation and Policy Analysis, 40*(1), 29–58. https://doi.org/10.3102/0162373717715708

Bogaerds-Hazenberg, S. T. M., Evers-Vermeul, J., & van den Bergh, H. (2020). A meta-analysis on the effects of text structure instruction on reading comprehension in the upper elementary grades. *Reading Research Quarterly, online first April 2020*. https://doi.org/10.1002/rrq.311

Chen, X. (2016, September). *Remedial coursetaking at U.S. public 2- and 4-year institutions: Scope, experiences, and outcomes (NCES 2001—405). U.S. Department of Education.* National Center for Education Statistics. https://eric.ed.gov/?id=ED568682

Chen, X., Elliott, B. G., Kinney, S. K., Cooney, D., Pretlow, J., Bryan, M., Wu, J., Ramirez, N. A., & Campbell, T. (2019). Persistence, retention, and attainment of 2011–12 first-time beginning postsecondary students as of spring 2017. U.S. Department of Education. National Center for Education Statistics. https://eric.ed.gov/?id=ED593528

Cho, K., & MacArthur, C. (2011). Learning by reviewing. *Journal of Educational Psychology, 103*, 73–84. https://doi-org.udel.idm.oclc.org/10.1037/a0021950

Cho, S., Kopko, E., Jenkins, D., & Jaggars, S. S. (2012, December). *New evidence of success for community college remedial English students: Tracking the outcomes of students in the Accelerated Learning Program (ALP)*. Community College Research Center. https://eric.ed.gov/?id=ED538995

College Board. (2016). *Accuplacer program manual*. Author.

Council of Writing Program Administrators, National Council of Teachers of English, & National Writing Project. (2011). *Framework for success in postsecondary writing*. Authors. http://wpacouncil.org/files/framework-for-success-postsecondary-writing.pdf

Cullinan, D., Barnett, E. A., Kopko, E., Lopez, A., & Morton, T. (2019, December). *Expanding access to college-level courses: Early findings from an experimental study of multiple measures assessment and placement*. MDRC & Community College Research Center. https://eric.ed.gov/?id=ED602455

Daugherty, L., Gomez, C. J., Carew, D. G., Mendoza-Graf, A., & Miller, T. (2018). *Designing and implementing corequisite models of developmental education: Findings from Texas community colleges*. Rand Corp. https://www.rand.org/pubs/research_reports/RR2337.html

Douglas, D., & Attewell, P. (2014). The bridge and the troll underneath: Summer bridge programs and degree completion. *American Journal of Education, 121*(1), 87–109. https://doi.org/10.1086/677959

Edgecombe, N., Cormier, M. S., Bickerstaff, S., & Barragan, M. (2013). *Strengthening developmental education reforms: Evidence on implementation efforts from the scaling innovation project* (CCRC Working Paper No. 61). Community College Research Center. https://academiccommons.columbia.edu/doi/10.7916/D8QV3JJ3

Edgecombe, N., Jaggars, S. S., Xu, D., & Barragan, M. (2014). *Accelerating the integrated instruction of developmental reading and writing at Chabot College* (CCRC working paper no. 71). Community College Research Center. https://academiccommons.columbia.edu/doi/10.7916/D8CZ359B

Englert, C. S., Raphael, T. E., Anderson, L. M., Anthony, H. M., & Stevens, D. D. (1991). Making writing strategies and self-talk visible: Cognitive strategy instruction in writing in regular and special education classrooms. *American Educational Research Journal, 28*, 337–372. https://doi.org/10.3102/00028312028002337

Flippo, R. F., & Bean, T. (Eds.). (2018). *Handbook of college reading and study strategies* (3rd ed.). Routledge.

Goldrick-Rab, S. (2010). Challenges and opportunities for improving community college student success. *Review of Educational Research, 80*, 437–469. https://doi.org/10.3102/0034654310370163

Goudas, A. M., & Boylan, H. R. (2012). Addressing flawed research in developmental education. *Journal of Developmental Education, 12*(1), 2–13. https://eric.ed.gov/?id=EJ1035669

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal, 115*, 523–547. http://www.jstor.org/stable/10.1086/681947

Graham, S., Harris, K., & Chambers, A. B. (2016). Evidence-based practice and writing instruction: A review of reviews. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 211–226). Guilford Press.

Harris, K. R., & Graham, S. (2009). Self-regulated strategy development in writing: Premises, evolution, and the future. *British Journal of Educational Psychology Monograph Series II, 6*, 113–135. https://doi.org/10.1348/978185409x422542

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The Science of Writing* (pp. 1–27). Erlbaum.

Hillocks, G. (1986). *Research on written composition: New directions for teaching*. ERIC Clearinghouse on Reading and Communication Skills.

Hodara, M., & Jaggars, S. S. (2014). An examination of the impact of accelerating community college students' progression through developmental education. *The Journal of Higher Education, 85*(2), 246–276. https://doi.org/10.1080/00221546.2014.11777326

Hodges, R., McConnell, M. C., Lollar, J., Guckert, D. A., Owens, S., Gonzales, C., Hoff, M. A., Lussier, K. O., Wu, N., & Shinn, H. B. (2020). Developmental education policy and reforms: A 50-state snapshot. *Journal of Developmental Education, 44*(1), 2–17.

Jenkins, D., Speroni, C., Belfield, C., Jaggars, S., & Edgecombe, N. (2010). *A model for accelerating academic success of community college remedial English students: Is the Accelerated Learning Program (ALP) effective and affordable?* (CCRC working paper no. 21). Community College Research Center.

MacArthur, C. A. (2011). Strategies instruction. In K. R. Harris, S. Graham, & T. Urdan (Eds.), *Educational psychology handbook: Vol. 3. Application to learning and teaching* (pp. 379–401). American Psychological Association.

MacArthur, C. A. (2016). Instruction in evaluation and revision. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 272–287). Guilford.

MacArthur, C. A., & Graham, S. (2016). Writing research from a cognitive perspective. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 24–40). Guilford.

MacArthur, C. A., & Lembo, L. (2009). Strategy instruction in writing for adult literacy learners. *Reading and Writing, 22*, 1021–1032. https://doi.org/10.1007/s11145-008-9142-x

MacArthur, C. A., & Philippakos, Z. A. (2012). Strategy instruction with college basic writers: A design study. In C. Gelati, B. Arfé, & L. Mason (Eds.), *Issues in writing research* (pp. 87–106). CLEUP.

MacArthur, C. A., & Philippakos, Z. A. (2013). Self-regulated strategy instruction in developmental writing: A design research project. *Community College Review, 41*, 176–195. https://doi.org/10.1177/0091552113484580

MacArthur, C. A., Philippakos, Z. A., & Ianetta, M. (2015). Self-regulated strategy instruction in college developmental writing. *Journal of Educational Psychology, 107*, 855–867. https://doi.org/10.1037/edu0000011

MacArthur, C. A., Traga Philippakos, Z. A., May, H., & Compello, J. (2019, March). *Self-regulated strategy instruction for basic college writers: Results from a randomized experiment.* Paper presented at the annual conference of the Society for Research on Educational Effectiveness, Washington, DC.

MacArthur, C. A., Traga Philippakos, Z. A., & May, H. (2021, March). *Basic writers and the challenges of writing from sources: Experimental study of a strategy instruction approach.* Paper presented at the tri-annual conference of Writing Research Across Borders, Xi'an, China. (postponed from March, 2020).

Miller, T., Daugherty, L., Martorell, P., & Gerber, R. (2020). *Assessing the effect of corequisite English instruction using a randomized controlled trial*. American Institutes for Research. https://collegecompletionnetwork.org/sites/default/files/2020-05/ExpermntlEvidncCoreqRemed-508.pdf

Nastal-Dema, J. (2019). Beyond tradition: Fairness, placement, and success at a two-year college. *Journal of Writing Assessment, 12*(1), 1–18.

Nefferdorf, E. (2020). *Design, implementation and outcomes of a condensed curriculum for an accelerated developmental education English course*. [Unpublished doctoral dissertation]. University of Delaware.

Nicholas, K. R., & Menchetti, B. M. (2005). An exploratory investigation of structured writing strategy training for African-American college students with learning disabilities. *Journal of College Reading and Learning, 36*(1), 37–59. https://doi.org/10.1080/10790195.2005.10850179

Oyserman, D. (2007). Social identity and self-regulation. In A. W. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 432–453). Guilford Press.

Paulson, E. J., & Van Overschelde, J. P. (2019). Accelerated integrated reading and writing: A statewide natural experiment. *Community College Journal of Research and Practice, 45*(*1*), *13–30*. https://doi.org/10.1080/10668926.2019.1636733

Perin, D. (2006). Can community colleges protect both access and standards? The problem of remediation. *Teachers College Record, 108*(3), 339–373. https://doi.org/10.1111/j.1467-9620.2006.00654.x

Perin, D., & Holschuh, J. P. (2019). Teaching academically underprepared postsecondary students. *Review of Research in Education, 43*, 363–393. https://doi.org/10.3102/0091732x18821114

Perin, D., Bork, R. H., Peverly, S. T., & Mason, L. H. (2013). A contextualized curricular supplement for developmental reading and writing. *Journal of College Reading and Learning, 43*, 8–38. https://doi.org/10.1080/10790195.2013.10850365

Philippakos, Z. A., & MacArthur, C. A. (2016). The effects of giving feedback on the persuasive writing of fourth and fifth-grade students. *Reading Research Quarterly, 51*, 419–433. https://doi.org/10.1002/rrq.149

Ran, F. X., & Lin, Y. (2019, November). *The effects of corequisite remediation: Evidence from a statewide reform in Tennessee* (CCRC Working Paper No. 115). Center for Analysis of Postsecondary Readiness. https://eric.ed.gov/?id=ED600570

Rose, M. (1989). *Lives on the boundary: The struggles and achievements of America's underprepared*. Free Press.

Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: Genre, knowledge, and pedagogy in the Sydney School*. Equinox.

Rutschow, E. Z., Cormier, M. S., Dukes, D., & Cruz Zamora, D. (2019, November). *The changing landscape of developmental education practices: Findings from a national survey and interviews with postsecondary institutions*. Center for Analysis of Postsecondary Readiness. https://eric.ed.gov/?id=ED600433

Sablan, J. R. (2014). The challenge of summer bridge programs. *American Behavioral Scientist, 58*(8), 1035–1050. https://doi.org/10.1177/0002764213515234

Scott-Clayton, J. (2012). *Do high-stakes placement exams predict college success?* (CCRC working paper no. 41). Community College Research Center. https://eric.ed.gov/?id=ED529866

Snyder, T. D., de Brey, C., & Dillow, S. A. (2019). *Digest of education statistics 2018*. National Center for Education Statistics. https://nces.ed.gov/pubs2017/2017094.pdf

Traga Philippakos, Z. A., & MacArthur, C. (2020). Writing strategy instruction for low-skilled postsecondary students. In D. Perin (Ed.), *The Wiley handbook of adult literacy* (pp. 495–516). Wiley.

Valentine, J. C., Konstantopoulos, S., & Goldrick-Rab, S. (2017). What happens to students placed into developmental education? A meta-analysis of regression discontinuity studies. *Review of Educational Research, 87*(4), 806–833. https://doi.org/10.3102/0034654317709237

van Ockenburg, L., van Weijen, D., & Rijlaarsdam, G. (2019). Learning to write synthesis texts: A review of intervention studies. *Journal of Writing Research, 10*, 401–428. https://doi.org/10.17239/jowr-2019.10.03.01

Weinstein, C. E., Palmer, D., & Schulte, A. C. (1987). *Learning and Study Strategies Inventory (LASSI)*. H & H Publishing.

Whinnery, E., & Pompelia, S. (2018). *50-state comparison: Developmental education policies*. Education Commission of the States.

Woods, C. S., Park, T., Hu, S., & Bertrand Jones, T. (2019). Reading, writing, and English course pathways when developmental education is optional: Course enrollment and success for underprepared first-time-in-college students. *Community College Journal of Research and Practice, 43*(1), 5–25. https://doi.org/10.1080/10668926.2017.1391144

Woods-Groves, S., Hua, Y., Therrien, W. J., Kaldenberg, E. R., Lucas, K. G., Hendrickson, J. M., & McAninch, M. J. (2014). An investigation of strategic writing instruction for post-secondary students with developmental disabilities. *Education and Training in Autism and Developmental Disabilities, 49*(2), 248–262. https://www.jstor.org/stable/23880608

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*, 302–314. https://doi.org/10.1080/00461520.2012.722805

# Influence of Disclosure Topic and Linguistic Perspective on Expressive Writing

**Teresa Jacques, Ana P. Azzam, Francisca Costa, and Rui A. Alves**

**Abstract** Expressive writing is a form of writing in which a person discloses highly charged emotional episodes, such as the loss of a loved one or a life-threatening episode. Typically, such events are concealed and seem to exert a toll on health. Fortunately, written disclosure is frequently associated with increased well-being (Pennebaker & Smyth, 2016). In this chapter, we report on a writing study conducted with undergraduates in which we manipulated a disclosure topic (expressive writing or daily routine) and pronoun perspective (first person/self-immersed; third person/self-distanced). The linguistic perspective that pronouns convey on discourse is likely not innocuous to a consideration of content and emotions shown in expressive writing. Studies have shown that using third person seems to facilitate self-distancing from the actual emotional experience (Kross and Ayduk, Advances in experimental social psychology. Elsevier Academic Press, 2017). Perhaps leading to increased expressiveness and increased self-regulation. One hundred and ten texts collected in 15 min writing sessions were analysed by automated linguistic analyses (using HandSpy 3.0) and evaluated by independent judges, to study the linguistic features and emotional content of the texts. We found that the trauma groups wrote using a higher number of different function words and higher lexical density and the self-distanced groups showed higher idea density, in comparison to the other groups. In addition, the self-distanced trauma group wrote using a higher number of positive words, in comparison to the self-immersed group. This is a push forward in the field of expressive writing as it might encourage others to start using non-traditional expressive writ-

This research was funded by a research grant from BIAL Foundation (312/16) awarded to the last author. The first author was supported by a doctoral grant from the Portuguese Foundation for Science and Technology (grant 2020.05024.BD).

T. Jacques · A. P. Azzam · F. Costa
University of Porto, Porto, Portugal

R. A. Alves (✉)
Faculdade de Psicologia e de Ciências de Educação, University of Porto, Porto, Portugal
e-mail: ralves@fpce.up.pt



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_19

ing prompts and to analyse linguistic and emotional content used during writing. These findings are framed and discussed at the light of the well-known Graham's Writer(s)-Within-Community model (Educ Psychol 53:258–279, 2018a).

**Keywords** Expressive writing · Emotion regulation · Self-distancing · Linguistic perspective

As noted by Graham (1982), composition is a powerful tool for exploring thought and fulfil emotional needs. In 1986, Pennebaker and Beall created a writing task that uses writing to face and manage emotions – the expressive writing paradigm. This paradigm involves writing about a personally traumatic experience, encouraging the writer to immerse in the feelings and emotions associated with the event. Particiants are usually asked to write about the trauma for one or more consecutive days, for 15–30 min each day. This paradigm has shown that when people transform their feelings and thoughts about personally upsetting experiences into language, their physical and mental health often improve (Frattaroli, 2006; Smyth, 1998; Valtonen, 2020). The positive link between expressive writing and well-being has been attributed to several sources: writing about difficult events can facilitate the purging of unwanted thoughts, increase the stress relieving effects of self-disclosure and help make sense of emotionally troubling events, leading to a more efficient management of personal feelings and emotions (D'Mello & Mills, 2014; Smyth & Pennebaker, 2008). Even when participants cry or report being upset by expressive writing, they generally find that it was a valuable and meaningful experience (Pennebaker & Chung, 2007).

When people remember negative autobiographical experiences, they tend to adopt a self-immersed perspective in which they visualise their experience happening all over again through their own eyes (Grossmann & Kross, 2010). Linguistically this is often reflected in the use of first-person pronouns, which was the case in the original expressive writing experiment (Pennebaker & Beall, 1986). However, Kross and collaborators (2005) noticed that when writing about negative events in the third person perspective, participants reported lower levels of distress in comparison to first person. They coined this as a "self-distanced perspective". Self-distanced writing is thought to be beneficial because it helps reflect about the event from an outsider's point of view, reducing negative affect and rumination (Kross et al., 2005), and increase the ability to recognize and reappraise emotions (Kross & Ayduk, 2017). Linguistically, perspective also appears to influence affect, with participants who complete third person expressive writing reporting less negative affect in comparison to those who complete first person expressive writing (Yasinski et al., 2016), particularly when self-distancing is spontaneous (Ayduk & Kross, 2010). Overall, self-distancing appears to alter the representation of the negative experience (Park et al., 2016).

The recent Graham's Writer(s)-Within-community model (WWC; 2018a) integrates and expands various writing conceptualizations, situating writing within communities. This model specifies the cognitive resources that community members bring to writing, such as long-term memory, working memory and executive control. The memories associated with a traumatic event are stored in long-term memory, while also housing all of the writing knowledge acquired as a result of the collective experiences in writing communities, including knowledge about emotional reactions to writing under different conditions (Graham, 2018a, b).

The ability to write about traumatic events is also dependent on working memory. Working memory refers to limited and temporary storage that can hold information necessary for action (Diamond, 2006). Likely, emotional content needs to be stored in long-term memory and be activated within working memory where it can be processed (Graham, 2018a). Working memory enables the regulation of attention, any writing production processes, and emotions, particularly it should be engaged in emotion regulation during writing (Graham, 2018a), key components of expressive writing. This connection between working memory and long-term memory also seems to be underlying the cognitive reserve mechanism (Sandry et al., 2015; Sandry & Sumowski, 2014), a beneficial mechanism for those who have experienced traumatic events (Rakesh et al., 2019).

Some important control processes that are involved in remembering and describing a traumatic event are related to executive control – goal setting, planning, monitoring, and reviewing (Zimmerman & Risemberg, 1997). These four processes can be applied to all aspects of the writing process (Zimmerman & Risemberg, 1997), and happen in conjunction with the purpose set by the communities where writing happens (Graham, 2018a). For example, when participants are assigned to write about their most upsetting life event, they will use these processes to shape their narrative, personalizing it in the process, with their own vision of the experience. Formulating goals can occur at any point during writing, before starting to write (e.g. control emotions) and while writing (e.g. including examples and understandable facts about the traumatic event). Use of long-term memory resources during production processes are always moderated by emotions (Graham, 2018a). Positive or negative emotions can increase writing effort and can merge with information stored in long-term memory, to moderate the relationship between emotions and writing (Graham, 2018a). Ultimately, emotions influence what a writer decides their final text will be (Pekrun et al., 2007).

An important method to study expressive writing is through automated linguistic analysis. These tools can help study expressive writing in two ways: facilitate routine word analysis and help identify recurring patterns occurring in the text (Pirnay-Dummer, 2016). When conducting linguistic analysis on expressive writing texts, Pennebaker (2011) found that the more participants naturally changed their use of first person singular pronouns (e.g., I, me) compared with other pronouns (e.g., we, they), the better their health became after writing. This is an example of a situation where language style affected the benefits of expressive writing.

Language style – how people speak or write – can therefore reveal aspects of someone's life (Pennebaker, 2011). This can be studied by analysing two types of words: Content words and function words. Content words are words charged with meaning, used for labelling, and necessary to convey an idea. They include nouns, verbs and most modifiers, such as adjectives and adverbs. Function words are used to connect and organize content words. By themselves function words have little meaning (Pennebaker, 2011), but can reflect objects and events in our lives, since with function words we can change linguistic perspective without even noticing (Pennebaker, 2011). They serve as the cement that holds the content words together (Pennebaker & Chung, 2007). For example, a content word like car can trigger a mental image but the word she, by itself, can not. However, function words – pronouns, or propositions, articles, conjunctions, and auxiliary verbs are used frequently and often for social purposes (Pennebaker & Chung, 2007). When people write using first person singular pronouns, they are focusing the attention of the text on the self, while most other pronouns focus on others (Pennebaker & Chung, 2007). The use of third person pronouns appears to be linked to adaptive coping and physical health benefits (Pennebaker & Chung, 2007). For example, switching from high rates of "I" to high rates of other personal pronouns when writing about upsetting live events, has been linked to health improvements after expressive writing (Campbell & Pennebaker, 2003). Sequence of perspective switching also showed positive health related outcomes, with changing from first person to third person perspective reducing distress associated with memories in people with a diagnosis of dysphoria (Williams & Moulds, 2008). Overall, perspective switching – what happens when someone starts thinking about their own traumatic experience in a self-distanced perspective – appears to result in higher cognitive mechanism word usage (for example, realise and understand; Seih et al., 2011; Pennebaker et al., 1997). Function word usage has also been associated with decreased anxiety symptoms and distress caused by a traumatic event (Mackenzie et al., 2008).

Two other important linguistic features in the study of emotionally charged texts are lexical density and idea density. Lexical density refers to the ratio between the number of content words and the total number of words in a text (Halliday, 1985; Johansson, 2008). Texts with high lexical density usually occur when the writer uses more subordinate clauses, which might be indicative of few function words used (Laufer & Nation, 1995). Because of this, texts with lower density tend to be more understandable (Gregori-Signes & Clavel-Arroitia, 2015). A text can include many different words and still have low lexical density, if it includes a large variety of pronouns and auxiliary words (types of function words) instead of nouns and lexical verbs (Gregori-Signes & Clavel-Arroitia, 2015). These authors have found for example, that lexical density was lower in first year university students in comparison to fourth year students. Finally, idea density (also known as prepositional density) is a measure of how much information was contained in a sentence relative to the number of words used (Chand et al., 2012; Farias et al., 2012). For example, the sentence "I grew up in France" has two prepositions ("I grew up" and "in France"), therefore idea density for this sentence would be 2 divided by 5 (total number of words in the sentence). A study that followed nuns from young adulthood

to old age found that the nuns that showed lower idea density in diaries during young adulthood were more likely to have a cognitive impairment/dementia diagnosis when they grew older (Snowdon et al., 1996).

Several measures have been proposed to measure emotional content during writing, such as emotional impact, emotional valence, and emotional integration. Emotional intensity refers to the arousal associated with an object, event, or situation, giving information about the way participants reflect on their experiences (Bradley & Lang, 1994). For example, if the tone of a text is negative, the intensity of the experience is higher than a text with a more positive tone or value. Differences in levels of affect intensity also reflect differences in the intensity of the emotional experience (Bachorowski & Braaten, 1994). Emotion intensity is important to study in expressive writing since it is related to emotion regulation processes (Dixon-Gordon et al., 2015). Since in this study we are not measuring the intensity of emotions on the writer, but on the reader, we call it emotional impact. Emotional valance has been defined by VandenBos (2007) as the value associated with a stimulus expressed on a continuum from negative to positive, describing the attractiveness (positive valence) or aversiveness (negative valence) of stimuli along a continuum (Costanzi et al., 2019). A previous study using linguistic analysis showed that using a high number of positive words combined with a moderate number of negative words during writing is related to improvements in health (Pennebaker, 1997).

Another important measure is emotional integration, particularly when studying traumatic memories. Emotional integration is a type of "healthy" emotion regulation that involves accessing and expressing negative as well as positive feelings (Roth et al., 2009). It seems that the extent to which emotions are linked with the momories of the event (cognitive-emotional distinctiveness; CED), is negatively associated with PTSD, with those with PTSD showing lower levels of CED compared to those without (Boals & Rubin, 2011).

## 1 Objectives and Aims

In this study, we aimed to examine the role of linguistic features and emotional content during expressive writing, within the homeward of the Writer(s)-Within-community model of writing (Graham, 2018a). Participants were randomly assigned to one of four groups: one of two trauma groups, writing about a traumatic or stressful life experience, either in a self-immersed or self-distanced perspective, and one of two control groups, writing about their daily routine either in a self-immersed or self-distanced perspective. Participants in all four groups were compared on linguistic measures: number of words, different words, function words, different function words, content words, different content words, average word length, function and content word length, number of sentences, lexical density and idea density and positive and negative emotion words; and emotional impact, valence and emotional integration. We expected the linguistic features of

the written texts, particularly content and function words, to show differences between topics and perspectives. More specifically, we expected participants in the trauma groups to write using more function and content words in comparison to the control groups, and the self-distanced trauma group to write using more function words in comparison to the self-immersed trauma group (Hypothesis 1, H1). Regarding emotional word usage, we expected participants in the self-distanced trauma group to write using more positive emotional words in comparison to the self-immersed trauma group, since they would have the opportunity to watch their own experience from an outsider's eyes (Hypothesis 2, H2). Finally, we expected emotional integration to be more present in self-distanced trauma writing (Hypothesis 3, H3).

## 2 Method

### *2.1 Participants*

We recruited 110 undergraduates and randomly assigned them to one of four groups: two trauma groups writing either in the first person ($n$ = 29; range = 17–51 years, $M$ = 20.1, $SD$ = 6.50; 16% male), or in the third person ($n$ = 28; range = 18–26 years, $M$ = 19.6, $SD$ = 2.12; 10,7% male), two control groups writing either in the first person ($n$ = 28; range = 18–39 years, $M$ = 19.9, $SD$ = 4.54; 21.7% male) or in the third person ($n$ = 29; range = 18–53 years, $M$ = 21.1, $SD$ = 6.84; 20,7% male). Permission from the ethics committee at the University of Porto, Faculty of Psychology and Education Science was received, to conduct the study (process 3–12/2016). All participants were native Portuguese speakers and received course credits for their participation.

### *2.2 Instruments*

#### HandSpy 3.0

To record and analyse the participants written text we used HandSpy 3.0 (Alves et al., 2019) and Neo smartpens M1 (Neo Smartpen Inc., Scotts Valley, CA, USA). HandSpy measured: (1) the linguistic features of the written texts: number of words, different words, function words, different function words and content words, different content words, average word length, function and content word length; number of sentences; lexical density and idea density; and (2) the positive and negative emotional words. The emotional vocabulary analysis was done using the Portuguese version of EMOTAIX (Costa, 2012; Piolat & Bannour, 2009), which is included in HandSpy 3.0. EMOTAIX lists Portuguese emotion words according to their positive or negative valence. We calculated the total positive emotion words and total of negative emotion words used by each group.

### Trauma Categories

Participants wrote about a variety of traumatic experiences classified by two independent judges. From the extensive content analysis of the texts, two independent judges identified a set of 11 categories representing the emotional content written by the participants. The 11 categories that emerged were: bullying, death, disease, romantic and intimacy relationships, harassment or sexual abuse, threat to physical integrity, domestic violence, divorce and/or parental conflicts, family conflicts, mental health at risk and others topics. As expected, for the control group, the only category that emerged was "daily routine". Regarding the prevalence of the emotional categories in the sample, mental health at risk was at 16%, death 14%, threat to physical integrity 12%, parents' conflicts and divorce 11%, romantic and intimacy relations 9%; the categories harassment and sexual abuse and domestic violence and family conflicts were both at 7%, followed by bullying (5%) and disease (4%). The category "other topics" represented 9% of the emotional texts.

### Emotional Impact and Valence and Integration

To measure emotional impact, valence and emotional integration, two independent blind judges were asked to analyse all written texts. Emotional impact and valence was measured using the Self-Assessment Manikin (SAM) scale (Bradley & Lang, 1994). SAM is an affective assessment scale showing a manikin depicting different levels of emotional activation. SAM was used to measure the impact on the reader of emotional content and valence of the emotion in the text. Emotional impact was measured through a set of graphic figures that represent a continuum according to excitement felt in the body, in a scale of 1 (most relaxed manikin) to 5 (great disturbance). Emotional valence was measured using 5 manikins ranging from a very sad face to a very happy one. For emotional integration, the judges rated the level of emotional integration that was shown by each text, using a Likert scale, from 1 (not integrated at all) to 5 (completely integrated).

Agreement between judges was measured using intraclass correlation coefficient was calculated for all measures. The correlation was significantly high for both emotional impact, (ICC = .978) and valence of emotion (ICC = .904). Finally, the correlation regarding emotional integration (ICC = .984) was also high.

## *2.3 Procedure*

The data collection took place in a dedicated laboratory. Prior to data collection, students were informed about the experimental procedures and informed of the voluntary nature of their participation. Anonymity of their participation and confidentiality of all data were collected, along with informed consent.

After signing the informed consent participants were taken to an isolated part of the room, and received a closed envelope containing the writing prompts. Researchers conducting the experiment were blind regarding the participants' experimental condition, as the envelopes were shuffled. Participants in the emotional group wrote about the most traumatic experience of their life, and those in the control group wrote about their daily routine, either in the first person singular (I; self-immersed) or using third person pronouns (She/He; self-distanced). Every participant wrote for 15 min. All participants were debriefed at the end.

### *2.4 Data Analysis*

The emotional and control group were compared on linguistic features, emotional words used, emotional impact, emotional valence and emotional integration. Before data analysis four participants were removed from the analysis for not complying to the instructions given during the writing task (e.g. writing in the first person when asked to write in the third person). Outliers were identified through box plot analysis based on the interquartile range (1.5 × IQR) and removed from the database by assigning missing values to the outlier variables. All measures were analysed using 2 × 2 ANOVA, with topic (emotional group, control group) and perspective ("I", "She/He") as between-subject factors, with the α threshold set at .05.

## 3 Results

### *3.1 Descriptive Statistics*

Table 1 displays descriptive statistics, as well as *F* and *p* values, for all linguistic features, including positive and negative emotion words, emotional impact and valence, and emotional integration.

### *3.2 Linguistic Features (H1)*

To analyse the linguistic features of the written text we conducted a 2 × 2 ANOVA, with the linguistic features as dependent variables and *topic* and *perspective* as fixed factors. Results showed that for *topic* there were no between subject differences for number of words ($F < 1$), different words ($F < 1$), function words ($F < 1$), content words,

$F(1, 106) = 2.91$, $p = .09$, different content words, $F(1, 106) = 2.91$, $p = .09$, number of sentences ($F < 1$) and idea density ($F < 1$). There were significant between

**Table 1** Descriptive statistics, *f* and *p* values for linguistic features, positive and negative emotion words, emotional impact, valance, and emotional integration

| | Trauma group | | Control group | |
|---|---|---|---|---|
| | Self-immersed ($n = 29$) | Self-distanced ($n = 26$) | Self-immersed ($n = 28$) | Self-distanced ($n = 27$) |
| Measures | *M (SD)* | *M (SD)* | *M (SD)* | *M (SD)* |
| Number of words | 321 (64.11) | 338.58 (74.25) | 333.04 (68.17) | 345.7 (71.77) |
| Different words | 140.97 (22.1) | 143.85 (21.82) | 143.14 (23.10) | 142.67 (24.18) |
| Function words | 167.93 (37.51) | 175.65 (42.62) | 166.39 (39.82) | 175.19 (39.21) |
| Different function words | 39.79 (4.79) | 38.85 (5.56) | 32.75 (5.89) | 30.89 (6.30) |
| Content words | 152.52 (30.91) | 162.31 (34.27) | 166.04 (30.51) | 169.93 (34.25) |
| Different content words | 107.17 (20.38) | 111.27 (18.94) | 115.32 (19.36) | 116.26 (21.90) |
| Word length | 4.06 (0.22) | 4.12 (0.24) | 3.86 (0.16) | 3.93 (0.20) |
| Function word length | 2.53 (0.15) | 2.5 (0.15) | 2.16 (0.14) | 2.19 (0.16) |
| Content word length | 5.74 (0.38) | 5.82 (0.39) | 5.55 (0.25) | 5.72 (0.35) |
| Number of sentences | 17.31 (5.23) | 16.12 (4.78) | 15.82 (6.03) | 17.04 (5.29) |
| Lexical density | 0.48 (0.04) | 0.48 (0.03) | 0.50 (0.03) | 0.49 (0.02) |
| Idea density | 0.37 (0.10) | 0.40 (0.18) | 0.34 (0.10) | 0.40 (0.10) |
| Positive emotion words | 6.83 (5.33) | 12.5 (5.65) | 3.96 (2.81) | 5.44 (4.26) |
| Negative emotion words | 15.03 (6.00) | 16.15 (7.00) | 0.71 (1.65) | 1.19 (1.82) |
| Emotional impact | 3.48 (0.76) | 3.75 (0.90) | 1.09 (0.39) | 1.07 (0.23) |
| Emotional valence | 1.79 (.45) | 1.73 (.53) | 3.04 (.34) | 3.03 (.31) |
| Emotional integration | 2.29 (1.85) | 2.15 (1.41) | | |

Descriptive statistics for all linguistic features, positive and negative emotion words, emotional impact, valance, and emotional integration

group differences for *topic* in different function words, $F(1, 106) = 48.35$, $p < .001$, $\eta^2_p = .31$ with the trauma group writing, on average 7.50 more distinct function words. In addition, average word length, $F(1, 106) = 22.67$, $p < .001$, $\eta^2_p = .18$, average function, $F(1, 106) = 142.69$, $p < .001$, $\eta^2_p = .57$ and content word length, $F(1, 106) = 5.08$, $p = .03$, $\eta^2_p = .05$, had significant between group differences, with the trauma groups writing longer function and content words in comparison to the control group. Lexical density was lower in the trauma group, $F(1, 106) = 9.70$, $p = .002$, $\eta^2_p = .08$, in comparison to the control group. Regarding *perspective* there were no between subject differences for number of words $F(1, 106) = 1.3$, $p = .257$, different words ($F < 1$), function words, $F(1, 106) = 1.19$, $p = .278$, content words, $F(1, 106) = 1.22$, $p = .272$, different function words, $F(1, 106) = 1.69$, $p = .196$, different content words ($F < 1$), average word length, $F(1, 106) = 2.23$, $p = .278$, average

function word, ($F < 1$) and content word length, $F(1, 106) = 3.63$, $p = .06$, number of sentences ($F < 1$), and lexical density, ($F < 1$). There were only significant between groups differences in *perspective* for idea density, $F(1, 106) = 4.37$, $p = .04$, $\eta^2_p = .04$, with the third person groups showing higher idea density. For interaction between *condition* and *perspective* no linguistic feature was significantly different.

## Positive and Negative Emotional Words (H2)

To analyse the positive and negative emotional words in the text we conducted a 2 × 2 ANOVA, with total positive emotional words and total negative emotional words as dependent variables and *topic* and *perspective* as fixed factors. The between-subjects ANOVA showed that there were significant differences between *topic* (trauma vs. control), for both positive, $F(1, 106) = 31.39$, $p < .001$, $\eta^2_p = .23$, and negative emotional words, $F(1, 106) = 260.68$, $p < .001$, $\eta^2_p = .71$, with the trauma group using on average 4.96 more positive words and 14.65 more negative words in comparison to the control group. In addition there were significant between group differences for *perspective* in positive emotional words, $F(1, 106) = 16.32$, $p < .001$, $\eta^2_p = .13$, but not for negative words, ($F < 1$). Results also showed a significant interaction between *topic* and *perspective* for positive emotional words, $F(1, 106) = 5.61$, $p = .02$, $\eta^2_p = .05$, with the self-distanced trauma group using on average 5.8 more positive words (see Fig. 1). The same was not found for negative words ($F < 1$).

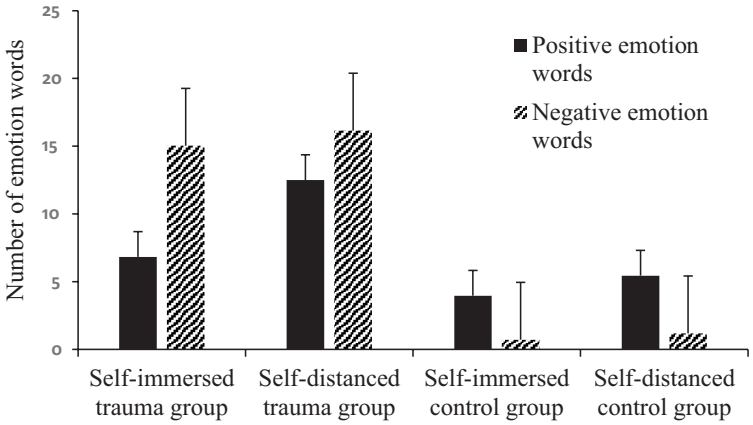



**Fig. 1** Average total number of positive and negative emotion words used for all groups. Note. Average total number of positive and negative emotion words used for both trauma/expressive groups (self-immersed and self-distanced) and both control groups (self-immersed and self-distanced). The separate bars represent the trauma and the control group. Greater values indicate higher number of words used. Line bars represent standard error of the mean

### *3.3 Emotional Impact, Valence and Emotional Integration (H3)*

Regarding the analysis for emotional impact, the test of between subjects effects showed significant differences between *topic*, $F(1,106) = 443.59$, $p < .001$, $\eta^2_p = .81$, with *perspective* showing similar between subject results, $F(1,106) = 1.09$, $p = .298$. Valence of emotional valence, showed similar results, with significant between subject effects of *condition*, $F(1,106) = 257.31$, $p < .001$, $\eta^2_p = .709$, while *perspective* did not show significant differences ($F < 1$). For emotional integration, a 2 × 2 ANOVA showed significant between subject differences for *topic*, $F(1,106) = 914.44$, $p < .001$, $\eta^2_p = .896$, while *perspective* did not ($F < 1$).

## 4 Discussion

Using the Writer(s)-Within-community model of writing (Graham, 2018a) and the expressive writing paradigm we designed the current study to test if topic and perspective had an impact on the linguistic features and emotional content of the written texts. We expected the participants in the trauma groups to show differences in linguistic features, particularly in content and function words in comparison to the control groups. We also expected the self-distanced trauma group to write using more function words in comparison to the self-immersed trauma group (H1). In addition, we expected participants in the self-distanced trauma group to use more positive emotional words during writing in comparison to the self-immersed trauma group (H2). Finally, we hypothesised that emotional integration would be more prevalent in the self-distanced trauma writing group, in comparison to the other three groups (H3).

### *4.1 Linguistic Features (H1)*

Results showed that the trauma writing groups used a higher number of different function words in comparison to the control groups. There was no significant difference in different content words used. Frequent change of pronouns (a type of function word) may be linked to improved health. For example, Campbell and Pennebaker (2003) found that those who frequently switched the use of personal pronouns, particularly from first to third person showed greater health improvements in the months following, compared to those who did not. These improvements have been linked to the notion that first person singular pronouns suggest attention on the self, while most other pronouns suggest that the person is attending to other individuals, and the more people make reference to others, the healthier they become (Chung & Pennebaker, 2007). Function word usage has also been liked to decreased anxiety

symptoms and distress directly caused by the traumatic event (Mackenzie et al., 2008). In this study, trauma group participants used more diverse function words, which can be indicative of a complex storytelling emerging from expressive writing. We propose that this might be one of the reasons why expressive writing keeps showing positive health effects (see Frattaroli, 2006). Future studies should continue to investigate if function word usage is one of the reasons why expressive writing leads to health improvements, as Pennebaker & Smyth (2016) has previously suggested.

Lexical density results further show that content words usage was lower in the trauma groups. Texts with lower density usually indicate that the writers used more pronouns and auxiliary words (types of function words) instead of nouns and lexical verbs, for example, proficient university level writers have been shown to write texts with lower lexical density when compared to less proficient writers (Gregori-Signes & Clavel-Arroitia, 2015).

In addition, self-distanced groups had higher idea density than self-immersed groups. Higher idea density has been shown to act as a possible indicator of cognitive reserve (Engelman, et al., 2010). The opposite (lower idea density) has been used as a marker of likelihood for dementia (Engelman, et al., 2010; Snowdon et al., 1996), and decline in cognitive functioning (Farias et al., 2012). It appears that writing about a traumatic event can lead to more diversity of function word usage and higher idea density, which might reflect improved well-being and higher cognitive reserve, respectively. Previous studies (Sandry et al., 2015; Sandry & Sumowski, 2014) have linked cognitive reserve with both working memory and long-term memory, suggesting that working memory might be the mechanism underlying the positive relationship between cognitive reserve and long-term memory. It appears that self-distanced writing could possibly have an impact on both working memory and long-term memory. This makes sense, as working memory houses the processes involving executive control of remembering and describing (Zimmerman & Risemberg, 1997) and long-term memory houses the memories associated with "traumatic" experiences (Graham, 2018a). Future studies should analyse word usage in more detail since it can provide further evidence for the benefits of expressive writing and self-distanced expressive writing.

## *4.2 Positive and Negative Emotional Words (H2)*

As far as we know this is the first study that analysed emotional word usage in self-distanced expressive writing. Our results showed that the trauma writing groups used on average more negative and positive words in comparison to the control groups. More negative word usage in trauma groups in comparison to control groups is to be expected, since the participants in the trauma group wrote about a traumatic event, and the control group about their daily routine, which was not expected to include significant emotion word usage, if at all. These results are similar to Cohn et al. (2004), that analysed blog entries following 9/11. These entries showed very

high use of negative emotions following the event, along with a drop of positive emotion word usage. However, positive emotion word usage increased to higher than before the event, 10 days after.

In our study, the self-distanced trauma group used on average more positive words in comparison to the self-immersed trauma writing group. An increase of positive emotional word usage was expected, since writing in third-person perspective is associated with new interpretations of events and the promotion of insight and closure (Kross et al., 2012). A previous study showed different results, finding lower emotional reactivity during self-distanced writing, since third-person pronouns were associated with fewer use of anxiety words during writing (Giovanetti et al., 2019). More research is needed that measures emotion word usage during writing. In the WWC model, Graham (2018a) proposed that positive and negative emotions can help store information in long-term memory, to moderate the relationship between emotions and writing. This is because emotions influence what a writer decides their final text will be (Pekrun et al., 2007).

### *4.3 Emotional Impact, Valence & Emotional Integration (H3)*

Emotional impact, valence and integration only showed significant differences between the trauma and the control groups. This was to be expected since the control group was not supposed to write an emotional text. These findings are nevertheless relevant for the use of both self-immersed and self-distanced writing as possible tools to re-think a traumatic event, since in our study, both appear to have equivalent impact, valence levels and levels of integration of the event. Emotion impact is important to consider when comparing the use of emotion regulation strategies in response to different emotions, since it plays a role in emotion regulation (Dixon-Gordon et al., 2015). Emotional integration shows the extent to which the traumatic event has become integrated into one's life (Boals & Rubin, 2011), and in our study, perspective did not affect integration.

Our study should be read keeping in mind the following limitations: First, the participants reported on a variety of traumatic experiences. Future studies might consider using a population sharing similar traumatic events. Second, the expressive writing prompts used were heavily based on the original expressive writing paradigm. Using different and/or novel expressive writing prompts could shed light on the role of the writing instruction on the effects of writing.

## 5 Conclusion

With this study we wanted to examine the role of the linguistic features and emotional content during expressive writing. To our knowledge, this is the first time linguistic and emotional measures during expressive writing have been compared

and integrated into the Writer(s)-Within-community model of writing. This model proposes three mechanisms involved in writing: Long-term memory resources, control mechanisms and production processes, all of which moderated by emotions, personality traits and physical states. Our study, viewed under the light of this model, shed light on how working memory and long-term memory might affect writing style (different function words, lexical density, idea density, and positive and negative emotional words).

We found that the trauma groups wrote using a higher number of different function words and lower lexical density and the self-distanced groups showing higher idea density. These results might be indicators of more control mechanisms at work during writing, particularly working memory. In addition, the self-distanced trauma group wrote using a higher number of positive words, in comparison to the immersion group, suggesting that expressive writing might affect long-term memory resources. This is a push forward in the field of expressive writing as it might encourage others to start using a non-traditional expressive writing prompt: self-distanced expressive writing. In addition, we hope it might encourage other researchers to analyse the linguistic features and emotional vocabulary usage during writing.

## References

Alves, R. A., Leal, J. P., & Limpo, T. (2019). Using HandSpy to study writing in real time: A comparison between low- and high-quality texts in grade 2. In E. Lindgren & K. Sullivan (Eds.), *Observing writing: Insights from keystroke logging and handwriting* (p. 50). Brill. https://doi.org/10.1163/9789004392526_004

Ayduk, O., & Kross, E. (2010). From a distance: Implications of spontaneous self-distancing for adaptive self-reflection. *Journal of Personality and Social Psychology, 98*(5), 809–829. https://doi.org/10.1037/a0019205

Bachorowski, J.-A., & Braaten, E. (1994). Emotional intensity: Measurement and theoretical implications. *Personality and Individual Differences, 17*, 191–199. https://doi.org/10.1016/0191-8869(94)90025-6

Boals, A., & Rubin, D. C. (2011). The integration of emotions in memories: Cognitive-emotional distinctiveness and posttraumatic stress disorder. *Applied Cognitive Psychology, 25*(5), 811–816. https://doi.org/10.1002/acp.1752

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 49–59. https://doi.org/10.1016/0005-7916(94)90063-9

Campbell, R. S., & Pennebaker, J. W. (2003). The secret life of pronouns: Flexibility in writing style and physical health. *Psychological Science, 14*(1), 60–65. https://doi.org/10.1111/1467-9280.01419

Chand, V., Baynes, K., Bonnici, L. M., & Farias, S. T. (2012). A rubric for extracting idea density from oral language samples. *Current Protocols in Neuroscience, 10*(10), 5. https://doi.org/10.1002/0471142301.ns1005s58

Chung, C., & Pennebaker, J. W. (2007). The psychological functions of function words. In K. Fiedler (Ed.), *Social communication* (pp. 343–359). Psychology Press.

Cohn, M. A., Mehl, M. R., & Pennebaker, J. W. (2004). Linguistic markers of psychological change surrounding September 11, 2001. *Psychological Science, 15*(10), 687–693. https://doi.org/10.1111/j.0956-7976.2004.00741.x

Costa, S. (2012). *Adaptação e teste de uma base lexical de palavras emocionais para o português europeu (EMOTAIX.PT)* [Adaptation and testing of a lexical database of emotional words to european portuguese (EMOTAIX.PT)]. Unpublished master's thesis. Faculty of Psychology and Education Sciences, University of Porto.

Costanzi, M., Cianfanelli, B., Saraulli, D., Lasaponara, S., Doricchi, F., Cestari, V., & Rossi-Arnaud, C. (2019). The effect of emotional valence and arousal on visuo-spatial working memory: Incidental emotional learning and memory for object-location. *Frontiers in Psychology, 10*(2587). https://doi.org/10.3389/fpsyg.2019.02587

D'Mello, S., & Mills, C. (2014). Emotions while writing about emotional and non-emotional topics. *Motivation and Emotion, 38*(1), 140–156. https://doi.org/10.1007/s11031-013-9358-1

Diamond, A. (2006). The early development of executive functions. In E. Bialystok & F. I. M. Craik (Eds.), *Lifespan cognition: Mechanisms of change* (pp. 70–95). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195169539.003.0006

Dixon-Gordon, K. L., Aldao, A., & De Los Reyes, A. (2015). Emotion regulation in context: Examining the spontaneous use of strategies across emotional intensity and type of emotion. *Personality and Individual Differences, 86*, 271–276. https://doi.org/10.1016/j.paid.2015.06.011

Engelman, M., Agree, E., Meoni, L., & Klag, M. (2010). Propositional density and cognitive function in later life: Findings from the precursors study. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 65*, 706–711. https://doi.org/10.1093/geronb/gbq064

Farias, S. T., Chand, V., Bonnici, L., Baynes, K., Harvey, D., Mungas, D., Simon, C., & Reed, B. (2012). Idea density measured in late life predicts subsequent cognitive trajectories: Implications for the measurement of cognitive reserve. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 67*(6), 677–686. https://doi.org/10.1093/geronb/gbr162

Frattaroli, J. (2006). Experimental disclosure and its moderators: A meta-analysis. *Psychological Bulletin, 132*(6), 823–865. https://doi.org/10.1037/0033-2909.132.6.823

Giovanetti, A. K., Revord, J. C., Sasso, M. P., & Haeffel, G. J. (2019). Self-distancing may be harmful: Third-person writing increases levels of depressive symptoms compared to traditional expressive writing and no writing. *Journal of Social and Clinical Psychology, 38*(1), 50–69. https://doi.org/10.1521/jscp.2019.38.1.50

Graham, S. (1982). Composition research and practice: A unified approach. *Focus on Exceptional Children, 14*, 10.17161/fec.v14i8.7447.

Graham, S. (2018a). A revised writer(s)-within-community model of writing. *Educational Psychologist, 53*(4), 258–279. https://doi.org/10.1080/00461520.2018.1481406

Graham, S. (2018b). A writer(s)-within-community model of writing. In C. Bazerman, V. W. Berninger, D. Brandt, S. Graham, J. Langer, S. Murphy, & M. Schleppegrell (Eds.), *The lifespan development of writing* (pp. 272–325). National Council of English.

Gregori-Signes, C., & Clavel-Arroitia, B. (2015). Analysing lexical density and lexical diversity in university students' written discourse. *Procedia – Social and Behavioral Sciences, 198*, 546–556. https://doi.org/10.1016/j.sbspro.2015.07.477

Grossmann, I., & Kross, E. (2010). The impact of culture on adaptive versus maladaptive self-reflection. *Psychological Science, 21*(8), 1150–1157. https://doi.org/10.1177/0956797610376655

Halliday, M. (1985). *An introduction to functional grammar* (1st ed.). Edward Arnold.

Johansson, V. (2008). Lexical diversity and lexical density in speech and writing: A developmental perspective. *Working Papers/Lund University, Department of Linguistics and Phonetics, 53*, 61–79.

Kross, E., & Ayduk, O. (2017). Self-distancing: Theory, research, and current directions. In J. M. Olson (Ed.), *Advances in experimental social psychology* (pp. 81–136). Elsevier Academic Press.

Kross, E., Ayduk, O., & Mischel, W. (2005). When asking "why" does not hurt: Distinguishing rumination from reflective processing of negative emotions. *Psychological Science, 16*(9), 709–715. https://doi.org/10.1111/j.1467-9280.2005.01600.x

Kross, E., Gard, D., Deldin, P., Clifton, J., & Ayduk, O. (2012). "Asking why" from a distance: Its cognitive and emotional consequences for people with major depressive disorder. *Journal of Abnormal Psychology, 121*, 559–569. https://doi.org/10.1037/a0028808

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics, 16*, 307–322. https://doi.org/10.1093/applin/16.3.307

Mackenzie, C. S., Wiprzycka, U. J., Hasher, L., & Goldstein, D. (2008). Seeing the glass half full: Optimistic expressive writing improves mental health among chronically stressed caregivers. *British Journal of Health Psychology, 13*(1), 73–76. https://doi.org/10.1348/135910707x251153

Park, J., Ayduk, O., & Kross, E. (2016). Stepping back to move forward: Expressive writing promotes self-distancing. *Emotion, 16*(3), 349–364. https://doi.org/10.1037/emo0000121

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 13–36). Academic. https://doi.org/10.1016/B978-012372545-5/50003-4

Pennebaker, J. W. (1997). Writing about emotional experiences as a therapeutic process. *Psychological Science, 8*(3), 162–166. https://doi.org/10.1111/j.1467-9280.1997.tb00403.x

Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. Bloomsbury Press/Bloomsbury Publishing.

Pennebaker, J. W., & Beall, S. K. (1986). Confronting a traumatic event: Toward an understanding of inhibition and disease. *Journal of Abnormal Psychology, 95*(3), 274–281. https://doi.org/10.1037/0021-843X.95.3.274

Pennebaker, J. W., & Chung, C. K. (2007). Expressive writing, emotional upheavals, and health. In *Foundations of health psychology* (pp. 263–284). Oxford University Press.

Pennebaker, J. W., & Smyth, J. (2016). *Opening up by writing it down: The healing power of expressive writing* (3rd ed.) Guilford.

Pennebaker, J. W., Mayne, T. J., & Francis, M. E. (1997). Linguistic predictors of adaptive bereavement. *Journal of Personality and Social Psychology, 72*(4), 863–871. https://doi.org/10.1037/0022-3514.72.4.863

Piolat, A., & Bannour, R. (2009). EMOTAIX : Un scénario de tropes pour l'identification automatisée du lexique émotionnel et affectif. *L'Année psychologique, 109*, 655–698. https://doi.org/10.3917/anpsy.094.0655

Pirnay-Dummer, P. (2016). Linguistic analysis tools. In C. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed.). Guilford Press.

Rakesh, G., Morey, R. A., Zannas, A. S., Malik, Z., Marx, C. E., Clausen, A. N., Kritzer, M. D., & Szabo, S. T. (2019). Resilience as a translational endpoint in the treatment of PTSD. *Molecular Psychiatry, 24*(9), 1268–1283. https://doi.org/10.1038/s41380-019-0383-7

Roth, G., Assor, A., Niemiec, C. P., Ryan, R. M., & Deci, E. L. (2009). The emotional and academic consequences of parental conditional regard: Comparing conditional positive regard, conditional negative regard, and autonomy support as parenting practices. *Developmental Psychology, 45*(4), 1119–1142. https://doi.org/10.1037/a0015272

Sandry, J., & Sumowski, J. F. (2014). Working memory mediates the relationship between intellectual enrichment and long-term memory in multiple sclerosis: An exploratory analysis of cognitive reserve. *Journal of the International Neuropsychological Society: JINS, 20*(8), 868–872. https://doi.org/10.1017/S1355617714000630

Sandry, J., DeLuca, J., & Chiaravalloti, N. (2015). Working memory capacity links cognitive reserve with long-term memory in moderate to severe TBI: A translational approach. *Journal of Neurology, 262*(1), 59–64. https://doi.org/10.1007/s00415-014-7523-4

Seih, Y. T., Chung, C. K., & Pennebaker, J. W. (2011). Experimental manipulations of perspective taking and perspective switching in expressive writing. *Cognition and Emotion, 25*(5), 926–938. https://doi.org/10.1080/02699931.2010.512123

Smyth, J. M. (1998). Written emotional expression: Effect sizes, outcome types, and moderating variables. *Journal of Consulting and Clinical Psychology, 66*(1), 174–184. https://doi.org/10.1037//0022-006x.66.1.174

Smyth, J. M., & Pennebaker, J. W. (2008). Exploring the boundary conditions of expressive writing: In search of the right recipe. *British Journal of Health Psychology, 13*(1), 1–7. https://doi.org/10.1348/135910707X260117

Snowdon, D., Kemper, S., Mortimer, J., Greiner, L., Wekstein, D., & Markesbery, W. (1996). Linguistic ability in early life and cognitive function and Alzheimer's disease in late life: Findings from the nun study. *JAMA: The Journal of the American Medical Association, 275*, 528–532. https://doi.org/10.1001/jama.275.7.528

Valtonen, J. (2020). The health benefits of autobiographical writing: An interdisciplinary perspective. *The Journal of Medical Humanities*. https://doi.org/10.1007/s10912-020-09631-9

VandenBos, G. R. (Ed.). (2007). *APA dictionary of psychology*. American Psychological Association.

Williams, A. D., & Moulds, M. L. (2008). Manipulating recall vantage perspective of intrusive memories in dysphoria. *Memory, 16*(7), 742–750. https://doi.org/10.1080/09658210802290453

Yasinski, C., Hayes, A. M., & Laurenceau, J.-P. (2016). Rumination in everyday life: The influence of distancing, immersion, and distraction. *Journal of Experimental Psychopathology, 7*(2), 225–245. https://doi.org/10.5127/jep.042714

Zimmerman, B. J., & Risemberg, R. (1997). Becoming a self-regulated writer: A social cognitive perspective. *Contemporary Educational Psychology, 22*(1), 73–101. https://doi.org/10.1006/ceps.1997.0919

# Attitudinal Evaluation in Written Peer Feedback: An Appraisal Study


**Xinghua Liu and Xuan Jiang**

**Abstract** This paper aims to reveal textual evaluations in students' written peer feedback. One English reading and writing course in China was connected with one writing-as-processes course in an American university. During the 10 weeks of the research, 21 out of 24 students from the American university, many of whom were ESL writers, and 20 participants from the university in China, all of whom were EFL writers, read the same article and wrote with the same essay prompt. Students were assigned with one overseas partner (or two partners on rare occasions) for peer feedback and then revised their essay with their decisions about the feedback accordingly. From this peer feedback activity, four corpora consisting of Chinese and American university students' written comments and responses were established. The Attitude system within the Appraisal framework (Martin and White, Language of evaluation: appraisal in English. Palgrave Macmillan, 2005) developed within the Systemic Functional Linguistics paradigm was adopted as the analytical tool. The study found that while giving feedback, both Chinese and American students exhibited a prevalence of Appreciation items over Affect and Judgement ones; American students employed significantly more positive Affect items than Chinese students did, and Chinese students used more negative Judgement items than American students did. When responding to peers' written comments, Chinese students foregrounded Judgement values while American students focused more on Appreciation subcategory; Chinese students used significantly more positive Affect items and Judgement items than American students did. We conclude that the nature of writing task, cultural traditions and personal learning experiences are some of the factors contributing to student reviewers' use of evaluative resources in their written comments.

X. Liu (✉)
School of Foreign Languages, Yang Yong Man Building, Shanghai Jiao Tong University, Shanghai, China
e-mail: liuxinghua@sjtu.edu.cn

X. Jiang
Florida International University, Miami, FL, USA



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_20

**Keywords** Peer feedback · Textual evaluation · Cross-cultural rhetoric · Appraisal framework

## 1 Introduction

Peer feedback, which is named differently as "peer review", "peer editing", "peer evaluation", or "peer response" in various literature, refers to the practice of students assuming responsibilities in commenting on each other's drafts in written and oral formats during the process of writing (Liu & Hansen, 2002, p. 1). It has been informed by various theoretical sources, such as process theory of writing, collaborative learning, communicative language teaching, and sociocultural theories (e.g. Vygotsky's Zone of Proximal Development). Peer feedback has been long and widely practiced in first language (L1) and second language (L2) classrooms, and has been reported to have cognitive, social, linguistic and affective benefits for language learning (for details, see Hyland & Hyland, 2006; Liu & Hansen, 2002).

The past decades have witnessed an exponential increase in the study of peer feedback from various perspectives, such as the effect of peer feedback on L2 student revisions and writing quality (Liu & Sadler, 2003; Lu & Law, 2012), the effects of training on students' peer feedback performance (Min, 2006, 2008; Rahimi, 2013), student interaction and participation in peer feedback (Villamil & de Guerrero, 1998), the comparison between teacher and peer feedback behaviors (Zhao, 2010), and computer-mediated peer feedback (Yang, 2015; Woo et al., 2013). However, few studies have examined students' written peer feedback as a site of interaction between feedback givers and feedback receivers, nor examined how student reviewers would employ evaluative strategies linguistically.

It has been widely acknowledged that writing is seen as a site of interaction between writers and readers, and through written text, writers construct solidarity and alignment with potential or target readers (Thompson, 2001). This kind of textual evaluation and interaction has been a well-established research area in academic context and different scholars for different purposes employ different terms to account for this phenomena, such as *attitude* (Halliday, 1994), *stance* (Biber & Finegan, 1989; Hyland, 1999), *evaluation* (Hunston & Thompson, 2000), *metadiscourse* (Crismore, 1989) and *Appraisal* (Martin, 2000). Though addressed from divergent approaches, previous studies on evaluation and interaction (e.g. Hood, 2004; Hyland, 2002; Thompson, 2001) have all shown that a proper management of interpersonal language is essential in constructing a critical voice and building persuasion.

Written peer feedback is interactive in nature, as that is the place where readers take on the role of evaluators and assess writers' performance in an explicit manner. However, little research has examined written peer feedback from the perspective of evaluation.

The current project was conducted between a university in China and a Hispanic-Serving Institution (HSI) in the U.S. The data were from students' written comments on their overseas peers' essays over the course of 10 weeks from both writing-related classes. The purpose of this project was to study how Chinese and American university students adopted evaluative strategies in their cross-cultural written peer feedback activities.

## 2 Literature Review

### *2.1 Cross-cultural Rhetoric Research*

Cross-cultural rhetoric (hereafter CR) is an area of research in second language acquisition that "examines differences and similarities in ESL (English as a second language) and EFL (English as a foreign language) writing across languages and cultures as well as across such different contexts as education and commerce" (Connor, 2002, p. 493). It originates from Kaplan's (1966) pioneering work on the examination of paragraph organization of ESL students' compositions. CR research has now become the flourishing domain of applied linguistics and one of the most studied areas within second language writing. Among them, the comparison between Chinese and English writing is one of the most studied areas.

One of the myths haunting cross-cultural rhetoric research between Chinese and English is that English writing by Chinese writers lacks critical stance and/or personal voice. Traditional social-cultural perspective has held that collectivism is a cultural pattern in China, and the Chinese people belong to groups of some kind which mainly function to maintain group cohesion and harmony. For example, Carson and Nelson (1994, 1996) found that in American university writing conferences, Chinese students were normally not able to respond critically and constructively, but instead they took harmony-maintenance, or face-saving strategies which might not be helpful for writing development. Another widely circulated article is Shen (1989), which presented a vivid narration of his struggling process of learning English composition in America. He argued that Chinese society is collectivist-oriented and must always subordinate the personal identity "I" to the collective body "we" as being "timid, humble, modest" (Shen, 1989, p. 462). After studying for many years in an American individual-centered culture, he put on a new English "I" who was "confident, assertive, and aggressive" (ibidem). However, doubts are raised as to what extent these findings remain valid for current Chinese society, considering these studies were conducted more than three decades ago. While describing the current education system in China, Jin and Cortazzi (2011, p.2) rightfully pointed out that, "since the pace of development is rapid, and education in China is expanding and undergoing reforms in a context of social change, many specific figures will be outdated as soon as they are

published." Although they wrote specifically about the constantly rapid development in the education system, this is actually the case for almost every aspect of Chinese society as a whole. An up-to-date account of Chinese students' English writing is therefore warranted.

There have been increasing efforts to seek a linguistic account of the issue of personal voice in writing. For example, Wu and Rubin (2000) examined the relationship between writers' collectivist/individualist tendencies and their uses of a multitude of linguistic features. In their study, 40 Taiwanese university EFL students wrote in both English and Chinese, and 40 first-year American students wrote in English. All participants were assessed for their collectivist/individualist tendencies through a sentence completion test. All essays were coded based on a set of categories alleged to indicate collectivist/individualist world views, such as directness/indirectness, personal disclosure, and assertiveness. For example, the placement of the thesis statement indicates the level of directness/indirectness, and the use of first and third person pronouns signify the degree of personal disclosure and collectivist concepts. The study found that Taiwanese students' Chinese writing showed a relative indirectness, contained expressions of Confucian principles of humaneness and collective virtue, and placed great reliance on proverbs and other canonical expressions, whereas their English writing was characterized by little self-disclosure and low level of assertiveness.

The debate on whether Chinese students can inscribe their personal voice or express their explicit attitudes in their English writing is sure to continue, but previous studies are limited by having predominantly taken students' timed writing in response to given prompt as the source of investigation. Alternative data sources, such as students' written peer feedback can also be usefully considered to render a more comprehensive view on this debate.

## *2.2 Linguistic Study of Evaluation in Writing: Appraisal Framework*

Evaluation is the broad umbrella term for the expression of the writer's "attitude or stance towards, viewpoint on, or feelings about the entities or propositions that he or she is talking about" (Hunston & Thompson, 2000, p. 5). Textual evaluations can fulfill multiple functions. They can express the writer's opinions and reflect the value system of the writer and their related community, align readers, and organize the discourse (Hunston & Thompson, 2000, p. 6). As mentioned above, there are various approaches for the study of textual interaction and evaluation. The one adopted for the current project is the Appraisal framework (Martin & White, 2005), evolved from Systemic Functional Linguistics (SFL) paradigm (Halliday & Matthiessen, 2004) which sees any language use perform three meta-functions simultaneously, namely ideational, interpersonal, and textual. The Appraisal

framework was developed to describe the various ways of linguistic realization of interpersonal meanings in language use.

The Appraisal framework is useful for the study of textual interaction for two practical considerations. First, the Appraisal framework has been by far the most comprehensive linguistic analytical framework for the analysis of interpersonal accounts. Hyland (2005) describes Appraisal as the "most systematic approach" for the study of evaluative language uses and states that it "offers a typology of evaluative resources available in English" (p. 174). Second, previous studies concerning the interpersonal features in writing have mainly focused on lexico-grammar analysis with few taking on a discourse-semantic approach. Although both approaches deal with meanings, the lexico-grammatical description "focuses on wordings at the level of the clause and below", whereas the discourse-semantic approach adopted by the Appraisal framework "operates at a higher level of abstraction to model meanings above the clause, in the context of text" (White, 1998, p. 11). According to White (1998), compared to the lexico-grammar approach, the "discourse-semantic orientated descriptions always have the potential to be more specific than general, to be more closely tailored to the communicative concerns of a particular context of situation, or group of related contexts of situation" (p. 73).

Appraisal framework has three subsystems: Attitude, Engagement, and Graduation, and each subsystem has its own subcategories. Figure 1 gives an overall view of the whole framework. In the current study, we focused only on the Attitude subsystem.

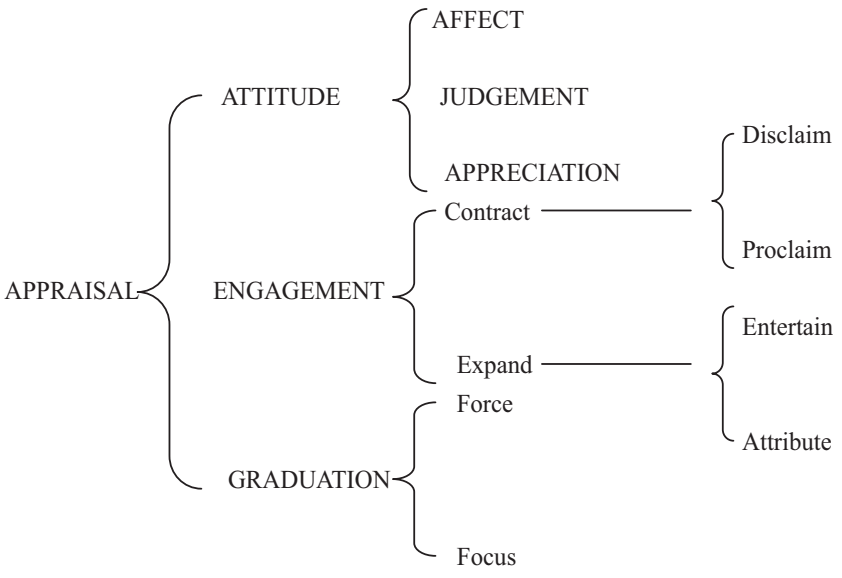



**Fig. 1** Overview of appraisal framework. (Adapted from Martin & White, 2005, p. 38)

Attitude is probably the most studied subsystem within Appraisal framework and is the umbrella term for evaluative language in attitudinal positioning in texts. It consists of three sub-categories: Affect, Judgement and Appreciation. Affect is concerned with people's emotional responses and feelings, such as *happy* or *frightened*. One useful dimension of classifying affectual values is concerned with whether, or not, the writer takes responsibility for the attitudinal value assessment, namely the Authorial-affect (1st-person) vs. non-Authorial-affect (2nd & 3rd person) (Martin & White, 2005). Authorial-affect means the writer is the source of emotional responses and takes responsibility for the evaluation conveyed. Example 1 below shows by using the first-person pronoun, the student writer demonstrated in an explicit manner his or her strong satisfaction with the peer's suggestions. The rhetorical effect is that through foregrounding their subjective presence in the communicative process, a writer establishes an interpersonal rapport with the reader and invites the reader's evaluation of their emotional responses (White, 2005).

**Example 1**
I like your word choices!!!

Judgement involves attitudinal evaluation of human beings and their behaviors by reference to social norms, customs, or rules, such as *capable, honest*. For writers, Judgement resources provide one of the most explicit means to inscribe their positions in texts, and they are concerned with esteem-oriented assessment, as well as moral and legal evaluations. Example 2 below illustrates that the student reviewer inscribed a positive judgement of the writer's capacity in English writing.

**Example 2**
Other than that, and a few spelling mistakes, well done.

Appreciation subsystem consists of resources for aesthetic evaluation of things and entities, such as *complex* or *important*. Different from Judgement system, which focuses on human being and their behaviors, Appreciation primarily assesses the nature, value, and composition of inanimate entities. Example 3 below shows how the student reviewer thought about the tone in the given essay. We can see although the topic was regarded as interesting, the student reviewer assessed the level of formality as too high.

**Example 3**
It's interesting [+appreciation], I thought yours was rather too formal [−appreciation].

Affect, Judgement and Appreciation constitute an interconnected and interactive system of evaluation. They are all motivated by affectual response with Judgement institutionalizing affectual positioning with respect to human behaviors and Appreciation institutionalizing affectual positioning with respect to product and process. Another feature of the Attitude system is that Attitude values can be positive or negative. For example, to be *happy* is a kind of positive affect while to be *afraid* refers to the feeling of insecurity, thus coded as negative. In Appraisal coding practice, positive and negative values are normally indicated by "+" and "−" respectively.

### *2.3 Related Studies on Peer Feedback Strategies*

Recent years have witnessed a growing interest in studying students' strategies for giving and receiving feedback orally or in written forms (e.g., Nguyen, 2008a, b; Patchan et al., 2016; Yu & Lee, 2015, 2016). It has been found that student reviewers employed linguistic, social, and cognitive strategies during the process of providing feedback, which attracted different reader responses and led to different revision behaviors. For example, drawing on activity theory, Yu and Lee (2016) explored the peer feedback strategies adopted by four Chinese EFL learners in group peer feedback activities. Data sources involved video recordings of peer feedback sessions, semi-structured interviews, stimulated recalls, and drafts of student L2 writing. It was found that the four students employed five major strategies in peer feedback, namely using L1 Mandarin Chinese, employing L2 writing criteria, adopting rules of group activity, seeking help from teachers, and playing different roles. In another case study of two Chinese EFL university students' peer feedback practices and the factors influencing their feedback practices, Yu and Lee (2015) found that, although one participant in the study primarily focused on form and surface-level issues (like grammar errors), another participant endeavored to balance his feedback coverage and paid more attention to fluency of writing, content and idea development, and vocabulary use. Overall, the two participants in the study both provided specific and negative and critical feedback. This finding is different from conventional belief (e.g., Carson & Nelson, 1994, 1996) that due to the negative influencer of Chinese cultural traditions, Chinese students prefer to value practices such as face protection and maintenance of group harmony. Yu and Lee (2015, 2016) argued that, apart from explaining Chinese students' peer feedback processes, beliefs and behaviors from a cultural tradition perspective, it is equally important to situate students' feedback practices within individual students' unique learning experiences.

Patchan et al. (2016) have attempted to examine how features of peer feedback influence students' likelihood of implementation and how implementable comments enhance the writers' ability to revise. They obtained written samples and reviewer comments from 432 students from an Introduction to Psychological Science course in an American university. From the written comments, they identified two general types of feedback features, namely cognitive and affectual. In the study, cognitive feedback features included summarization, feedback specificity, explanations and scope, and affectual features referred to affective language, particularly the use of praises and negative comments. The study found that 64% of all comments included some form of praise and 52% of all comments included only praise. Furthermore, 12% of all comments contained praises aiming to mitigate a criticism comment. It was also found that praise comments affected students' likelihood of implementation, but not the quality of their revisions. Another interesting finding from the study is that students were 10% less likely to implement the comments if they received mitigating praises. In other words, although praise overall may have a motivating effect on a writer, praise being used to mitigate or soften the tone of a critique may have an unintended effect of allowing the writer to overlook the problem.

Nguyen (2008a, b) studied pragmatic failure phenomenon in Vietnamese English learners' criticisms in peer-feedback. Thirty-six Vietnamese English learners completed a 250-word argumentative essay and were then arranged into dyads for peer-feedback conversations. It was found that Vietnamese English learners varied in their choice of criticism strategies and formulae compared to the Australian native English speakers. For example, Vietnamese English learners tended to be less direct than native English speakers with less use of direct criticisms; instead, they tended to use more offensive indirect criticisms such as demands, like "you must." The students were more likely to announce their peers' problems rather than just describe them. When suggesting solutions, their suggestions sounded more imposing and assertive than native English speakers. The Vietnamese students tended to use less modifiers to reduce the potential face-threatening effects of their criticisms.

In short, the study on students' peer feedback strategies has been attracting increasing attention. However, previous studies mainly examined students' practices in giving feedback while few studies looked at how students responded to feedback received. Secondly, few studies have investigated student reviewers' strategy use in giving and responding to feedback from an evaluation perspective. Hence, the purpose of this study is to investigate how Chinese and American students used evaluative language in providing and responding to their peers' written feedback by applying the Appraisal framework. The specific research questions of the current study are as follows:

1. What sub-categories of the Attitude subsystem of the Appraisal framework are present in Chinese university students' written feedback to American university students' English writing, and in Chinese university students' written responses to American university students' written feedback?
2. What sub-categories of the Attitude subsystem of the Appraisal framework are present in American university students' written feedback to Chinese university students' English writing, and in American university students' written responses to Chinese university students' written feedback?
3. What are the similar or different patterns in the use of sub-categories of the Attitude subsystem of the Appraisal framework in Chinese and American university students' provision of written feedback and their responses to peers' written feedback?

## 3 Methodology

### *3.1 Context of the Study*

The study was IRB-approved and conducted at an American research university and a prestigious university in a coastal metropolis in China. The university in China has approximately 42,000 students, of which around 2500 are international students. The American university has approximately 56,000 students, with over 60% of its undergraduates being Latino/Hispanic (National Center for Education Statistics, n.d.).

For the current study, one English reading and writing course in China was paired with one writing-as-processes course in the American university. In the 10 weeks of the research, 21 out of 24 students from the American university, many of whom were ESL writers, and 20 participants from the university in China, all of whom were EFL writers, read the same article and wrote with the same essay prompt. The article was published in New Yorker, that is "Live and Learn: Why We Have College", providing a conceptual framework to understand the nature of higher education. After reading this paper, students wrote an essay on the topic, "Why I Attend College."

Students were assigned with one overseas partner (or two partners for three pairs) for peer feedback and then revised their essay based on their decisions about the feedback received; they submitted their final draft for grading.

## *3.2 Stages of Peer Feedback*

There were three stages from the first draft to the final draft in the current study. During writing processes, written peer feedback, occurring from the second to the third stage, acted as an important measure to improve student writers' reader awareness and knowledge about writing. In the first stage, students completed their first draft of the essays and received a brief peer feedback training session. The training introduced benefits of giving and taking feedback, procedures of peer feedback, and the proper language use in written feedback. During the second stage, the Chinese students were randomly paired with their American counterparts. Within each pair, participating students were asked to read and comment on the overseas partner' drafts with the help of a feedback rubric. The rubric consists of eight questions and revolves around issues like essay structure (i.e., paragraphing, topic sentence, thesis statement), idea development, evidence and supports, grammar, and vocabulary. Students were encouraged to provide as much feedback as possible related to those categories. Students were also asked to annotate within the essay their comments by using the Comment and Track Changes function of MS Word.

During the third stage, students responded to the comments provided by partners from the other country. They were asked to evaluate how they would incorporate their partner's comments into the revision process. After this round of re-feedback, students had to revise their writing considering the peer feedback activities and submit the final version to instructor for scoring.

## *3.3 Corpora of Written Peer Feedback*

As mentioned above, all participating students first provided initial feedback to their counterparts' English writing and then responded to the feedback they received. Therefore, by the end of this project, four types of students' written feedback were collected, namely Chinese and American university students' initial feedback on their peers' English writing, and their responses to feedback they have received. While

**Table 1** General information of the four corpora

| Corpus name | Sources of comments | Size |
|---|---|---|
| C-A-1 | Chinese university students' feedback on American university students' writing | 9194 words |
| A-C-1 | American university students' feedback on Chinese university students' writing | 10,779 words |
| C-A-2 | Chinese university students' responses to feedback given by American university students | 5476 words |
| A-C-2 | American university students' responses to feedback given by Chinese university students | 5068 words |

compiling the corpora, we noticed that students included in their comments many direct quotations from their peers' essays. To have these corpora represent only students' own use of language, we decided to delete direct quotations students included from their peers' essays. In the end, we have four corpora amounting to 30, 517 words in total. Table 1 gives an overview of the nature of the four corpora and their size respectively.

### *3.4 Appraisal Coding*

To conduct a valid coding of Appraisal items, a combination of bottom-up and top-down approach was adopted in the current study (Martin & White, 2005). By taking a bottom-up approach, coders could focus on the lexical and grammatical expressions of appraisal values. However, one danger of this bottom-up approach of coding Appraisal values is over-coding because, at times, too much focus on specific linguistic entities can run the risk of the coder ignoring their surrounding context and thus coding more than necessary. At this time, a top-down approach can help minimize over-coding because this approach takes a top-down reading of the coding work, namely a "prosodic perspective" (Martin & White, 2005, p. 85) and can detect and iron out inappropriate coding. In the present study, a top-down reading of the coding work was performed once a bottom-up coding of the essay was completed.

To establish coder reliability, the first author and a research assistant independently coded Appraisal items in the four corpora. In total, we achieved an 81% point by point agreement for all coding. Differences in coding were resolved through discussions and consensus.

## 4 Results and Discussion

In this section, we report on the results of the analysis of Attitude coding on two dimensions. The first dimension was to see how Chinese and American students employed Attitude items in their initial provision of feedback and their responses to

feedback received. Within this dimension, we examined how Chinese and American students used Attitude items in their written feedback from a within group angle, and then compared the two groups' Attitude uses from a cross-group angle. The second dimension examined how Chinese and American students used Attitude items across the initial giving feedback to their responses to feedback received.

As the four corpora are of different size in the following analysis, all occurrences of Attitude items were normalized to an article with 1000 words to make the data comparable. To compare the mean differences in the use of Attitude items, the Mann-Whitney test, a non-parameter equivalent of the *t*-test, was employed. Statistic package SPSS was used to run the Mann-Whitney tests.

## *4.1 Attitude Items in the Initial Provision of Feedback*

Table 2 includes the frequency of Attitude items in Chinese and American university students' initial written feedback to their peers' essays. Overall, American students employed 822 instances of Attitude items and Chinese students used 749 items, but there was no significant difference ($U$ (24,24) = 248, $p > .05$).

Chinese and American students exhibited a similar pattern of Attitude resources. That is, there were predominantly more Appreciation items in both C-A-1 (72.0%) and A-C-1 (72.5%) than Affect and Judgement ones. In fact, in both C-A-1 and A-C-1, Affect and Judgement items only took up similarly small proportions. Specifically, in C-A-1, Affect items took up 12.6% and Judgement 15.5%; in A-C-1, Affect items amounted to 13.7% and Judgement 13.8%. The predominance of Appreciation resources pertains to the very nature of the writing task. In this round of reviewing, students' foci were on the technical aspects of the essay, such as structure, content, and grammar; these are all inanimate entities, hence more Appreciation items. The resultant rhetorical effect of the predominance of Appreciation values is that students' written feedback sound more appreciative than emotional or judgmental (Hood, 2004). In other words, while evaluating peers' essays, both Chinese and American university students seemed to refrain from disclosing too many personal emotions, and they avoided direct ethical or moral evaluations.

**Table 2** Attitude items in C-A-1 and A-C-1 during the initial provision of feedback

| | | Chinese students' written feedback (C-A-1) | | | American students' written feedback (A-C-1) | | |
|---|---|---|---|---|---|---|---|
| | | Mean | SD | % | Mean | SD | % |
| Affect | Positive | 6.07 | 9.36 | 8.9 | 8.55 | 7.10 | 11.3 |
| | Negative | 3.05 | 3.30 | 3.6 | 2.18 | 2.99 | 2.4 |
| Appreciation | Positive | 40.03 | 18.15 | 47.4 | 41.03 | 21.03 | 51.2 |
| | Negative | 19.21 | 13.02 | 24.6 | 15.67 | 9.86 | 21.3 |
| Judgement | Positive | 6.56 | 6.07 | 8.3 | 7.79 | 6.64 | 10.2 |
| | Negative | 5.99 | 5.81 | 7.2 | 2.64 | 3.26 | 3.5 |

Regarding the polarity of Attitude evaluations, more than half of Attitude resources were positive in both C-A-1 (64.63%) and A-C-1 (72.75%). Specifically, this prevalence of positive tone was observed across all the subsystems, namely Affect, Appreciation, and Judgement, across the two corpora. This consistent pattern of positive tone suggests that, overall, both Chinese and American students seemed to show satisfaction with and recognition of the value of peers' essays, and a positive assessment of peers' writing practices.

Apart from the above mentioned overall patterns, there are another two noteworthy differences between C-A-1 and A-C-1. First, American students employed significantly more positive Affect items (M = 8.55) than Chinese students (M = 6.07) ($U(24,24) = 160$, $p < .01$). Compared with Chinese students, American students used more Authorial-affect in their feedback, which indicates that the writer takes responsibility for the attitudinal value assessment, functioning to pose a strong personal voice or evaluative stance. This finding seems to accord with what traditional CR research has disclosed about Chinese students' unwillingness to explicitly reveal their personal affect in writing (Carson & Nelson, 1994, 1996; Shen, 1989). Linguistically, as Examples 4 and 5 show, American students were more likely to explicitly encode their positive affect of satisfaction towards Chinese students' writing by the structure of "I like/love."

**Example 4**
This is a strong example, I like [+affect] it. (C1B, A-C-1).

**Example 5**
I really like [+affect] the use of this word and the image it creates. (C3B, A-C-1).

This finding is different from what Liu and McCabe (2017) had found out in their corpora of Chinese and British university students' English writing. In Liu and McCabe's (2017) study, they found that Chinese university students employed significantly more Authorial-affect items in the same argumentative writing than their British counterparts did, and they suggested that the resultant rhetorical effect was to make Chinese students' English writing sound more personal and informal, having the potential effect of negatively assessed by writing instructors. However, in the current study, as peer feedback is interactional in nature, this kind of foregrounded Affect has the potential to project a clear writer identity (Hyland, 2002) and help position the potential audience attitudinally.

Secondly, Chinese students used more negative Judgement items (M = 5.99) than American students (M = 2.64) ($U(24,24) = 175$, $p < .05$). This result supports findings from Nguyen's (2008a, b) study, in which Vietnamese English learners also used more criticisms than Australian writers. In both C-A-1and A-C-1, most of the negative Judgement items were attributed to the negative evaluation of feedback receivers' incompetence in writing. In other words, the appraisees were their peers. For example, in Example 6, the Chinese student identified some misspelling mistakes in the American counterpart's essay, whereas in Example 7, the American student thought the Chinese student did not write the introductory paragraph well.

**Example 6**
There are some issues of word choices. First, the writer misspells the word "illustrate" in the first paragraph [-judgement] (E1, C-A-1).

**Example 7**
I feel like you finished the introductory paragraph too quickly [-judgement], so you didn't [-judgement] really have much of a chance to introduce. (C1A, A-C-1).

However, a careful examination of C-A-1 reveals that, in the Chinese students' corpora, there were still a sizable proportion of the negative Judgement items coming from Chinese students' negative evaluation of their own inability to understand, nor comment on their peers' writing. In other words, the appraisees of these negative Judgments were Chinese students themselves as Examples 8 and 9 below show.

**Example 8**
But maybe I misunderstand the author's idea [-judgement], aren't there two topic sentences in the passage? (E14, C-A-1).

**Example 9**
I don't quite understand this sentence [-judgement], would you mind explaining it? (E10, C-A-1).

These types of statements of uncertainty about writing problems are regarded as one of the strategies for indirect criticisms (Nguyen, 2008a, b), and its purpose is to raise appraisees' awareness of the inappropriateness of their writing practices.

## *4.2 Attitude Items in Students' Responses to Feedback Received*

Table 3 shows the distribution of Attitude items across the three subsystems in Chinese and American students' written responses to the feedback their overseas partners have provided. Numerically speaking, Chinese students employed significantly more Attitude items (M = 137.58) than American students (M = 64.50) ($U(24, 24) = 27$, $p < .01$). More specifically, there were both more positive ($U(24, 24) = 44$, $p < .01$) and negative Attitude items ($U(24, 24) = 192$, $p < .05$) in Chinese students' written responses than in American students' ones.

As to the three subsystems, it is evident from Table 3 that in Chinese students' responses to their American counterparts' comments (C-A-2), Judgement items took up the biggest proportion, amounting to 46.7% while Appreciation and Affect items were 36.3% and 17.0% respectively. In American students' responses to Chinese peers' comments (A-C-2), more than half of the Attitudinal items were encoded as Appreciation, amounting to 57.2%, whereas Judgement and Affect items were 23.0% and 19.8%, respectively. This revealed a different attitude orientation between the two groups of students, namely the Chinese students foregrounded judgements of American students' capacities and their reviewing practices, whereas American students focused more on the appreciation of Chinese students' essays.

**Table 3** Attitudinal items in students' responses to feedback received

| | | Chinese students' written feedback (C-A-2) | | | American students' written feedback (A-C-2) | | |
|---|---|---|---|---|---|---|---|
| | | Mean | SD | % | Mean | SD | % |
| Affect | Positive | 18.49(104) | 11.19 | 14.0 | 9.48(55) | 7.88 | 14.9 |
| | Negative | 3.87(22) | 4.69 | 3.0 | 2.71(18) | 4.20 | 4.9 |
| Appreciation | Positive | 37.39(198) | 19.39 | 26.7 | 27.42(156) | 19.28 | 42.3 |
| | Negative | 10.56(72) | 10.44 | 9.7 | 9.17(55) | 11.57 | 14.9 |
| Judgement | Positive | 57.50(289) | 19.82 | 38.9 | 12.62(69) | 12.97 | 18.7 |
| | Negative | 9.77(58) | 8.75 | 7.8 | 3.11(16) | 5.43 | 4.3 |

Taking a between group perspective, Chinese students used significantly more Judgement items (M = 67.27) than American students (M = 15.72) ($U$(24,24) = 17, $p < .01$). A closer examination of students' written responses showed that a large proportion of Chinese students' Judgement items were encoded by expressing their gratitude to American students' provision of revision suggestions, as Example 10 shows. However, this type of explicit "thank you" evaluation of feedback givers' efforts was rare in American students' feedback.

**Example 10**
Thank you [+judgement] for your suggestions. (C1A, C-A-2).

Another noteworthy difference is that Chinese students employed significantly more positive Affect statements (M = 18.49) than American students (M = 9.48) ($U$(24,24) = 142, $p < .05$). Specifically, Chinese students encoded more affectual inclination items, indicating their readiness to adopt their American peers' advice and revise their essays accordingly. This type of evaluation was mostly realized through the structure of "I'll try to do something" as shown in Examples 11 and 12.

**Example 11**
Thank you. I'll try to [+affect] build more on my introduction and your example is helpful. (C4A, C-A-2).

**Example 12**
I had not found an adequate one when I wrote the essay, but I will try to [+affect] find one if necessary. (C8, C-A-2).

### *4.3 Attitude Items in Chinese Students' Initial Written Feedback and Their Responses to Feedback Received*

Table 4 shows the use of Attitude items in Chinese students' initial provision of feedback and their further responses to feedback received. Numerically speaking, Chinese students used significantly more Attitude values in their written responses to feedback received (C-A-2, Mean = 137.58) than in their initial feedback (C-A-1, Mean = 80.91) ($U$ (24,24) = 27, $p < 0.01$). Specifically, this significant difference between C-A-2 and C-A-1 can be observed across Affect, Judgement and Appreciation.

**Table 4** Attitude items in Chinese students' initial written feedback and their responses to feedback received

| | | Chinese students' written feedback (C-A-1) | | | Chinese students' written feedback (C-A-2) | | |
|---|---|---|---|---|---|---|---|
| | | Mean | SD | % | Mean | SD | % |
| Affect | Positive | 6.07 | 9.36 | 9.0 | 18.49 | 11.19 | 14.0 |
| | Negative | 3.05 | 3.30 | 3.6 | 3.87 | 4.69 | 3.0 |
| Appreciation | Positive | 40.03 | 18.15 | 47.4 | 37.39 | 19.39 | 26.7 |
| | Negative | 19.21 | 13.02 | 24.6 | 10.56 | 10.44 | 9.7 |
| Judgement | Positive | 6.56 | 6.07 | 8.3 | 57.50 | 19.82 | 38.9 |
| | Negative | 5.99 | 5.81 | 7.2 | 9.77 | 8.75 | 7.8 |

As far as the polarity of Attitude values are concerned, in both C-A-1 and C-A-2, there were significantly more positive Attitude items than negative ones. Notably, there were significantly more positive Attitude items in C-A-2 (Mean = 113.39) than those in C-A-1 (Mean = 52.67) ($U(24,24) = 24$, $p < 0.01$). This shows that Chinese university students instilled more positive emotional responses than negative ones in their evaluation of peers' written products and peers' feedback practices. They sounded more emotionally positive and assertive when responding to peer's feedback than when giving feedback.

Across each of the three subsystems, there were three other important differences. First, there were significantly more positive Affect items in C-A-2 (M = 18.49) than in C-A-1 (M = 6.07) ($U(24,24) = 88$, $p < 0.01$). This difference was mainly attributed to the positive encoding of affectual inclinations in C-A-2 through expressing their readiness to take American counterparts' suggestions and revise their essays accordingly as shown in Examples 11 and 12 above.

Second, there were significantly more negative Appreciation items in C-A-1 than in C-A-2 ($U(24,24) = 164$, $p < .01$). This suggests that while giving feedback, Chinese students paid more attention to identify problems in their peers' writing. Third, there were significantly more positive Judgement items in C-A-2 than in C-A-1($U(24,24) = 3$, $p < .01$). This difference is attributed to Chinese students' expression of gratitude to their peers' comments in C-A-2 as shown in Example 10 above. These two differences seem to signify Chinese students' changing perception of the targets for their textual evaluation. In the first round of reviewing, their focus of evaluation was the essay itself, whereas in the second round of reviewing, the artifact for evaluation was a combination of essays and peers' reviewing practices.

## *4.4 Attitude Items in American Students' Initial Written Feedback and Their Responses to Feedback Received*

Table 5 shows the use of Attitude items in American students' initial provision of feedback and their further responses to feedback received. Overall speaking, compared to their Chinese counterparts, American students exhibited less variations in

**Table 5** Attitude items in American students' initial written feedback and their responses to feedback received

| | | American students' written feedback (A-C-1) | | | American students' written feedback (A-C-2) | | |
|---|---|---|---|---|---|---|---|
| | | Mean | SD | % | Mean | SD | % |
| Affect | Positive | 8.55(93) | 7.10 | 11.3 | 9.48(55) | 7.88 | 14.9 |
| | Negative | 2.18(20) | 2.99 | 2.4 | 2.71(18) | 4.20 | 4.9 |
| Appreciation | Positive | 41.03(421) | 21.03 | 51.2 | 27.42(156) | 19.28 | 42.3 |
| | Negative | 15.67(175) | 9.86 | 21.3 | 9.17(55) | 11.57 | 14.9 |
| Judgement | Positive | 7.79(84) | 6.64 | 10.2 | 12.62(69) | 12.97 | 18.7 |
| | Negative | 2.64(29) | 3.26 | 3.5 | 3.11(16) | 5.43 | 4.3 |

their uses of Attitude items in both giving feedback and responding to feedback received.

The only significant difference was observed in American students' use of Appreciation items. Specifically, there were significantly more Appreciation items in A-C-1 (M = 56.70) than in A-C-2 (M = 36.59) ($U(24,24) = 158$, $p < 0.01$). This is easy to understand as in the first round of feedback, students' attention was on the paper writing, the inanimate entities, thus more Appreciation values.

## 5 Conclusions

This study reported some preliminary findings from a larger project on cross-cultural written peer feedback between Chinese and American university students in a reading and writing course. The aim of this chapter was to present numerical patterns of the use of evaluative language in Chinese and American students' written comments from a linguistic perspective. For this purpose, the Appraisal framework (Martin & White, 2005) developed from a Systemic Functional Linguistics perspective has been adopted as the analytical tool. There were some noteworthy patterns for the use of Appraisal items between the two groups of student reviewers at the two different stages of reviewing.

First, while giving feedback, a consistent pattern for the use of Attitude items is that both Chinese and American students exhibited a prevalence of Appreciation items over Affect and Judgement ones. There were two major differences at this stage of reviewing. One was that American students employed significantly more positive Affect items than Chinese students did, and another was that Chinese students used more negative Judgement items than American students did.

Second, while responding to peers' written comments, Chinese and American students demonstrated more variations than similarities. For example, Chinese students foregrounded Judgement values while American students focused more on Appreciation subcategory; Chinese students used significantly more positive Affect items and Judgement items than American students did.

Third, during the process of giving feedback and responding to feedback received, Chinese students exhibited more variations than American students did. On the one hand, the only significant difference observed in American students' written comments across the two stages is that they used more Appreciation items in giving feedback than in responding to feedback received. On the other hand, Chinese students used more negative Appreciation items in giving feedback than in responding to feedback received; they also used significantly more positive Affect and Judgement items in responding to feedback received than in giving feedback.

Future studies can usefully expand this work from two perspectives. First, the current paper mainly focused on numerical description of Attitude items, so a fine-grained textual analysis for the linguistic realization of Attitude items is warranted. Second, this paper only included an examination of the Attitude subsystem. Thus, future studies may examine how students would position and align imagined audience by employing Engagement and Graduation values in their written comments.

## References


Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text, 9*, 93–124.

Carson, J. G., & Nelson, G. L. (1994). Writing groups: Cross-cultural issues. *Journal of Second Language Writing, 3*(1), 17–30.

Carson, J. G., & Nelson, G. L. (1996). Chinese students' perceptions of ESL peer response group interaction. *Journal of Second Language Writing, 5*(1), 1–19.

Connor, U. (2002). New directions in contrastive rhetoric. *TESOL Quarterly, 36*(4), 493–510.

Crismore, A. (1989). *Talking with readers: Meta-discourse as rhetorical act*. Peter Lang.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.).

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd ed.).

Hood, S. (2004). *Appraising research: Taking a stance in academic writing*. Unpublished Ph.D Dissertation. University of Technology, Sydney.

Hunston, S., & Thompson, G. (2000). *Evaluation in text: Authorial stance and the construction of discourse*. Oxford University Press.

Hyland, K. (1999). Talking to students: Metadiscourse in introductory coursebooks. *English for Specific Purposes, 18*(1), 3–26.

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics, 34*(8), 1091–1112.

Hyland, K. (2005). Stance and engagement: A model of interaction in academic discourse. *Discourse Studies, 7*(2), 173–192.

Hyland, K., & Hyland, F. (2006). Feedback on second language students' writing. *Language Teaching, 39*(2), 83–101.

Jin, L., & Cortazzi, M. (2011). *Researching Chinese learners: Skills, perceptions and intercultural adaptations*. Palgrave Macmillan.

Kaplan, R. B. (1966). Cultural thought patterns in inter-cultural education. *Language Learning, 16*(1–2), 1–20.

Liu, J., & Hansen, J. (2002). *Peer response in second language writing classrooms*. University of Michigan Press.

Liu, X., & McCabe, A. (2017). *Attitudinal evaluation in Chinese university students' English writing: A contrastive perspective*. Springer.

Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for Academic Purposes, 2*(3), 193–227. https://doi.org/10.1016/S1475-1585(03)00025-0

Lu, J., & Law, N. (2012). Online peer assessment: Effects of cognitive and affective feedback. *Instructional Science, 40*, 257–275. https://doi.org/10.1007/s11251-011-9177-2

Martin, J. R. (2000). Beyond exchange: APPRAISAL systems in English. In S. Hunston & G. Thompson (Eds.), *Evaluation in text: Authorial stance and the construction of discourse* (pp. 142–177). Oxford University Press.

Martin, J. R., & White, P. R. R. (2005). *Language of evaluation: Appraisal in English*. Palgrave Macmillan.

Min, T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing, 15*, 118–141. https://doi.org/10.1016/j.jslw.2006.01.003

Min, T. (2008). Reviewer stances and writer perceptions in EFL peer review training. *English for Specific Purposes, 27*, 285–305. https://doi.org/10.1016/j.esp.2008.02.002

Nguyen, T. T. M. (2008a). Criticizing in an L2: Pragmatic strategies used by Vietnamese EFL learners. *Intercultural Pragmatics, 5*, 41–66.

Nguyen, T. T. M. (2008b). Modifying L2 criticisms: How learners do it. *Journal of Pragmatics, 40*, 768–791.

Patchan, M. M., Schunn, C. D., & Correnti, R. J. (2016). The nature of feedback: How peer feedback features affect students' implementation rate and quality of revisions. *Journal of Educational Psychology, 108*(8), 1098–1120.

Rahimi, M. (2013). Is training student reviewers worth its while? A study of how training influences the quality of students' feedback and writing. *Language Teaching Research, 17*(1), 67–89. https://doi.org/10.1177/1362168812459151

Shen, F. (1989). The classroom and the wider culture: Identity as a key to learning English composition. *College Composition and Communication, 40*(4), 459–466.

Thompson, G. (2001). Interaction in academic writing: Learning to argue with the reader. *Applied Linguistics, 22*(1), 58–78.

Villamil, O. S., & de Guerrero, M. C. M. (1998). Assessing the impact of peer revision on L2 writing. *Applied Linguistics, 19*(4), 491–514.

White, P. R. R. (1998). *Telling media tales: The news story as rhetoric*. Unpublished Ph.D dissertation. University of Sydney.

Woo, M., Chu, S., & Li, X. (2013). Peer-feedback and revision process in a wiki mediated collaborative writing. *Educational Technology Research and Development, 61*, 279–309. https://doi.org/10.1007/s11423-012-9285-y

Wu, S. Y., & Rubin, D. L. (2000). Evaluating the impact of collectivism and individualism on argumentative writing by Chinese and North American college students. *Research in the Teaching of English, 35*(2), 148–178.

Yang, Y. (2015). Transforming and constructing academic knowledge through online peer feedback in summary writing. *Computer Assisted Language Learning, 29*(4), 683–702. https://doi.org/10.1080/09588221.2015.1016440

Yu, S., & Lee, I. (2015). Understanding EFL students' participation in group peer feedback of L2 writing: A case study from an activity theory perspective. *Language Teaching Research, 19*(5), 572–593.

Yu, S., & Lee, I. (2016). Exploring Chinese students' strategy use in a cooperative peer feedback writing group. *System, 58*, 1–11.

Zhao, H. (2010). Investigating learners' use and understanding of peer and teacher feedback on writing: A comparative study in a Chinese English writing classroom. *Assessing Writing, 15*, 3–17. https://doi.org/10.1016/j.asw.2010.01.002

# Part VII
# Epilogue

# Hitching a Ride with Steve Graham Through the Galaxy of Writing Research


**Tanya Santangelo, Michael Hebert, and Pamela Shanahan Bazis**

**Abstract** Steve Graham's legacy as a writing researcher may be best represented through his systematic reviews and meta-analyses. Through his reviews, he has hitchhiked through the literature and provided the field with a guide to the galaxy of writing research. The purpose of this chapter was to review Graham's systematic reviews and place them within the context of a scoping review of all of the other systematic reviews of writing research. To do this, we conducted two separate reviews. First, we identified all of the reviews of writing research on Steve Graham's curriculum vitae and grouped and reviewed them by type. Second, we conducted a scoping review of the writing research. We identified a total of 317 systematic reviews of the writing literature, of which 40 (13%) involved Steve Graham as an author. Of the 277 reviews that did not include Steve Graham as an author, 117 (42%) cited Graham at-least once. Graham's reviews have spanned 44 years (and counting) and included more than 2900 studies. We classified Graham's reviews into three primary types: (1) instructional effectiveness, (2) group comparisons, and (3) general non-systematic reviews. Some of the major findings of Graham work include (a) the identification of more than 30 effective general practices for improving writing outcomes for students, (b) writing and writing instruction improve reading and content learning, (c) students who are at-risk for learning difficulties due to reading, language, or ADHD also tend to perform lower than their typically developing peers in writing, and (d) some of his general reviews of the literature show a more complete picture of the writing literature in a particular area. The discussion includes ideas for how Graham's work might be used to inform future writing research.

T. Santangelo
Arcadia University, Glenside, PA, USA

M. Hebert (✉)
School of Education, University of California, Irvine, Irvine, CA, USA
e-mail: mhebert1@uci.edu

P. S. Bazis
University of Nebraska–Lincoln, Lincoln, NE, USA



X. Liu et al. (eds.), *The Hitchhiker's Guide to Writing Research*, Literacy Studies 25, https://doi.org/10.1007/978-3-031-36472-3_21

**Keywords** Writing instruction · Writing skills · Systematic review · Meta-analysis · Galaxy

## 1 Introduction: Sticking Our Thumb Out

Steve Graham has been one of the most prolific writing researchers of our time. He has published more than 450 manuscripts, including journal articles, book chapters, and white papers. His work covers an incredible array of topics around writing in education, including studies of the effectiveness of writing instructional approaches (e.g., Graham et al., 2005), books on best practices of writing instruction (e.g., Graham et al., 2019), examinations of writing assessment (e.g., Graham, 1987), surveys of writing teachers (e.g., Graham et al., 2014), correlational studies examining the relationships between writing and other variables (e.g., Kim & Graham, 2022), and the development of theoretical models of writing (e.g., Graham, 2018). Alongside his wife Dr. Karen Harris, Graham also helped to develop Self-Regulated Strategy Development for writing, which has been one of the most widely-researched and effective writing approaches for improving writing outcomes for students, especially those with disabilities and writing difficulties. That said, some of Steve Graham's most important contributions are arguably his systematic reviews and meta-analyses of the writing literature.

In his paper, *Through the Looking Glass: Reflections of a Writing Scholar*, Graham (2021) stated that he has found conducting systematic reviews of the literature useful for three reasons: (1) they have provided readers with a roadmap[1] of what science tells us about writing, (2) they have broadened his knowledge about writing, and (3) they have helped him improved his own research through examining the strengths and weaknesses in the work of the field. Given Steve's impacts on the field of writing, the fact that two of his reasons focused on how his reviews have influenced his own work speaks volumes about the importance of them. For the purposes of this chapter, however, we're more interested in his first reason; if one systematic review is a roadmap, Steve's collection of reviews and meta-analyses are an atlas that we can use to navigate our way through the galaxy of writing research.

### *1.1 A Vast and Unexplored Galaxy?*

Although we describe the field of writing research as a "galaxy," writing is often considered a neglected subject area. In 2003, the National Commission on Writing in America's Schools and Colleges dubbed writing the Neglected "*R*" and pointed

[1] Steve conveniently used this road map metaphor within a paper in which he also tells his tales of hitchhiking.

out that it had been left out of the school reform movement. Not much has changed since then. The National Commission on Writing no longer exists, and writing has generally been left out of the national conversation around assessing writing in grades K-12 in the United States. As of this writing, the most recent writing assessment conducted by the National Assessment of Education Progress (NAEP) was in 2011 (although a pilot test of digital writing was conducted in 2017), and the next National writing assessment is not scheduled until 2030 (U.S. Department of Education, 2022). In comparison, the NAEP assessments of students' reading and mathematics skills occur every two years. The lack of attention on writing assessment is alarming.

In 2011a, Graham et al. suggested that states either (a) throw high-stakes writing assessments out or (b) make them better. At least one of the authors of that paper (i.e., Hebert) hoped states would choose the latter. Although researchers such as Hillocks (2002), have presented some compelling arguments opposing state writing assessment due to how they can control learning in unintended ways, it is important to consider that *not assessing* student writing might lead to a lack of control altogether. As it stands, many U.S. states have historically assessed writing in only a few grade levels, such as grade 4, grade 8, and grade 12. It isn't a stretch to hypothesize that this may be at-least part of the reason many teachers report not teaching writing, or report teaching writing only sporadically across many grade levels (Goodrich et al., 2022; Graham et al., 2014; Hebert & Savaiano, 2020; Namkung et al. 2022; Powell et al., 2021).

## *1.2 Hitching a Ride on Steve Graham's Rocketship*

Despite the relative neglect of writing in schools, and the relative lack of writing research compared to reading research (Graham & Harris, 2019; Newell et al., 2014; Tate & Warschauer, 2022), an examination of the Steve Graham's reviews and meta-analyses show that we actually know quite a lot about how to teach and assess writing effectively. In fact, Graham has published more than 47 reviews and meta-analyses, many of which identify effective writing practices. He writes them so fast, we sometimes think he must have 1.6 million pounds of rocket fuel[2] propelling him. In fact, Steve recommends always having one or more systematic reviews in progress (see Graham, 2021), and he has influenced his students and collaborators to conduct their own meta-analyses and systematic reviews involving writing research (e.g., Cooper et al., 2022; Hebert et al., 2016; Powell et al., 2017). The point is, although writing may be the neglected R in practice, there has been a robust effort to study writing in education in the research community. As such, it is important to

[2]This was the amount of rocket fuel needed to send NASA's Space Shuttle into orbit (NASA, 2001). It wouldn't surprise us if Steve used an equivalent amount of unsweetened iced tea and lemons for his meta-analyses.

take stock of these reviews so that we can synthesize what we know about writing in schools and develop new research questions to push the field forward.

Based on the nature of this book as a collection of writings to honor Steve Graham, and the limited space we available for this chapter, we focus our attention primarily on Graham's contributions to the writing *review* literature. However, we have also attempted to contextualize his contributions within the broader scope of systematic reviews and meta-analyses of writing conducted in the field of education. Specifically, we conducted a scoping review of all of the systematic reviews and meta-analyses that have been conducted on writing in the field. We also identified that Steve's review research seemed to fall into three broad categories: (1) systematic reviews of writing instructional/intervention research, (2) systematic reviews of population differences on writing outcomes, and (3) broad, non-systematic reviews of topics in the writing literature.

## Reviews of Writing Instruction/Intervention

Many reviews of the writing literature have been conducted to identify effective ways of improving writing outcomes for students. Meta-analysis is a special systematic review tool often used for this purpose, as it can be used to statistically average the effect of an instructional practice across multiple experiments. The use of meta-analysis has grown rapidly in its use for examining the effectiveness of educational approaches across the years; from a single meta-analysis in the field of education published in 1980 to 151 meta-analyses published in 2021 (based on a search of education related meta-analyses in Web of Science, 2022). Meta-analyses of writing have been published over nearly as long of a period, beginning in 1984 with the first meta-analysis of writing instruction published by George Hillocks (i.e., Hillocks, 1984).

Steve Graham has conducted a multitude of these systematic reviews and meta-analyses designed to examine he impacts of instruction on writing outcomes. Meta-analyses of writing instruction may include many different types of instructional practices, including multi-component writing programs, writing interventions focused on teaching students how to write in a specific genre (e.g., narrative or argument writing), or approaches to improving writing (e.g., strategy development). They may also include narrow instructional elements, such as using models or providing explicit handwriting instruction, which may be included within an instructional approach but are less comprehensive. For the purposes of this paper, we classify any study or review aimed at examining the effectiveness of some aspect of instruction or intervention for improving student writing into this category.

In addition to examining instructional impacts on writing outcomes, researchers have conducted reviews to examine whether writing has impacts on other learning outcomes. For example, Bangert-Drowns et al. (2004) examined the impacts of writing on learning across content areas such as science, social studies, and mathematics. Steve Graham has contributed reviews to this set of literature, as well. Therefore, we categorize and examine a subset of Graham's reviews examining the impacts of writing or writing instruction on other learning outcomes.

### Reviews Comparing Populations

Although most used for examining instructional effectiveness, systematic reviews and meta-analyses can also be used to statistically compare non-experimental groups on a range of outcome measures. For example, one might compare the writing performance of boys with the writing performance of girls. As a special education researcher, Graham has conducted several reviews to compare the writing of special populations of students with typically developing peers. Therefore, we identified reviews Steve conducted to examine comparisons between subpopulations of school-aged writers and summarize his findings.

### Broad Literature Reviews on a Topic or Population

Finally, some reviews of the writing literature are broader examinations of a specific aspects of writing (e.g., handwriting) or examinations of practices used across a broad population of learners (e.g., elementary or secondary grades). Steve Graham has also published several of these general reviews. In some cases, Graham has provided models of writing development or composition. In others, his purpose was to make recommendations for practice. Regardless, the purpose of these reviews appears to be to provide a broader understanding of a subset of writing literature. Graham also conducted several such reviews prior to the widespread use of systematic review and meta-analytic methods. Therefore, we examine these within a set of broad reviews of the literature.

## 2 Purpose of the Current Review: Mapping the Galaxy of Steve Graham's Contributions

The purpose of this study is to honor and contextualize Steve Graham's contributions to writing research through his considerable commitment to conducting systematic reviews and meta-analyses of the educational writing literature. These contributions provide overviews and summaries of the research literature, generate new knowledge about writing through syntheses and analyses of findings across related studies, and help to identify potential gaps in the literature or promising areas for new writing research questions and inquiry. To provide context for Graham's contributions, we conducted a scoping review of the educational writing research literature to identify the corpus of systematic reviews and meta-analyses. We provide an overview of Graham's contributions by identifying and summarizing his systematic reviews and meta-analyses, and then situating them within the context of the larger corpus of systematic reviews we identified.

## 3 Method: Blasting Off

Our method included two related searches. One was a search for reviews and meta-analyses conducted by Steve Graham. The second was a scoping review of the literature for all reviews and meta-analyses to put Steve Graham's contributions into the context of the larger educational writing literature.

### *3.1 Mapping the Galaxy of Steve's Reviews*

For Graham's work, we obtained the latest copy of his curriculum vitae (CV) from his webpage. Due to high overlap between Graham's published manuscripts and conference presentations, we limited the studies included to reviews published in journals, in book chapters, or as white papers. When identifying studies from Graham's CV, we searched for titles containing the words *review*, *synthesis*, *meta-analysis*, and *overview*. We then obtained all potential reviews and meta-analyses.

### *3.2 Scoping Review: The Constellation of Other Reviews of Writing*

To identify writing reviews and meta-analyses conducted by the field, we included a systematic and comprehensive search of the literature involving databases, online resources, hand searches of selected sources, and ancestral searches. First, we used EBSCOhost to conduct a multi-database search of *Academic Search Ultimate*, *ERIC*, *OminFile Full Text Mega (H.W. Wilson)*, *APA PsychArticles*, and *APA PsychInfo* with two sets of Boolean search terms:

- Set 1 (9782 unique results): ((review of literature OR literature review OR meta-* OR systematic review OR synthesis) AND (written OR write OR writing OR compose OR composing))
- Set 2 (1377 unique results): ((review of literature OR literature review OR meta-* OR systematic review OR synthesis) AND (spell* OR keyboard* OR handwrit* OR grammar OR vocabulary))

Second, we searched websites developed by organizations committed to collecting and sharing research on evidence-based practices, including: (1) *Best Evidence Encyclopedia*, (2) *Campbell Collaboration*, (3) *Carnegie Corporation*, (4) *Center on Instruction*, (5) *EPPI-Center,* (6) *Evidence for ESSA*, and (7) *What Works Clearinghouse*. All studies related to writing or literacy were identified for further review.

Third, we conducted hand-searches of three education-focused, synthesis-oriented journals:

- *Educational Psychology Review* [1989–2020, 1(*1*)–32(*4*)];
- *Educational Research Review* [2006–2021, 1(*1*)–32]; and
- *Review of Educational Research* [1979–2021, 49(*1*)–91(*1*)]

Fourth, we conducted hand-searches of 18 handbooks that might include writing research, including: *Best Practices in Writing Instruction* (Graham et al., 2019); *Cambridge Handbook of Literacy* (Olson & Torrance, 2009); *Handbook of Family Literacy* (Wasik, 2012); *Handbook of Language and Literacy* (Stone et al., 2016); *Handbook of Orthography and Literacy* (Joshi & Aaron, 2005); *Handbook of Reading Disability Research* (McGill-Franzen & Allington, 2011); *Handbook of Reading Research* (Birr Moje et al., 2020); *Handbook of Research in Second Language Teaching and Learning* (Hinkel, 2017); *Handbook of Writing Research* (MacArthur et al., 2016); *International Handbook of Literacy and Technology* (McKenna et al., 2006); *Routledge Handbook of English Language Teaching* (Hall, 2016); *Routledge Handbook of Language Learning and Technology* (Farr & Murray, 2016); *Routledge Handbook of Literacy Studies* (Rowsell & Pahl, 2015); *Routledge Handbook of Second Language Acquisition* (Gass & Mackey, 2012); *Routledge International Handbook on English, Language, and Literacy Teaching* (Wyse et al., 2010); *SAGE Handbook of Dyslexia* (Reid et al., 2008); and *SAGE Handbook of Early Childhood Literacy* (Marsh & Larson, 2012).

Fifth, we manually screened the reference lists for all of the studies identified from other search methods that met our inclusion criteria. We evaluated references for inclusion/exclusion criteria and obtained any references that could not be excluded based on the title, if they were not already identified through other searches.

### *3.3 Inclusion Criteria*

Following the searches, we identified relevant reviews using the following inclusion/exclusion criteria:

1. The review involved systematic syntheses or reviews of extant, primary research. To be considered systematic, the studies needed to include at least a basic description of critical methods, such as search procedures for locating studies, inclusion/exclusion criteria, or analytic approaches. (Note: We made exceptions for Graham's general reviews of the literature to understand these contributions and due to their comprehensiveness.)
2. The review included results for primary research that included:
   (a) writing instruction and/or
   (b) writing measures.
3. The review involved research conducted for education-related purpose(s) and/or in education-related setting(s).

4. The review included analyses involving students from pre-kindergarten through higher education and/or their writing teachers/instructors.
5. The review was published in 1979 or later. We used 1979 as a benchmark because it is when Steve Graham published his first literature review, and it also seemed to be the start of documentation of systematic search procedures for systematic searches of writing research (e.g., Hillocks, 1984). Without such documentation it was not possible to determine whether the review was systematic or targeted.
6. The review was published in a location accessible to the authors, including:
   (a) peer-reviewed journals;
   (b) prominent, publicly-accessible organizational websites; or
   (c) academic, peer-reviewed/edited handbooks.
7. The review was written in English.

## *3.4 Coding Studies*

### Steve Graham's Reviews

We had similar—but separate—coding procedures for Steve Graham's reviews and other reviews. We first coded classified Steve's reviews by whether they were (1) reviews of the effectiveness of educational practices, (2) comparisons of subpopulations on their writing performance, or (3) general reviews of the literature for a specific writing skill or population. We further classified effectiveness reviews into two subcategories: (1a) those examining the impacts of practices on writing outcomes, and (1b) those examining the impacts of practices or writing on other academic outcomes (e.g., reading, learning).

Next, we coded parameters of the review, including: type of review (i.e., systematic review, meta-analysis, systematic review with a meta-analysis, and general literature reviews); types of studies included in the review (e.g., experiments, quasi-experiments, single-case design, qualitative investigations); publication type (i.e., journal, book chapter, white paper); grade level of included students; type of included students (e.g., full-range of writing achievement levels, specific subpopulations of students with disabilities); and number of studies included in the review. We then coded the type of instruction, comparison, or purpose for the review, as well as the primary outcomes examined. Finally, we coded for the effect sizes, general findings of the review, or recommendations, depending on the type of review.

### Other Reviews

Because this scoping review was primarily concerned with summarizing Steve Graham's contributions to the review literature for writing, we limited the coding of other reviews beyond the inclusion/exclusion criteria. To examine Graham's

influence and contributions to these reviews of educational research on writing, we coded each review for:

- The total number of unique citations of Steve Graham's work.
- The number of citations of one of Graham's reviews or meta-analyses.
- Authors who were previously Graham's students. We cross-referenced the author list from each review with Graham's curriculum vitae doctoral advisees, post-doctoral advisees, and Master's student advisees. We also counted authors as one of Graham's students if he served on their dissertation or thesis committee, to account for those types of mentorship roles. Although we can't know whether Graham's former students conducted reviews based on his influence, we know that a certain percentage of them did (including the first two authors of this scoping review).
- Authors who have collaborated with Graham in some way, but were not former students. It is difficult (or impossible) to know whether Graham may have influenced his collaborators to conduct their own reviews, whether his collaborators were influenced by his work when they conducted their reviews, or whether their collaboration began before or after they conducted their review. However, it is at-least possible that he may have influenced their work in some way. When cross referencing Graham's CV to identify collaborators, we included co-authors of manuscripts or grants, co-presenters, key personnel on grants, editors of books Graham wrote chapters for, and authors of books Graham edited. We limited colleagues to individuals named on Graham's CV.

### *3.5 Analyses*

Following coding, we descriptively summarized the findings of Graham's reviews and meta-analyses. We describe the findings within each category of review, identify potential gaps, and provided ranges of effects. We also use the number of studies in each review and effect sizes to provide recommendations about the relative strengths of the findings. Finally, we identify potential gaps in the literature and discuss areas where Graham's reviews may need updating.

Next, we examined Graham's corpus of reviews and meta-analyses in the context of all of the meta-analyses and writing reviews we identified in our scoping review. We compared Graham's productivity to the total number of reviews, and examine his potential influence on the larger body of writing review literature.

## 4 Results: Galaxy Defining Discoveries

In total, we identified 40 reviews of the educational writing literature conducted by Steve Graham with various colleagues. As we examined the reviews, we were able to categorize them into three primary types of reviews: (1) instructional

effectiveness, (2) group comparisons, and (3) general. Instructional effectiveness reviews examined the effects of some type of instruction or intervention, which we further classified into two subcategories (i.e., impacts of practices on writing outcomes, and impacts of practices or writing on other outcomes). Group comparison reviews examined differences in writing outcomes across two or more populations, modes of writing, and/or writing assessment procedures. General reviews were defined as comprehensive, but non-systematic reviews that were conducted by Graham and colleagues to provide an overview of an area of writing or research.

Graham's reviews have spanned 44 years (and counting) and included more than 2900 studies (although there is likely considerable overlap among the studies across reviews, and we counted all of the references in the general reviews). Thirty of the reviews are in peer-reviewed journals, five are book chapters, and five are white papers. Twenty-eight of the reviews included a meta-analysis, and seven reviews included examinations of qualitative research. A large majority of the reviews (90%) included studies involving multiple levels of schooling (e.g., elementary, middle, and high schools) and 27 (68%) involved studies that included students across the full-range of achievement levels. Nineteen reviews included examinations of studies specifically including analyses of students with disabilities. Table 1 includes a summary of the reviews by type.

## *4.1 Reviews of Instructional or Intervention Effectiveness*

Of the 40 reviews Graham conducted, reviews of instructional effectiveness made up the largest group. These include 30 reviews of the effectiveness of interventions/instruction aimed at improving academic outcomes for students in pre-kindergarten through college. Graham's reviews of instructional effectiveness included more than 1900 studies, combined. Twenty-four (80%) of the effectiveness reviews included meta-analyses. Twenty-five (83%) had inclusion criteria that spanned the full-range of achievement levels, whereas 11 (37%) examined the effectiveness of instruction/interventions for students with disabilities. Two (7%) included analyses for English learners.

**Impacts of Educational Practices on Writing Outcomes** Graham conducted 24 reviews examining impacts of educational practices on writing outcomes (see Table 2). Of those, 19 examined writing instructional or intervention practices (with one a review of other meta-analyses), two examined the impacts of formative assessment and feedback, one examined the impacts of reading on writing, one examined the impacts of balancing reading and writing instruction on writing, and one identified common themes from qualitative studies of effective writing teachers. The average number of studies included in the reviews was 79.3, with a range of 12–151 studies included.

**Table 1** Summary of Graham's reviews of the writing literature by type

| | Instructional effectiveness | | | |
|---|---|---|---|---|
| | Writing outcomes | Other outcomes | Group comparisons | General |
| Total | 24 | 6 | 6 | 6 |
| Included studies (range) | 1538 (12–151) | 363 (19–95) | 298 (17–111) | 727 (1—208)[a] |
| Type of publication | | | | |
| Journal article | 15 | 5 | 6 | 6 |
| Book chapter | 5 | | 0 | 0 |
| White paper | 4 | 1 | 0 | 0 |
| Type of review | | | | |
| Meta-analysis | 14 | 6 | 6 | 0 |
| Systematic review and meta-analysis | 4 | 0 | 0 | 0 |
| Systematic review | 2 | 0 | 0 | 1 |
| Meta-analysis & meta-synthesis | 1 | 0 | 0 | 0 |
| General | 0 | 0 | 0 | 6 |
| Included designs | | | | |
| Quantitative | 22 | 6 | 6 | |
| (Quasi-)Experimental | 22 | 6 | 1 | 6 |
| Regression discontinuity | 1 | 0 | 0 | 0 |
| Single-case | 10 | 0 | 0 | 6 |
| Qualitative | 1 | 0 | 0 | 6 |
| Included students | | | | |
| Full-range of achievement | 19 | 6 | 0 | 4 |
| Diagnosed disabilities | 9 | 2 | 4 | 4 |
| English learners | 2 | 0 | 0 | 0 |
| Average achievement | 3 | 0 | 0 | 0 |
| Above average achievement | 4 | 0 | 0 | 0 |
| Below average achievement | 5 | 0 | 0 | 0 |
| Included grade levels | | | | |
| Elementary (PK-5) | 23 | 6 | 6 | 6 |
| Middle (6–8) | 22 | 6 | 6 | 5 |
| Secondary (9–12) | 19 | 6 | 6 | 5 |
| Post-secondary (13+) | 1 | 0 | 1 | 0 |
| Multiple | 21 | 6 | 6 | 5 |

[a]Some general reviews did not have explicit inclusion criteria, so this number includes all references from the study

**Table 2** Summary of Graham's instructional effectiveness reviews

| References | Pub | Type of review | Included studies | | | | Key findings | ES (PND) range | Cited by |
|---|---|---|---|---|---|---|---|---|---|
| | | | Designs | Grades | Students | *k* | | | |
| Writing outcomes | | | | | | | | | |
| Gillespie & Graham (2014) | J | MA | E, QE | 1–12 | LD | 43 | Identified 4 effective practices: dictation, goal setting, process writing, and strategy instruction. | 0.43–1.09 | 288 |
| Graham (2006) | C | MA | E, QE, SC | 1–12 | A, AA, BA, FR, LD | 39 | Strategy instruction improved writing elements, length, mechanics, quality, and revisions. Effects varied by subgroup. | 0.30–1.32 (87–92) | 776 |
| Graham et al. (2012a) | W | SR | E, QE, SC | K-5 | FR | 34 | Identified effective general practices and made 4 recommendations: (1) create an engaged community of writers; (2) provide daily time for writing; (3) teach students to become fluent with handwriting, spelling, sentence construction, typing, and word processing; and (4) teach students to use the writing process for a variety of purposes. | NA | 375 |
| Graham et al. (2016a) | W | SR | E, QE, RD, SC | 6–12 | FR | 15 | Identified effective general practices and made three recommendations: (1) explicitly teach appropriate writing strategies using a Model-Practice-Reflect instructional cycle, (2) integrate reading and writing to emphasize key writing features, and (3) use assessments of writing to inform instruction and feedback. | NA | 2 |
| Graham and Harris (2003) | C | MA, SR | E, QE, SC | 2–8 | A, AA, ADHD, BA, LD | 18 | SRSD improved writing elements, length, quality, and story grammar. Effects were maintained and varied by subgroup. | 1.47–3.52 | 834 |
| Graham and Harris (2018) | C | SR | E, QE | 1–12 | FR | 437[a] | Identified and categorized effective practices: (1) explicit writing instruction [11 practices], (2) students write [4 practices], (3) supporting students' writing [8 practices], and (4) writing assessment [4 practices]. Some effects varied by subgroup. | −0.17–1.59 | 107 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Graham et al. (2011a) | W | MA, SR | E, QE | 2–12 | FR, LD | 136 | Identified 5 effective assessment practices: minimize presentation effects, monitor progress, provide feedback, teach students to assess their writing, and use preferred mode of writing. | 0.27–0.77 | 333 |
| Graham et al. (2013) | C | MA, SR | E, QE, SC | 1–12 | BA, EBD, FR, LD | 84 | SRSD improved writing elements, length, and quality. Effects were maintained for quality and elements, generalized to different types of writing, and varied by subgroup. | 0.47–2.93 (62–89) | 118 |
| Graham et al. (2015a) | J | MA, MS | E, QE, QL, SC | K-8 | EL, FR, LD | 35 | Identified 13 effective practices aligned with Common Core State Standards: collaborative writing, creating routines, foundational skills instruction, goal setting, prewriting, process writing, sentence combining, strategy instruction, text models, text structure instruction, vocabulary instruction, word processing, and writing to learn. Some effects varied by subgroup. | 0.22–1.00 | 328 |
| Graham et al. (2015b) | J | MA | E, QE | 1–8 | FR | 35 | Feedback from adults, computers, peers, and self improved writing. | 0.38–0.87 | 288 |
| Graham et al. (2018) | J | MA | E, QE | PK-12 | FR | 54 | Reading instruction and increasing students' interactions with text (through reading or observing others read) improved writing. | 0.37–0.63 | 158 |
| Graham et al. (2017b)[b] | J | MA | E, QE | PK-12 | FR | 47 | Balanced reading and writing instruction, with no more than 60% devoted to either, improved writing mechanics, output, and quality. | 0.18–0.69 | 156 |
| Graham et al. (2012b) | J | MA | E, QE | K-5 | A, AA, BA, EL, FR, VD | 115 | Identified 12 effective practices: assessment, comprehensive writing programs, creativity/imagery instruction, extra writing, peer assistance, prewriting, product goals, SRSD, strategy instruction, text structure instruction, transcription skill instruction, and word processing. | 0.42–1.02 | 1084 |

(continued)

**Table 2** (continued)

| References | Pub | Type of review | Included studies | | | | Key findings | ES (PND) range | Cited by |
|---|---|---|---|---|---|---|---|---|---|
| | | | Designs | Grades | Students | *k* | | | |
| Graham and Perin (2007a) | J | MA | E, QE | 4-12 | FR | 123 | Identified 10 effective practices for improving writing: inquiry activities, peer assistance, prewriting, process writing, product goals, sentence combining, strategy instruction, studying models, summarization, and word processing. | 0.25–0.82 | 2319 |
| Graham and Perin (2007b) | J | MA, SR | QL, SC | 4–12 | A, AA, BA, FR, LD | 43 | Identified 6 effective writing instruction practices from SC studies: behavioral reinforcement, direct instruction, grammar, self-monitoring, strategy instruction, and word processing. Identified 10 common themes of effective writing teachers from QL studies (e.g., adapt writing assignments and instruction to meet students' individual needs; dedicate time to writing and writing instruction; model, explain, and provide guided assistance when teaching). | (56–91) | 314 |
| Graham and Perin (2007c) | W | MA | E, QE | 4–12 | FR | 142 | Identified 11 effective writing instruction practices: collaborative writing, inquiry activities, prewriting, process writing, product goals, sentence combining, strategy instruction, studying models, summarization, word processing, and writing for content learning. | 0.23–0.82 | 2373 |
| Graham and Santangelo (2014)[b] | J | MA | E, QE | K-12 | FR | 53 | Teaching spelling improved spelling skills. Effects were maintained and generalized to spelling when writing. | 0.43–0.94 | 385 |
| Graham et al. (2022) | J | MA | E, QE | K-PS | FR | 77 | Identified 4 effective practices for students in Turkey: peer assistance, prewriting, process writing, and strategy instruction. | 0.92–1.55 | 7 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mason and Graham (2008) | J | SR | E, QE, SC | 4-12 | LD | 40 | Provided recommendations for Tier 1 instruction (whole group; e.g., collaborative writing) and Tier 2 instruction (more individualized support; e.g., teach self-regulating strategies). | NA | 136 |
| Reid et al. (2014) | J | SR | SC | 2–12 | ADHD | 12 | SRSD improved writing elements, length, planning, quality, time, transition words, and vocabulary. | (0–100) | 37 |
| Rodgers and Graham (2008) | J | MA | SC | 1-12 | FR | 88 | Identified 9 effective writing instruction practices: goal setting; grammar instruction; prewriting; reinforcing specific writing outcomes; strategy instruction for planning/drafting, editing, and paragraph writing; teaching sentence construction; and word processing. | (51–99) | 507 |
| Sandmel and Graham (2011) | J | MA | E, QE | 1-12 | EL, FR, VD | 29 | Process writing improved writing quality for students in general education classrooms. Effects were not replicated for students in special education classrooms or ELs. | 0.34 | 481 |
| Santangelo and Graham (2016) | J | MA | E, QE | K-12 | FR | 80 | Handwriting instruction improved handwriting fluency and legibility. Effects generalized to writing fluency, length, and quality. | 0.48–0.84 | 207 |
| Santangelo et al. (2016) | C | MA | E, QE | K-12 | FR | ~88 | Zimmerman & Risemberg's 5 self-regulation processes (i.e., cognitive strategies, goal setting, mental imagery, self-selected models and tutors, self-evaluative standards) improved writing outcomes. | 0.30–1.06 | 70 |
| Other outcomes | | | | | | | | | |
| Graham and Hebert (2010)[c] | W | MA | E, QE | 1–12 | FR, LD | 93 | Identified 3 connections between writing and reading: teaching writing skills improved reading skills, writing about reading improved reading comprehension, and writing more improved reading comprehension. | 0.27–0.79 | 820 |
| Graham and Hebert (2011)[c] | J | MA | E, QE | 1–12 | FR, LD | 95 | Writing improved reading skills and comprehension. | 0.22–0.68 | 744 |

(continued)

**Table 2** (continued)

| References | Pub | Type of review | Included studies | | | | Key findings | ES (PND) range | Cited by |
|---|---|---|---|---|---|---|---|---|---|
| | | | Designs | Grades | Students | *k* | | | |
| Graham et al. (2020b) | J | MA | E, QE | 1–12 | FR | 56 | Writing about content (math, science, and social studies) improved learning. | 0.30 | |
| Graham et al. (2017b)[a] | J | MA | E, QE | PK-12 | FR | 47 | Balanced reading and writing instruction, with no more than 60% devoted to either, improved reading comprehension, decoding, and vocabulary. | 0.35–0.53 | 156 |
| Graham and Santangelo (2014)[a] | J | MA | E, QE | K-12 | FR | 53 | Teaching spelling improved phonological awareness and reading skills. | 0.44–0.51 | 385 |
| Hebert et al. (2013) | J | MA | E, QE | 1–12 | FR | 19 | Writing tasks were equally effective for improving reading, generally. Some writing activities were more effective for specific reading outcomes. | -0.20 to 0.24 | 141 |

*Note. Pub* publication, *ES* effect size, *PND* percentage of non-overlapping data, *J* journal, *MA* meta-analysis, *E* experimental, *QE* quasi-experimental, *LD* learning disability, *C* chapter, *SC* single-case, *A* average achievement, *AA* above average achievement, *BA* below average achievement, *FR* full range of achievement, *W* white paper, *SR* systematic review, *K* kindergarten, *RD* regression discontinuity, *ADHD* attention-deficit/hyperactivity disorder, *SRSD* self-regulated strategy development, *EBD* emotional/behavioral disorder, *MS* meta-synthesis, *QL* qualitative, *EL* English learner, *PK* pre-kindergarten, *VD* varying disabilities, *PS* post-secondary

[a]Number of unique experiments from the included syntheses is not provided; this is total number of studies used to compute effect sizes for all interventions listed in Table 2

[b]Included in both sections of this table

[c]Included most of the same studies, with one published in a peer-reviewed journal

Positive and statistically significant effect sizes of practices in the reviews ranged from 0.18 to 3.52, although not all practices were identified as effective. That said, nearly all of the instructional approaches led to positive and statistically significant effect sizes on a range of writing outcomes. Across all of the reviews, grammar instruction and motor instruction for handwriting seem to have been the only practices that led to negative effects on writing outcomes; grammar instruction led to negative effects in more than one review. One caveat to consider is that grammar instruction was sometimes compared to other writing approaches in these reviews. Graham has cautioned that it may not be that grammar instruction is ineffective, but that it is not as effective as the other writing instruction to which it has been compared.

On the other end of the spectrum, strategy instruction, including Self-Regulated Strategy Development, was the most consistently effective approach, as it was identified as an effective approach in 15 of the reviews. This is partly because it was the primary approach examined in several reviews, but it was also consistently one of the practices with the largest effect sizes in studies that included other practices. It was also found to be effective for the full range of students, struggling writers, average writers, good writers, English learners, students with learning disabilities, students with ADHD, and students in Turkey. It was also found to be effective in all levels of schooling (elementary, middle, high school, and college).

Graham and his colleagues also found more than 30 other practices to be effective for improving student writing outcomes (see Table 2). We identified three distinct groups of practices across the reviews. The first group involved instruction in specific writing practices, such as process writing, sentence combining, text structures, summarization, spelling skills, or handwriting skills, among others. The second group included instructional approaches educators use when teaching writing, such as creating an engaged community of writers, integrating reading and writing, modeling, creating routines, teaching students to make product goals, using different group sizes, and providing daily time to write. The third group involved assessing writing, providing students with feedback, and teaching students to evaluate their own writing. There was considerable overlap in instructional approaches examined across the different studies, often due to examining practices with different populations (e.g., elementary, secondary, students with learning disabilities, English learners), across different types of studies (e.g., group design, single-case design), or across different countries (e.g., U.S., Turkey). The consistency of the approaches across these variations is important and shows that there are many writing practices that are effective across populations and contexts.

**Studies of Educational Practices or Writing on Other Academic Outcomes** Graham conducted six reviews to examine whether writing or writing instruction impacted other academic outcomes, including reading and content area learning. These reviews included a range of 19–95 studies. All of the reviews revealed positive effects for writing on academic outcomes when compared to non-writing conditions. Additionally, comparisons made by Hebert et al. (2013) showed no statistically significant difference in the impact of different writing activities on reading

outcomes; however, the authors indicated there may have been too few studies to detect such effects. The findings across Graham's reviews indicate that having students write about reading and other content statistically significant impacts on student learning. Specifically:

- Asking students to write about their reading improved reading comprehension.
- Teaching students writing skills improved reading skills.
- Having students write more often improved reading comprehension.
- Teaching reading and writing together improved reading outcomes.
- Having students write about content in science, social studies, and math classes improved course grades.
- Teaching spelling skills improved reading outcomes.

Essentially, this body of work illustrates that writing influenced skill building, comprehension, and content learning in other academic areas.

## *4.2 Group Sub-population Comparisons*

Beyond understanding effective approaches for improving writing or learning, Graham has also conducted a substantial amount of research involving students with disabilities and their writing. Providing a foundation of his work on writing with students with disabilities are his systematic reviews and meta-analyses involving population comparisons. Such comparisons can include comparisons of groups or experimental comparisons of non-learning conditions. Six of Graham's syntheses fell into this category (see Table 3).

In four of the six reviews, Graham and his colleagues compared the writing skills of students in a group at-risk for learning problems (i.e., reading disabilities, learning disabilities, ADHD, and speech-language impairments) to typically developing peers. Across those four reviews, the at-risk student populations scored statistically significantly lower than same-aged peers without identified disabilities. Negative effect sizes from these meta-analyses indicated students in these populations consistently scored lower than their peers in writing, with an alarming range of moderate (−0.38) to very large (−1.42) effect sizes. It is also important to note that these populations were not necessarily identified with writing disabilities, and instead identified risks in areas that might be related to writing, such as language, reading, and attention. In addition to comparing subgroups to typically developing peers, two of the reviews (Graham et al., 2021, 2020a) also included meta-analyses of studies comparing students in at-risk populations (i.e., students with reading disabilities, students with speech-language impairments) to younger peers with similar reading or language achievement. In both reviews, the at-risk populations scored significantly lower than their peers on writing outcomes, suggesting the writing difficulties of these populations lag even further behind their language or reading skills.

In the two other comparison reviews in this group, Graham and his co-authors examined writing tools and assessment procedures. Morphy and Graham (2012)

**Table 3** Summary of Graham's group comparison reviews

| | Pub | Included Studies | | | | Key findings | ES range | Cited by |
|---|---|---|---|---|---|---|---|---|
| | | Designs | Grades | Students | *k* | | | |
| Graham et al. (2021) | J | E, Q, QE | PK-12 | RD, TD, YP | RQ1: 87 RQ2: 24 | RQ1: Compared to TD same-age peers, students with RD scored significantly lower on all writing measures (i.e., handwriting, organization, output, quality, sentences, spelling, syntax, vocabulary). RQ2: Compared to YPs with similar reading skills, students with RD scored significantly lower on writing measures (i.e., spelling). | RQ1: −1.42 to −0.64 RQ2: −0.94 to −0.93 | 14 |
| Graham et al. (2017a) | J | Q | 1–12 | LD, TD | 53 | Compared to TD same-age peers, students with LD scored significantly lower on all writing measures (i.e., conventions, genre elements, motivation, organization, output, quality, sentence fluency, vocabulary). | −1.14 to −0.42 | 185 |
| Graham et al. (2016b) | J | Q | 1–12 | ADHD, TD | 45 | Compared to TD same-age peers, students with ADHD scored significantly lower on all writing measures (i.e., genre elements, handwriting, output, quality, spelling, vocabulary). | –0.80 to −0.62 | 99 |
| Graham et al. (2011b) | J | Q | 1-PS | FR | 17 | Identified 4 factors that negatively influence evaluations of student writing: grammar errors, poor handwriting, spelling errors, and typed/printed text. | −1.03 to −0.38 | 185 |
| Graham et al. (2020a) | J | Q | 1–12 | SLI, TD, YP | RQ1: 39 RQ2: 6 | RQ1: Compared to TD same-age peers, students with SLI scored significantly lower on all writing measures (i.e., grammar, output, quality, spelling, vocabulary). RQ2: Compared to YPs with similar language skills, students with SLI scored significantly lower on general writing measures. | RQ1: −1.17 to −0.68 RQ2: −0.47 | 33 |
| Morphy and Graham (2012) | J | E, QE | 1–12 | RWD | 27 | Compared to writing by hand, word processing significantly improved writing development/organization, length, mechanical correctness, motivation, and quality. Students preferred word processing. | 0.48–1.46 | 229 |

*Note. Pub* publication, *ES* effect size, *E* experimental, *Q* quantitative, *QE* quasi-experimental, *PK* pre-kindergarten, *RD* reading difficulties, *TD* typically developing, *YP* younger peers, *RQ* research question, *LD* learning disability, *ADHD* attention-deficit/hyperactivity disorder, *PS* post-secondary, *FR* full range of achievement, *SLI* specific language impairment, *RWD* reading and/or writing difficulties

reviewed studies comparing writing with word processing tools versus writing by hand. Results showed students in the word processing conditions outperformed students who wrote by hand, across a variety of writing outcomes. Graham et al. (2011b) examined presentation effects that might bias the scorers of writing assessments, including, for example: typed versus handwritten text, compositions with corrected versus uncorrected spelling or grammar, and papers with versus without student names. The results consistently demonstrated presentation effects, indicating that humans who score students' writing can be biased by a variety of factors related to how it is presented.

### *4.3 General Writing Reviews*

The third category of reviews conducted by Graham include more general literature reviews. Originally, we intended to only include Graham's systematic reviews in this manuscript. However, we decided to also include selective reviews in this category because of the comprehensiveness of the non-systematic reviews conducted. For these reviews, Graham's purpose was to provide an overview of a particular set of literature and provide recommendations for practice. For example, in three of the reviews, Graham presented a unified approach for teaching spelling (Graham & Miller, 1979), handwriting (Graham & Miller, 1980), and composition (Graham, 1982). The reviews we included in this category are listed in Table 4.

We classified all of the studies in this group as selective general reviews, although one (Graham & Weintraub, 1996) also included a systematic review. These reviews range from including 14 studies of whole-language writing instruction to a comprehensive review of handwriting research with 208 references. The size and depth of these reviews preclude us from being able to easily summarize the findings. However, in several of the reviews, Graham provides findings that build a theoretical basis for his recommendations citing specific studies; a model for instruction and assessment in handwriting, spelling, and composition; and descriptions of specific activities with step-by-step instructions based on the recommendations of the original researchers.

## 5 A Supernova: The Impact of Steve's Work on the Field

As with most research, it is difficult to have a complete picture of the impact of the work. For instance, it is impossible to measure or know how many school administrators or teachers were impacted by reading Graham's reviews, reading a summary of the findings, hearing about the findings in a conference presentation, or learning about the findings through word of mouth. It is similarly difficult to know how these reviews have influenced other scholars. However, the number of times a research paper has been cited by other authors can provide a proxy for its influence.

**Table 4** Summary of Graham's general literature reviews

| | Pub | Included studies | | | Purpose/Contribution | Cited by |
|---|---|---|---|---|---|---|
| | | Grades | Students | *k* | | |
| Graham (1982) | J | Any | LD | 199[a] | Provided a model of composition for students with LD, including: theoretical background, model of instruction, empirical evidence for selected studies, and specific practice recommendations, grounded in research studies. | 130 |
| Graham and Harris (1994) | J | K-6 | FR | 14 | Examined the impact of whole-language instruction on student writing outcomes and narratively concluded: there is some improvement for K-2 students associated with whole- language models, but no significant differences when compared to other approaches; and more research is needed to draw conclusions for older students. | 109 |
| Graham and Harris (2009) | J | K-HS | LD | 55 | Summarized the writing research program conducted by Graham and Harris during the previous three decades, with a specific focus on students with LD. | 265 |
| Graham and Miller (1979) | J | School-aged | LD | 170[a] | Presented a model of spelling development and offered recommendations for instruction and assessment, based on existing research. | CD |
| Graham and Miller (1980) | J | School-aged | LD | 81[a] | Presented a model of handwriting development and offered recommendations for instruction and assessment, based on existing research. | 266 |
| Graham and Weintraub (1996) | J | PK-Adult | FR | 208[a] | Provided a comprehensive overview of handwriting research published from 1980 to 1994. | 578 |

*Note. Pub* publication, *J* journal, *LD* learning disability, *K* kindergarten, *FR* full range of achievement, *HS* high school, *CD* can't determine
[a]Number of references

At the start of 2023, Graham's reviews included in this paper had been cited an astonishing 15,600 times, with a mean of 400 citations per manuscript (median = 229) and a range of 7–2373 citations. It is important to note that the study with only 7 citations was published in 2022, so it had less than 1 year to be read and cited before this review. Citations are partially dependent on how long the study has been published. Based on that, a better citation measure might be the number of citations per year. We calculated the number of citations per year for each manuscript, and the mean citations per year across all of the manuscripts. Graham's reviews have been cited and average of 32.5 times per year each, range 0.3–98.5 citations per year across all of the reviews. For those interested in H-index metrics, Graham's would have an H-index score of 36 from his reviews alone, meaning that he has 36 review papers that have been cited at-least 36 times each.

In addition to examining citations, we also conducted a scoping review to situate Graham's work in the larger writing review literature. Our scoping review identified 318 reviews of writing research. This was based on liberal inclusion criteria, as we included any review that included 2 or more studies involving writing, even if writing was not the primary focus of the review. For example, if the authors of a review examined the impacts of a reading or vocabulary practice on academic outcomes but found two studies that included writing outcome measures, we determined that the study should be included in our scoping review. In such cases, we felt that such reviews may have captured the state of writing research in a particular area, even when the research was sparse.

The 318 reviews of writing research we identified included the 40 that identified Steve Graham as an author, meaning 277 were conducted by other researchers. In other words, Graham conducted 13% of the writing reviews identified in our search.

We were also interested in exploring how Graham has impacted reviews of writing research conducted by other researchers. Of the 277 reviews that did not include Steve Graham as an author, 117 (42%) cited Graham at-least once, and 73 (26%) cited one of Graham's reviews. In total, these 277 reviews of writing research included 658 unique citations of Graham's work overall, and his reviews were cited 186 times in this corpus. These are remarkable numbers that illustrate Graham's influence on other researchers, including his students and colleagues.

We also cross-referenced Steve's students and collaborators on his curriculum vitae with the authors listed on each of the reviews. We counted Steve's students as authors of a review if he listed himself as their direct advisor or on their dissertation committee on his CV. Similarly, we counted an author as one of Steve's collaborators if they were included anywhere on his CV, but were not already counted as one of his students. In total, 15 reviews (5.4%) that did not include Steve as an author involved at-least one of his students, and 52 (19%) involved someone Steve has collaborated with in some way.

## 6 Discussion: Bringing the Findings Back Down to Earth

The purpose of this paper was to provide an overview of Steve Graham's vast number of reviews of writing in education, which we affectionately and lightheartedly refer to as his Hitchhikers' Guide to Writing Research. With 40 reviews of educational writing research and counting, Graham is nothing short of prolific, and one of the most influential educational researchers in any area of study. His review work has examined the effectiveness of writing practices, compared populations of at-risk learners to typical learners in their writing skills, and provided in-depth reviews of a variety of topics.

## 6.1 Impacts of Writing Instruction

Some of the major findings of Graham work include the identification of more than 30 effective general practices for improving writing outcomes for students, with many more specific writing practices within those categories. For example, within strategy instruction alone, Graham has identified subgroups of effective planning strategies, revision strategies, self-regulation strategies, strategies for specific genres, etc. Of this work, it is difficult to find any instructional practice that he has not covered in some way, other than writing approaches that may have had too few primary studies to examine. One area that could be a gap might be instructional approaches aimed at teaching students to be aware of their audience. Another might be examining the impacts of improving students' content or background knowledge on writing outcomes. In both cases, it isn't clear whether there would be enough studies to conduct a review, although we know that one research team is exploring this for audience awareness.

Graham's work in reviewing effective practices for improving writing outcomes is also primarily in school-aged children. Although pre-kindergarten students and college students are included in some of these reviews, these populations were not the primary focus of Graham's reviews. The findings of his reviews also may not generalize to those populations, due to the small number of studies included for students in those age ranges, and the unique developmental stages of those populations. Reviews of effective practices for emergent writers and college students would be good extensions of Graham's work. Graham has also not conducted systematic reviews of effective practices for students learning a foreign language, although we identified some such studies conducted by others in our scoping review.

Another important finding from this work is that assessing students' writing and providing students with feedback also improve writing outcomes. These findings indicate that elaborate instructional interventions may not always be necessary for improving writing. The drawback is that assessing writing and providing feedback takes time. Graham et al. (2011a) have also examined whether there is any promise for automated essay scoring to impact writing outcomes, because this would save assessment time. Results were promising, but there were too few studies to make strong conclusions in 2011. With the development of artificial intelligence tools like ChatGPT, we expect this area of study to grow quickly.

## 6.2 Impacts of Writing on Reading and Learning

Graham's reviews of writing have also shown that writing and writing instruction improve reading and content learning. These reviews included impacts of writing on science, math, and social studies, as well as foundational reading skills and reading comprehension. The implications of this work are difficult to overstate, especially

considering the relative lack of national attention given to writing in the United States. Many teachers report not being prepared to teach writing and large percentages of teachers include only a minimal amount of writing in their classrooms (e.g., Gillespie et al., 2014; Goodrich et al., 2022; Graham et al., 2014; Hebert & Savaiano, 2020; Namkung et al., 2022; Powell et al., 2021). In the studies in Graham's reviews, writing was compared to typical classroom instruction in reading and content areas, indicating that simply including writing more often in schools has the potential to lead to improvements in writing, reading, and content learning. Some gaps in Graham's reviews in this area are largely related to age, as studies of college writers, adults, and (pre-)kindergarten writers were absent from most of these reviews. At-least in the case of pre-kindergarten and kindergarten writing, it may be that there were simply not enough studies of those populations at the time of Graham's reviews.

### *6.3 Comparing Populations or Practices*

The number of studies Graham has conducted in this area is small ($k = 6$), but influential. One of the more important findings to come out of this work is that students who are at-risk for learning difficulties due to reading, language, or ADHD also tend to perform lower than their typically developing peers in writing. This further highlights the need for including writing as a focal area of intervention for students in at-risk populations. Teachers and schools need to consistently assess the writing performance of students at-risk for other learning difficulties to also identify their writing strengths and weaknesses. The fact that students with reading disabilities and speech-language impairments scored lower than younger peers with comparable reading or language skills also indicates that writing will not automatically improve when reading or language skills improve. Targeted writing interventions are needed. Although these findings are impactful, the gaps in this literature are obvious. Many other subpopulations of students could be studied in this way, including English learners, students with autism spectrum disorder, students with intellectual disabilities, etc. We have not examined the studies found in our scoping review in enough detail to determine if researchers other than Graham have already started some of this work, but future research should examine this.

### *6.4 General Reviews*

General reviews conducted by Steve Graham earlier in his career were comprehensive, but not always systematic. That said, these reviews provide a comprehensive examination of specific programs of research in writing. Such work can provide a jumping-off point for other work in those areas. Aside from not always being systematic, one of the bigger weaknesses of these reviews is that they are now a bit

outdated. Three of these reviews were conducted more than 40 years ago, and only one was conducted after the year 2000. Current methods used for reviewing the literature, such as including systematic searches and using strict inclusion and exclusion criteria, might make these reviews candidates for updating. Still, the approaches used for these reviews provide some advantages over some systematic reviews. For example, they include findings from studies that used different methodologies to show a more complete picture of the writing literature in a particular area. This work requires a painstaking amount of time and knowledge of the literature, and the literature has grown considerably since the time Graham started this work. Therefore, it may be more reasonable to expect that future researchers will work on smaller subsets of the literature, or limit the review of literature to specific and recent date ranges to ensure the literature is relevant to today's school contexts.

### *6.5 Limitations of This Review*

The primary limitation of this scoping review is the bias of the authors toward Steve Graham's work. Two of us are Graham's students and authors on several of his reviews and meta-analyses, so we may not have been as critical as we should have been, or we may have made some assumptions about some of the work that more objective authors may not have made. We also presented only the positive outcomes of Steve Graham's reviews, but did not provide information on non-significant or negative findings due to space.

Another limitation is that we based our search for Graham's manuscripts and collaborators on our ability to cross reference off his CV. This approach may have overlooked colleagues not on his CV that he also may have influenced in some way. For example, colleagues Graham worked with at his affiliated universities but did not formally collaborate with, colleagues he collaborated with on work that did not make it onto his CV, or colleagues that were influenced by things like conference talks or presentations. That said, this likely tempered our conclusions on his influence more than any other form of bias.

## 7 Conclusion: Future Hitchhiking Missions

Steve Graham's contributions to educational writing literature are unparalleled. The purpose of this current review was to simply provide an overview of his work reviewing the literature. However, his work provides us with a guide to the galaxy of writing research that can be used to further the development of new instructional approaches and interventions aimed at improving writing, advance theory around the mechanisms for how writing might inform reading and learning outcomes, identify populations of students that may need more support in developing writing skills, and help us develop deep understandings of comprehensive approaches to

improving writing. Our scoping review may be used as a jumping-off point to summarize and synthesize the broader universe of writing research. As Graham noted, this can give us a guide to what science tells us about writing and help us identify the strengths and weaknesses in the work of the field. It can also help us understand how writing is situated in the educational universe.

## References

Bangert-Drowns, R. L., Hurley, M. M., & Wilkenson, B. (2004). The effects of school-based writing-to-learn interventions on academic achievement: A meta-analysis. *Review of Educational Research, 74*(1), 29–58. https://doi.org/10.3102/00346543074001029

Birr Moje, E., Afflerback, P. P., Enciso, P., & Lesaux, N. K. (Eds.). (2020). *Handbook of reading research* (Vol. 5). Routledge.

Cooper, S., Hebert, M., Goodrich, J. M., Cardona, S. L., Lin, X., Peng, P., & Nelson, J. R. (2022). Effects of automaticity training of fundamental literacy component skills on reading performance: A meta-analysis. *Journal of Behavioral Education*. Advance online publication. https://doi.org/10.1007/s10864-022-09480-7

Farr, F., & Murray, L. (Eds.). (2016). *Routledge handbook of language learning and technology*. Routledge.

Gass, S. M., & Mackey, A. (Eds.). (2012). *Routledge handbook of second language acquisition*. Routledge.

Gillespie, A., & Graham, S. (2014). A meta-analysis of writing interventions for students with learning disabilities. *Exceptional Children, 80*(4), 454–473. https://doi.org/10.1177/0014402914527238

Gillespie, A., Graham, S., Kiuhara, S., & Hebert, M. (2014). High school teachers' use of writing to support students' learning: A national survey. *Reading and Writing, 27*, 1043–1072. https://doi.org/10.1007/s11145-013-9494-8

Goodrich, J. M., Hebert, M., & Namkung, J. (2022). Impacts of the COVID-19 pandemic on elementary school teachers' practices and perceptions across the spring and fall 2020 semesters. *Frontiers in Education, Educational Psychology, 6*, 793285. https://doi.org/10.3389/feduc.2021.793285

Graham, S. (1982). Composition and practice: A unified approach. *Focus on Exceptional Children, 14*(8), 1–16.

Graham, S. (1987). Writing assessment. In C. Reynolds & L. Mann (Eds.), *Encyclopedia of special education* (1st ed., pp. 1671–1673). Wiley.

Graham, S. (2006). Strategy instruction and the teaching of writing: A meta-analysis. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 187–207). Guilford.

Graham, S. (2018). A revised writer(s)-within-community model of writing. *Educational Psychologist, 53*(4), 258–279. https://psycnet.apa.org/doi/10.1080/00461520.2018.1481406

Graham, S. (2021). Through the looking glass: Reflections of a writing scholar. Acquired Wisdom Series, S. Nieto, F. Erickson & P. Winne (Eds.). *Education Review, 28,* 1–19. https://doi.org/10.14507/er.v28.3209.

Graham, S., & Harris, K. R. (1994). The effects of whole language on children's writing: A review of literature. *Educational Psychologist, 29*(4), 187–192. https://doi.org/10.1207/s15326985ep2904_2

Graham, S., & Harris, K. R. (2003). Students with learning disabilities and the process of writing: A meta-analysis of SRSD studies. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 323–344). Guilford.

Graham, S., & Harris, K. R. (2009). Almost 30 years of writing research: Making sense of it all with The Wrath of Khan. *Learning Disabilities Research & Practices, 24*(2), 58–68. https://doi.org/10.1111/j.1540-5826.2009.01277.x

Graham, S., & Harris, K. R. (2018). Evidence-based writing practices: A meta-analysis of existing meta-analyses. In R. Fidalgo, K. R. Harris, & M. Braaksma (Eds.), *Design principles for teaching effective writing: Theoretical and empirical grounded principles (pp. 13-37)*. Brill Editions. https://doi.org/10.1163/9789004270480_003

Graham, S., & Harris, K. R. (2019). Evidence-based practices in writing. In S. Graham, C. MacArthur, & M. Hebert (Eds.), *Best practices in writing instruction* (3rd ed., pp. 3–28).

Graham, S., & Hebert, M. (2010). *Writing to read: Evidence for how writing can improve reading. A Carnegie Corporation time to act report.* Alliance for Excellent Education. https://www.carnegie.org/publications/writing-to-read-evidence-for-how-writing-can-improve-reading/

Graham, S., & Hebert, M. (2011). Writing to read: A meta-analysis of the impact of writing and writing instruction on reading. *Harvard Educational Review, 81*(4), 710–744. https://doi.org/10.17763/haer.81.4.t2k0m13756113566

Graham, S., & Miller, L. (1979). Spelling research and practice: A unified approach. *Focus on Exceptional Children, 12*(2), 1–16.

Graham, S., & Miller, L. (1980). Handwriting research and practice: A unified approach. *Focus on Exceptional Children, 13*(2), 1–16.

Graham, S., & Perin, D. (2007a). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*(3), 445–476. https://doi.org/10.1037/0022-0663.99.3.445

Graham, S., & Perin, D. (2007b). What we know, what we still need to know: Teaching adolescents to write. *Scientific Studies of Reading, 11*(4), 313–335. https://doi.org/10.1080/10888430701530664

Graham, S., & Perin, D. (2007c). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools*. Alliance for Excellent Education. https://www.carnegie.org/publications/writing-next-effective-strategies-to-improve-writing-of-adolescents-in-middle-and-high-schools/

Graham, S., & Santangelo, T. (2014). Does spelling instruction make students better spellers, readers, and writers? A meta-analytic review. *Reading and Writing, 27*(9), 1703–1743. https://doi.org/10.1007/s11145-014-9517-0

Graham, S., & Weintraub, N. (1996). A review of handwriting research: Progress and prospects from 1980 to 1994. *Educational Psychology Review, 8*(1), 7–87. https://doi.org/10.1007/BF01761831

Graham, S., Harris, K. R., & Mason, L. (2005). Improving the writing performance, knowledge, and self-efficacy of struggling young writers: The effects of self-regulated strategy development. *Contemporary Educational Psychology, 30*(2), 207–241. https://psycnet.apa.org/doi/10.1016/j.cedpsych.2004.08.001

Graham, S., Harris, K. R., & Hebert, M. (2011a). *Informing writing: The benefits of formative assessment. A Carnegie Corporation time to act report*. Alliance for Excellent Education. https://www.carnegie.org/publications/informing-writing-the-benefits-of-formative-assessment/

Graham, S., Harris, K. R., & Hebert, M. (2011b). It is more than just the message: Presentation effects in scoring writing. *Focus on Exceptional Children, 44*(4), 1–12.

Graham, S., Bollinger, A., Booth Olson, C., D'Aoust, C., MacArthur, C., McCutchen, D., & Olinghouse, N. (2012a). *Teaching elementary school students to be effective writers* (NCEE 2012-4058). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/PracticeGuide/17

Graham, S., McKeown, D., Kiuhara, S., & Harris, K. R. (2012b). A meta-analysis of writing instruction for students in the elementary grades. *Journal of Educational Psychology, 104*(4), 879–896. https://doi.org/10.1037/a0029185

Graham, S., Harris, K. R., & McKeown, D. (2013). The writing of students with learning disabilities, meta-analysis of self-regulated strategy development writing intervention studies, and future directions: Redux. In H. L. Swanson, K. R. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (2nd ed., pp. 405–438).

Graham, S., Capizzi, A., Harris, K., Hebert, M., & Morphy, P. (2014). Teaching writing to middle school students: A national survey. *Reading and Writing, 27*, 1015–1042. https://doi.org/10.1007/s11145-013-9495-7

Graham, S., Harris, K. R., & Santangelo, T. (2015a). Research-based writing practices and the common core: Meta-analysis and meta-synthesis. *Elementary School Journal, 115*(4), 498–522. https://doi.org/10.1086/681964

Graham, S., Hebert, M., & Harris, K. R. (2015b). Formative assessment and writing: A meta-analysis. *Elementary School Journal, 115*(4), 523–547. https://doi.org/10.1086/681947

Graham, S., Bruch, J., Fitzgerald, J., Friedrich, L., Furgeson, J., Greene, K., Kim, J., Lyskawa, J., Booth Olson, C., & Smither Wulsin, C. (2016a). *Teaching secondary students to write effectively* (NCEE 2017-4002). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/PracticeGuide/22

Graham, S., Fishman, E. J., Reid, R., & Hebert, M. (2016b). Writing characteristics of students with attention deficit hyperactivity disorder: A meta-analysis. *Learning Disabilities Research & Practice, 31*(2), 75–89. https://doi.org/10.1111/ldrp.12099

Graham, S., Collins, A. A., & Rigby-Wills, H. (2017a). Writing characteristics of students with learning disabilities and typically achieving peers: A meta-analysis. *Exceptional Children, 83*(2), 199–218. https://doi.org/10.1177/0014402916664070

Graham, S., Liu, X., Ng, C., Bartlett, B., Harris, K. R., & Holzapfel, J. (2017b). Effectiveness of literacy programs balancing reading and writing instruction: A meta-analysis. *Reading Research Quarterly, 53*(3), 279–304. https://doi.org/10.1002/rrq.194

Graham, S., Liu, X., Bartlett, B., Ng, C., Harris, K. R., Aitken, A., Barkel, A., Kavanaugh, C., & Talukdar, J. (2018). Reading for writing: A meta-analysis of the impact of reading interventions on writing. *Review of Educational Research, 88*(2), 243–284. https://doi.org/10.3102/0034654317746927

Graham, S., MacArthur, C., & Hebert, M. (Eds.). (2019). *Best practices in writing instruction* (3rd ed.).

Graham, S., Hebert, M., Fishman, E., Ray, A. B., & Gillespie Rouse, A. (2020a). Do children classified with specific language impairment have a learning disability in writing? *A meta-analysis. Journal of Learning Disabilities, 53*(4), 292–310. https://doi.org/10.1177/0022219420917338

Graham, S., Kiuhara, S. A., & MacKay, M. (2020b). The effects of writing on learning in science, social studies, and mathematics: A meta-analysis. *Review of Educational Research, 90*(2), 179–226. https://doi.org/10.3102/0034654320914744

Graham, S., Aitken, A., Hebert, M., Santangelo, T., Camping, A., Harris, K. R., Eustice, K., Sweet, J. D., & Ng, C. (2021). Do children with reading difficulties experience writing difficulties? A meta-analysis. *Journal of Educational Psychology, 113*(8), 1481–1506. https://doi.org/10.1037/edu0000643

Graham, S., Tavsanli, O., & Kaldirim, A. (2022). Improving writing skills of students in Turkey: A meta-analysis of writing interventions. *Educational Psychology Review, 34*, 889–934. https://doi.org/10.1007/s10648-021-09639-0

Hall, G. (Ed.). (2016). *Routledge handbook of English language teaching*. Routledge.

Hebert, M., & Savaiano, M. (2020). A survey of the writing instructional practices of Nebraska teachers of students with visual impairments. *Exceptionality, 29*(4), 294–309. https://doi.org/10.1080/09362835.2020.1850450

Hebert, M., Gillespie, A., & Graham, S. (2013). Comparing effects of different writing activities on reading comprehension: A meta-analysis. *Reading and Writing, 26*(1), 111–138. https://doi.org/10.1007/s11145-012-9386-3

Hebert, M., Bohaty, J. J., Nelson, J. R., & Brown, J. A. (2016). The effects of text structure instruction on expository reading comprehension: A meta-analysis. *Journal of Educational Psychology, 108*(5), 609–629. https://doi.org/10.1037/edu0000082

Hillocks, G., Jr. (1984). What works in teaching composition: A meta-analysis of experimental treatment studies. *American Journal of Education, 93*(1), 133–170. https://www.jstor.org/stable/1085093

Hillocks, G., Jr. (2002). *The testing trap: How state writing assessments control learning*. Teachers College.

Hinkel, E. (Ed.). (2017). *Handbook of research in second language teaching and learning* (Vol. 3).

Joshi, R. M., & Aaron, P. G. (Eds.). (2005). *Handbook of orthography and literacy*. Routledge.

Kim, Y.-S., & Graham, S. (2022). Expanding the direct and indirect effects model of writing (DIEW): Reading-writing relations, and dynamic relations as a function of measurement/dimensions of written composition. *Journal of Educational Psychology, 114*(2), 215–238. https://doi.org/10.1037/edu0000564

MacArthur, C. A., Graham, S., & Fitzgerald, J. (Eds.). (2016). *Handbook of writing research* (2nd ed.).

Marsh, J., & Larson, J. (Eds.). (2012). *SAGE handbook of early childhood literacy*. Sage.

Mason, L. H., & Graham, S. (2008). Writing instruction for adolescents with learning disabilities: Programs of intervention research. *Learning Disabilities Research & Practice, 23*(2), 103–112. https://doi.org/10.1111/j.1540-5826.2008.00268.x

McGill-Franzen, A., & Allington, R. L. (Eds.). (2011). *Handbook of reading disability research*. Routledge.

McKenna, M. C., Labbo, L. D., Kieffer, R. D., & Reinking, D. (Eds.). (2006). *International handbook of literacy and technology* (Vol. 2). Routledge.

Morphy, P., & Graham, S. (2012). Word processing programs and weaker writers/readers: A meta-analysis of research findings. *Reading and Writing, 25*(3), 641–678. https://doi.org/10.1007/s11145-010-9292-5

Namkung, J., Goodrich, J. M., Hebert, M., & Koziol, N. (2022). Impacts of the COVID-19 pandemic on student learning and opportunity gaps across the 2020-2021 school year: A national survey of teachers. *Frontiers in Education, Educational Psychology, 7*, 921497. https://doi.org/10.3389/feduc.2022.921497

National Aeronautics and Space Administration. (2001). Space shuttle use of propellants and fluids. *John F. Kennedy Space Center.* FS-2001-09-015-KSC.

National Commission on Writing. (2003). *The neglected "R": The need for a writing revolution*. College Entrance Examination Board.

Newell, G. E., VanDerHeide, J., & Olsen, A. W. (2014). High school English language arts teachers' argumentative epistemologies for teaching writing. *Research in the Teaching of English, 49*(2), 95–119. https://www.jstor.org/stable/24398670

Olson, D. R., & Torrance, N. (Eds.). (2009). *Cambridge handbook of literacy*. Cambridge University Press.

Powell, S. R., Hebert, M. A., Cohen, J. A., Casa, T. M., & Firmender, J. M. (2017). A synthesis of mathematics writing: Assessments, interventions, and surveys. *Journal of Writing Research, 8*(3), 493–526. https://doi.org/10.17239/jowr-2017.08.03.04

Powell, S. R., Hebert, M. A., & Hughes, E. M. (2021). How educators use mathematics-writing in the classroom: A survey of mathematics educators. *Reading and Writing, 34*, 417–447. https://doi.org/10.1007/s11145-020-10076-8

Reid, G., Fawcett, A. J., Manis, F., & Siegel, L. S. (Eds.). (2008). *SAGE handbook of dyslexia*. Sage.

Reid, R., Hagaman, J. L., & Graham, S. (2014). Using self-regulated strategy development for written expression with students with attention deficit hyperactivity disorder. *Learning Disabilities: A Contemporary Journal, 12*(1), 21–42.

Rogers, L. A., & Graham, S. (2008). A meta-analysis of single subject design writing intervention research. *Journal of Educational Psychology, 100*(4), 879–906. https://doi.org/10.1037/0022-0663.100.4.879

Rowsell, J., & Pahl, K. (Eds.). (2015). *Routledge handbook of literacy studies*. Routledge.

Sandmel, K., & Graham, S. (2011). The process writing approach: A meta-analysis. *Journal of Educational Research, 104*(6), 396–407. https://doi.org/10.1080/00220671.2010.488703

Santangelo, T., & Graham, S. (2016). A comprehensive meta-analysis of handwriting instruction. *Educational Psychology Review, 28*(2), 225–265. https://doi.org/10.1007/s10648-015-9335-1

Santangelo, T., Harris, K. R., & Graham, S. (2016). Self-regulation and writing: Meta-analysis of the self-regulation processes in Zimmerman and Risemberg's model. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (2nd ed., pp. 174–193).

Stone, C. A., Silliman, E. R., Ehren, B. J., & Wallach, G. P. (Eds.). (2016). *Handbook of language and literacy: Development and disorders* (2nd ed.). Guilford.

Tate, T., & Warschauer, M. (2022). Access, digital writing, and achievement: The data in two diverse school districts. *Journal of Writing Assessment, 15*(1), 1–36. https://escholarship.org/uc/item/8rb1j91b

U.S. Department of Education. (2022). *National Assessment of Educational Progress: Assessment calendar*. https://nces.ed.gov/nationsreportcard/about/calendar.aspx

Wasik, B. H. (Ed.). (2012). *Handbook of family literacy* (2nd ed.). Routledge.

Web of Science. (2022). https://www.webofscience.com/wos/woscc/citation-report/4ddb39b5-bcaa-403b-ae79-9bb4e3732205-5bafa6b3

Wyse, D., Andrews, R., & Hoffman, J. (Eds.). (2010). *Routledge international handbook of English, language and literacy teaching*. Routledge.