

CHAPTER 10.

IMPLICATIONS OF LONGITUDINAL WRITING RESEARCH METHODS FOR LIFESPAN PERSPECTIVES ON WRITING DEVELOPMENT: A SYSTEMATIC REVIEW

Jonathan M. Marine

George Mason University

Paul Rogers

University of California, Santa Barbara

Teresa Jacques

University of Porto

Like all research methods, longitudinal research methods are linked to epistemologies, axiologies, and ideologies which help shape and undergird lines of inquiry and the particulars of individual research studies. Longitudinal research methods are common across many disciplines including psychology, biology, economics, education, neurology, gerontology, and other subfields in the health sciences (Hedeker & Gibbons, 2006; Fitzmaurice et al., 2012). Longitudinal research serves two primary purposes, “to describe patterns of change and to establish the direction and magnitude of causal relationships” (Lewis-Beck et al., 2004). Typically, longitudinal studies use ongoing, recurrent measures to follow individuals over prolonged periods of time or across ages frequently “without any external influences being applied” (Caruana et al., 2015). While cross-sectional methods attempt to analyze several variables at a given point in time in order to examine differences *between* cases, longitudinal studies instead foreground the influence of time on the variables being measured in order to examine changes *within* cases (Lewis-Beck et al., 2004). Given the importance of understanding the relationships among and between the many variables and factors associated

with acquiring writing abilities and how those abilities change over time, longitudinal methods are especially useful in studying writing development.

Longitudinal writing research methods can be broadly categorized as either qualitative, quantitative or mixed methods in nature (Hartley & Chesworth, 2000). Writing research benefits from work in all of these methodological categories because, broadly speaking, qualitative research methods allow researchers to explore new phenomena in order to generate hypotheses while quantitative research allows researchers to investigate causality and correlation in order to test hypotheses (Sullivan & Sergeant, 2011). Quantitative studies use positivist paradigms as their basis while investigating a single or limited set of related variables while qualitative research uses a more phenomenological paradigm (Firestone, 1987) and attends to the multiple variables associated with context. This distinction between quantitative and qualitative research methods plays an important role in our review, and in the section below we further elaborate on the importance of these distinctions.

QUALITATIVE LONGITUDINAL WRITING RESEARCH

Qualitative longitudinal writing research includes many different types of methods of research and research designs. From single-subject interview-based studies that follow a student throughout their entire school career (e.g., Svensson, 2018), to text-based studies that attempt to gauge how college students' literate lives beyond the academy shape their writing in the classroom (e.g., Roozen, 2008), to grounded theory-based accounts of how extracurricular writing helps to support students' development of voice through writing (e.g., Chen, 2017), qualitative studies of longitudinal writing all share the common concern of trying to understand how writers develop over time by attending closely to a wide range of complex contexts and situations.

QUANTITATIVE LONGITUDINAL WRITING RESEARCH

Typically experimental or correlational in nature, quantitative methods usually aim to combine a wide array of measures (textual, spoken, or otherwise) into a few key points of data collection in order to try and gauge how one variable (e.g., particular language abilities or skills) might track or correlate with another (e.g., literacy and vocabulary, spelling and reading comprehension, phonemic reading and written expression, etc.). From quasi-experimental studies that investigate what elements of neuropsychological development contribute to writing development in first graders (Hooper et al., 2010), to studies which seek to trace correlations between reading and writing motivation across multiple cohorts

of primary school students (Hamilton et al., 2013), to studies which attempt to describe the strategies used in undergraduate writing (Torrance et al., 2000), to studies which seek to account for extracurricular factors that might influence children's writing development from ages 4 to 7 (Dunsmuir & Blatchford, 2004), quantitative studies of longitudinal writing all share the common goal of trying to capture in what ways and to what degree the varying factors which influence writing development relate to, predict, and correlate with one another.

MIXED METHODS LONGITUDINAL WRITING RESEARCH

An important alternative research design is mixed methods research, which includes both quantitative and qualitative research used in a single study (Doyle et al., 2009). Mixed methods research can deepen the information retrieved from a study, ultimately leading to more informed findings; however, mixed methods can be time-consuming (Almalki, 2016; Greene et al., 1989). From studies which attempt to investigate changes in the nature and amount of preschoolers' parental writing support (Skibbe et al., 2013), to single-subject studies which attempt to examine the development of figurative competence in narrative writing from elementary school through high school and beyond (Svensson, 2018), longitudinal mixed methods studies of writing draw on both qualitative and quantitative methods in combination with one another in order to quantify the phenomena under investigation while also accounting for the contexts in which the phenomena occurs.

Putting the systematic review methodology into action, in the section below, we present the results of our longitudinal review in order to illuminate the methodological characteristics and trends of longitudinal writing research during the past 21 years. Although at this current juncture, these studies address shorter time sequences than the lifespan model ultimately demands, these smaller stories are important for continuing the pursuit of what Bazerman (2018) called, "the impossible dream" of an empirically grounded lifespan view of writing development.

RESULTS

In this section, we present the results of our systematic review of L1, international, longitudinal, studies of writing since the year 2000 across the following categories:

- Longitudinal Writing Research Design
- Educational and Study Settings

- Longitudinal and Participant Characteristics
- Methods of Data Collection and Analysis
- Publication of Longitudinal Writing Research
- Funding of Longitudinal Writing Research

LONGITUDINAL WRITING RESEARCH DESIGN

Research Questions. Research questions are at the center of research design, and the longitudinal studies of writing in our review provide a rich landscape of inquiry to consider in the light of lifespan approaches to writing. At the most general level, the research designs employed in the studies we reviewed used three primary methods to frame their inquiries: research questions, hypothesis driven studies, and what we refer to as purpose driven or goal-oriented studies.

Studies framed with research questions use typical interrogative words and phrases such as *how*, *what*, and *to what extent* to guide their studies. Researchers using primarily hypothesis driven approaches usually present their research with phrases like “we hypothesized that” Purpose driven or goal-oriented studies framed their work around categories of action like investigating, identifying, verifying, and describing.

Of the 54 studies we reviewed, 27 studies were framed primarily with research questions, 13 were framed primarily with hypotheses, and 15 studies used a goal orientation. This is not to say there was no overlap between these categories. Indeed, of the 27 studies primarily organized around research questions, six also included goal statements, and three included hypotheses, though these appeared in a subordinate fashion to the research questions. Likewise, of the 13 hypothesis driven studies, nine also included some kind of purpose statement.

In regards to studies framed around research questions, 12 studies presented a single research question, while 15 used multiple questions (ranging from 2 questions to 7 questions) with an average of almost 4 questions per study for studies with multiple questions. The total number of questions from all of the studies we marked as driven by research questions was 67. The most common question stems were *what* (21), *how* (16), and variations of to which extent and *to what degree* (7). Other interrogative question stems included: *are*, *does*, *will*, *which*, *do*, and *whether*. Of the 10 hypothesis driven studies, 9 pursued multiple hypotheses. Purpose driven or goal-oriented studies used a variety of keywords, the most common of which were *explore*, *examine*, and *investigate* with the objects of these goals most frequently relating to inquiries related to development, specific theoretical claims and models, or particular units of analysis, especially longitudinal relationships, between, for example, reading and writing.

In considering the degree to which longitudinal writing researchers replicate other studies using similar or the same questions, and the degree to which writing researchers conduct studies that can be aggregated with other studies to help build out areas of inquiry and knowledge (Haswell, 2005), the results are mixed. In the main, we found almost no precise replication studies in our corpus. That is, the detailed research design elements differ considerably from study to study whether quantitative or qualitative. Only one qualitative study (Myhill & Jones, 2007) provided enough detail to qualify as a potentially replicable study. However, there are research groups who continue to build out lines of inquiry based on the results of their previous studies (e.g., Bigozzi & Vettori, 2016).

Types of Research Questions. In terms of research questions, we identified three main types. Most common are questions that investigate how writing changes over time in relationship to specific elements, factors, or variables such as increases in content knowledge, the amount and quality of parental support, or the various impacts of particular skills (skills being a very common attribute across the studies). The second most common type of question focused on specific measures and the degree to which those measures are effective predictors of change in writing, usually with a very specific unit of analysis that was related to the measures. The third type of questions focus on descriptive research aims.

Framing of Hypothesis Driven Studies. Similarly, hypothesis driven longitudinal writing research includes three major categories of hypotheses. First are predictor-driven studies that seek to present clear evidence for the predictive value of particular elements and skills associated with writing. Next are hypothesis driven studies that attempt to define more clearly the relationships between and among factors and elements such as correlation studies. A number of the hypothesis driven studies focus narrowly on specific elements of a particular model and the degree to which that model is efficacious in predicting future writing growth. Finally, the purpose-driven studies in our review appear to be approaching their areas of inquiry with an eye towards building out further the larger programs of research and lines of inquiry.

Study Settings (Country, Grade Level, and Public or Private Schools). The longitudinal studies of writing in our corpus were conducted in 13 different countries: United States (23), Italy (5), the UK (3), Portugal (3), China (3), Israel (3), Canada (2), Netherlands (2), Hong Kong (2), Brazil (2), Denmark (2), South Korea (2), Sweden (1), France (1), Australia (1), Russia (1) and Argentina (1). In addition to coding for the geographical location of the studies, we also identified the language under investigation (L1) within each article. Our final data corpus included 14 different languages: Arabic, Cantonese, Chinese, Danish, Dutch, English, French, Hebrew, Italian, Korean, Portuguese, Russian,

Spanish, and Swedish. The three most common languages were English (28), Italian (5), and Portuguese (5). Studies took place across a wide range of grade levels. The most prevalent educational settings for studies of longitudinal writing in our corpus were elementary school (16), university/undergraduate (9), followed by kindergarten (4) and preschool (4), secondary education/high school (3), home (3), middle school (1), university/postgraduate (1), and out of school (1). Eleven studies included settings that crossed typical school designations. These included kindergarten + elementary (3), kindergarten + 1st grade (2), kindergarten + 2nd grade (1), kindergarten + 4th grade (1), kindergarten + 1st grade + 2nd grade (1), kindergarten through 12th grade (1), 1st grade through 6th grade (1), and undergraduate through graduate (1). While the majority of studies did not specify whether the schools were public or private (28/56), of those which did specify, the majority (22) were set in public schools, with only four studies taking place in a private school setting.

Longitudinal and Participant Characteristics. As a part of our coding schema, we looked at the longitudinal characteristics of each study by coding for the number of points of measure and the duration of each study in our corpus. The studies in our corpus of longitudinal writing research averaged 2.9 points of measure per study. Points of measure ranged from 2 (our minimum for inclusion) to 6 points across the 40 articles which specified this information. The remaining studies either had several points of measure (10) but didn't specify how many or were simply unclear (5).

In terms of study duration, the studies in our corpus averaged 3.8 years across which data was collected, ranging from 12 weeks (a quantitative study; Rosário et al., 2017) to 30 years (an ethnographic type of study; Smith & Prior, 2020). Although no study was precisely the mean of 3.8 years in length, one representative study (Yeung et al., 2013a) lasted for 4 years and had 3 points of measure. This funded study was conducted in China, with native Cantonese speakers from 1st to 4th grade, and aimed to examine the relationships between cognitive-linguistic skills that are important to Chinese children's writing development, based on the model of the developmental constraints on writing acquisition (Berninger et al., 1991).

We also coded for the total number of participants, reported age range, and gender distribution. The 54 studies in our corpus involved 6714 total participants at an average of 126.68 participants per study, ranging from 1 to 481 total participants. Coding for age ranges of participants proved problematic as many of the studies reported only grade levels and not age levels. From the studies which did report age averages and ranges, the average reported age was 5.9 years old, ranging from 3.61 (a number of studies looked at the transition from preschool to kindergarten e.g., Skibbe et al., 2013) to 30.95 years old (a number

of studies looked at the transition from university to the workplace or graduate school e.g., Chen, 2017).

More clear was the gender distribution of participants in these studies, with only 1.9 percent of articles not clearly reporting the gender distribution of their participants. 82 percent of the studies in our corpus had both male and female participants, with 7.7 percent reporting all male participants and 7.7 percent reporting all female participants. No studies in this corpus reported on gender categories besides male and female. In studies where gender distributions were reported, an average of 58.8 percent of participants were female.

Method of Data Collection and Analysis in Longitudinal Writing Research.

One of the most critical elements of any research study involves the systematic collection of data. In this review we coded for the data collection techniques of each study by creating both high-inference and low-inference categories for methods of data collection. While our preference was to use the actual language from the published studies in our analysis (i.e., the low-inference category), we used the high inference categories to generalize more broadly regarding methods of data collection. The high inference categories were: ethnography, observation, interviews, document and artifact collection, descriptive research, correlational research, experimental, quasi-experimental, other, and not specified. (More detailed information on our coding methods can be found in chapter 9.)

The most prevalent methods of longitudinal research in our corpus were quasi-experimental which was employed in almost half of the studies in our corpus (49 percent), followed by experimental research (11 percent), ethnography (9 percent), document and artifact collection (9 percent), other (7 percent), correlational research (6 percent), descriptive research (5 percent) and finally, interviews (4 percent). Within quantitative studies 26 were quasi-experimental (e.g., Aram & Levin, 2004; Kuzeva et al., 2015; Niedo et al., 2014), 6 experimental (e.g., Cordeiro et al., 2020; Drijbooms et al., 2017), and 2 correlational (e.g., Pinto et al., 2009). Three studies which did not fit our initial coding schema were coded as other (10.6 percent). These studies included qualitative studies where the method of data collection was either not stated, unclear, or included more than one type of method (e.g., causal and exploratory quantitative studies or mixed-methods).

Data Collection in Quantitative Studies. In the quantitative studies in our corpus, the most common measures used to research writing were spelling (see Abbott et al., 2010; Beers & Nagy, 2011; Cardoso-Martins et al., 2006; Cordeiro et al., 2020; Kim & Park, 2019; Kim et al., 2015; Limpo & Alves, 2013; Pinto et al., 2015; Treiman et al., 2013; Yeung et al., 2020) and essay and story writing (see Abbott et al., 2010; Bigozzi & Vettori, 2016; Coker, 2006; Cordeiro et al., 2020; Drijbooms et al., 2017; Limpo & Alves, 2013; Mäki et al., 2001;

Oppenheimer et al., 2017; Pinto et al., 2015; Rosário et al., 2017; Tong & McBride, 2016; Woodward-Kron, 2009; Yeung et al., 2013b; Yeung et al., 2020). Spelling was commonly measured through direct dictation tasks (e.g., Yan et al., 2012; Yeung et al., 2013a), the percentage of words spelled correctly in a story (Limpó & Alves, 2013), and by accurately copying characters (Fischer & Koch, 2016). Tasks used to assess writing include narrative writing (“Tell a story about a child who lost his or her pet”; Limpó & Alves, 2013), opinion essays (“Do you think teachers should give students homework every day?”; Limpó & Alves, 2013), persuasive opinion essays (Oppenheimer et al., 2017), descriptive writing (“Describe a happy birthday scene”; Yeung et al., 2013b), expository writing, and scientific writing (Oppenheimer et al., 2017).

Handwriting was another common measure used to investigate writing by the quantitative studies in our corpus (see Cordeiro et al., 2020; Kim & Park, 2019; Limpó & Alves, 2013; Yan et al., 2012; Yeung et al., 2013a). Handwriting was used as a measure in studying fluency (Kim & Park, 2019; Limpó & Alves, 2013; Yan et al., 2012), stroke order tasks (in Chinese; Yeung et al., 2013a) and letter writing automaticity (Kim et al., 2015). Orthographic skills (e.g., hyphenation, capitalization, punctuation, etc.) were also measured in various studies (Bigozzi & Vettori, 2016; Pinto et al., 2012; Pinto et al., 2015; Yeung et al., 2013a; Yeung et al., 2013b). Questionnaires and writing specific scales were also used, such as the motivational orientation writing scale, although less frequently than the measures above (Ahmed et al., 2014; Dunsmuir & Blatchford, 2004; Hamilton et al., 2013; Hooper et al., 2011; Torrance et al., 2000).

A few measures were used in only one or two studies: transcription skills, word length, writing a series of noun-adjective pairs, writing of sentences, writing of short words, early writing by hand and by keyboard (Beers & Nagy, 2011) and early writing concepts (Hooper et al., 2010), emergent literacy abilities, standardized assessments, and syntactic skills (Yeung et al., 2013b; Yeung et al., 2020).

We also investigated the types and number of scales and measures used to collect data by studies in our corpus. Scales or test batteries are one or more tests that aim to assess a particular factor of a person’s functioning (Frey, 2018). Our findings show that the majority of studies in our corpus (30/56) relied upon some type of pre-designed scale or measure, and often more than one at a time. The 30 studies in our corpus which used scales and measures used 51 different scales a total of 109 times for an average of 3.6 scales per study which used them. The most frequently used scales (and the number of studies which used the scales in parenthesis) were the WIAT (7), the WJ (5), Raven’s Colored Progressive Matrices (6), the WISC-IV (3), the Peabody Picture Vocabulary Test III (2), CELF3 (2), and the TOWRE PDE (2). (See Appendix A for the entire list

of scales). These top 7 scales were used 34 different times across the 30 studies in our corpus which used scales and measures. There were 42 other scales used in the studies in our corpus of longitudinal writing research.

Our final corpus of studies included 2 mixed method studies (Skibbe et al., 2013; Svensson, 2018). These studies measured lexicalized figurativeness and genuine figurativeness, with number of units (clauses) and percentage of these measures in the narrative texts (Svensson, 2018), and semistructured writing tasks for both parents and children (Skibbe et al., 2013; Svensson, 2018).

Data Collection in Qualitative Studies. In the qualitative studies in our corpus, the most used method of data collection was ethnography (Chapman, 2002; Compton-Lilly, 2014; Elf, 2016; Roozen, 2008; Smith & Prior, 2020), followed by document and artifact collection (Beaufort, 2004; Chen, 2018; Lammers & Marsh, 2018; Woodward-Kron, 2009), descriptive research (Johnson & Kruse, 2012; Sommers & Saltz, 2004; Turnbull et al., 2011), interviews (Driscoll & Powell, 2016; Lunsford et al., 2013), correlational research (Haswell, 2000), and a category we referred to as other (Myhill & Jones, 2007).

Many of these qualitative studies were unclear about the specific number of points of measure, but nonetheless collected data many times across the time-frame of their study (7/14), followed by three points of measure (3/14), five points of measure (1/14), four points of measure (1/14), two points of measure (1/14), and N/A (1/14). These qualitative studies ranged from 6 months to 30 years in duration, averaging 6.53 years of study. The qualitative studies in our corpus rarely reported the ages of participants (12/14), and when they did, they simply gave age ranges (2/14), instead choosing to defer to grade level. Most qualitative studies took place in the secondary grades (8/14), followed by undergraduate (2/14), primary (2/14), middle (1/14), and preschool/kindergarten (1/14). Ranging from 1 participant to 481, the qualitative studies in our corpus averaged 56.28 participants per study. Four of the qualitative studies in our corpus focused on students as the object of study, three focused on texts, and the remaining seven (50 percent) used both texts and students.

Across all qualitative studies, most focused on school-based assigned writing (7/14), or interviews (10/14). For studies which only focused on texts (3/14), two used assigned writing (Chen, 2018; Lammers & March, 2018) and one used extracurricular writing (Turnbull et al., 2011). For studies which only focused on students (Beaufort, 2004; Driscoll & Powell, 2016; Lunsford et al., 2013; Myhill & Jones, 2007), all used interviews and one used observations and interviews (Myhill & Jones, 2007). For studies which focused on both students and texts (7/14), all but one were a combination of assigned texts and interviews, with two of those studies also allowing for extracurricular writing (Roozen, 2008; Sommers & Saltz, 2004).

Data Analysis. We coded data analysis methods for each study in our corpus (the complete coding guide can be viewed online at https://osf.io/tjyu2/?view_only=72272c4f124b4b00bbd41667798edc76). Across all of the studies, predictive, causal, and inferential data analysis methods (i.e., quantitative methods of analysis) remain the most prominent data analysis methods in longitudinal writing research over the past twenty-one years. Among these methods, predictive analysis methods such as correlations (e.g., Hooper et al., 2010) and regressions (e.g., Pinto et al., 2015) were used in 32.1 percent of the articles in our corpus. Qualitative analysis methods were used in 14.3 percent and causal in 8.9 percent of the articles in our corpus. Inferential analysis methods such as ANOVA and t-tests were used in 5.4 percent of studies, the same as Grounded Theory (5.4 percent), and Exploratory (5.4 percent). Studies which used data analysis techniques that did not fit our coding schema were coded as Other (19.6 percent) and mostly involved more than one type of analysis. For quantitative studies specifically, 18 used predictive data analysis, ten used more than one type of analysis, four used causal analysis, three inferential, one descriptive and one exploratory.

Overall, research aims/questions and hypotheses of the quantitative studies in our corpus matched their data analysis methods. For example, studies that aimed to *predict* an outcome, to observe *effects*, or to investigate *relationships* (Sykes, 1993) tended to choose predictive data analysis options, such as regression (see Cordeiro et al., 2020) and CFA and SEM (see Kim & Park, 2019), while studies that used *how* questions, for example “to analyze *how* children are prepared for learning to write and *how* this skill is developed” (see Kuzeva et al., 2015) or *differences* between groups (see Silva et al., 2010), chose data collection methods that lead to the use of inferential data analysis methods such as ANOVA and t-tests (see Beers & Nagy, 2011; Oppenheimer et al., 2017

On the other hand, while qualitative studies asked similar questions to quantitative studies, such as *how* questions and to investigate relationships between writing and other aspects, it is, however, clear that these questions were broader and more intent upon generating new knowledge on a particular topic rather than identifying predictors. For example, the qualitative studies in our corpus which used ethnography used *what* and *how* questions (Chapman, 2002; Compton-Lilly, 2014; Elf, 2016; Roozen, 2008; Smith & Prior, 2020).

Quantitative Data Analysis. Quantitative studies used the following methods of data analysis, with regression being the most used method: ANOVA (Beers & Nagy, 2011; Bigozzi & Vettori, 2016; Levin et al., 2001) and t-tests (Frost, 2001; Oppenheimer et al., 2017) or non-parametric alternatives; correlation (Aram et al., 2013; Dunsmuir & Blatchford, 2004; Frost, 2001; Hooper et al., 2010; Levin et al., 2001; Niedo et al., 2014; Tong & McBride, 2016; Torrance et al., 2000),

regression, such as bivariate (Bigozzi & Vettori, 2016), multiple regression (Pinto et al., 2015), hierarchical regression (Yan et al., 2012) and logistic and stepwise regression (Pinto et al., 2012) and likelihood-based mixed-effects regression (Rosário et al., 2017); descriptive analysis (Coker, 2006); cluster analysis (Torrance et al., 2000), confirmatory factor analyses (CFA; Kim & Park, 2019), structural equation modeling (SEM; Kim & Park, 2019), content analysis (Yan et al., 2012), exploratory factor analysis (EFA; Pinto et al., 2009), hierarchical linear modeling analysis (Coker, 2006); latent change score modeling (Ahmed et al., 2014); mean proportions (Cardoso-Martins et al., 2006), path analyses (Yeung et al., 2013) and finally relative percentage (Fischer & Koch, 2016).

Qualitative Data Analysis. Qualitative studies of longitudinal writing in our corpus mostly deferred to asking single or multiple questions to frame their inquiry (9/14), with four studies stating a research purpose, and one study stating a research hypothesis; most of the question-driven studies posed a single question (7/9). Questions ranged from “What role do writing performances (particularly outside the classroom) play in early college students’ development as writers?” (Sommers & Saltz, 2004), to “How do changes in textual features over time (taken from writing in all classes, not just English/LA) demonstrate emergent genres (and thereby increasingly complex writing)?” (Chapman, 2002). Most of the studies of longitudinal writing which utilized qualitative methods in our corpus took place in the United States (10/14), with two in Canada, one in the UK, and one in Denmark. In all of these studies English was the first language. Only two qualitative studies in our corpus reported the type of community the schools were based in, and both were urban.

The frequency of data analysis methods used in qualitative longitudinal writing research were Grounded Theory (8/14), followed by Qualitative Content (4/14), Discourse Analysis (1/14), Exploratory studies (1/14), and Other (1/14). Five of the studies used some form of coding to perform their analysis (Beaufort, 2002; Myhill & Jones, 2007; Turnbull et al., 2011; Compton-Lilly, 2014; Driscoll & Powell, 2016). Across all studies, most focused on assigned writing (7/14), or interviews (10/14). For studies which only focused on texts (3/14), two used assigned writing and one used extracurricular writing. For studies which only focused on students (4/14), all used interviews and one used observations and interviews. For studies which focused on both students and texts (7/14), all but one were a combination of assigned texts and interviews, with two of those studies also allowing for extracurricular writing (Sommers & Saltz, 2004; Roozen, 2008), and one study strictly based on interviews (Compton-Lilly, 2014). The mixed method studies in our corpus used categorization of narrative texts and its method of analysis of the data (Svensson, 2018) and t-tests, regression, and observation (Skibble et al., 2013).

PUBLICATION OF LONGITUDINAL STUDIES OF WRITING

As a part of our coding schema, we looked at the general characteristics of the publication of each study by coding the authors, journals, and years of publication for each study in order to understand the broader ecosystem of longitudinal studies of writing, which includes the venues in which these studies become public. Our findings show that longitudinal research is on the rise; from 2000-2009 there were 17 articles on longitudinal writing, but from 2010-2020 there were 39 articles on longitudinal writing, meaning that there were 1.7 articles per year for the first decade of this century, but 3.9 articles per year in the most recent decade.

In examining which publication venues had the most longitudinal studies of writing, *Reading and Writing: An Interdisciplinary Journal* published the most with 12 total articles. Five other journals published more than one longitudinal study of writing: *The Journal of Educational Psychology* (5), *Learning and Individual Differences* (3), *European Journal of Psychology of Education* (3), and *Written Communication* (3). *The Journal of Child Language* and *British Journal of Educational Psychology* each had 2; no other journal in our corpus published more than one study.

In terms of authorship, longitudinal writing research proved to be very collaborative, with 49/56 articles written by more than one author with an overall average of 2.8 authors per article. Notably, five of the seven single-authored articles were single-subject case studies. Ten authors had more than one article in our corpus (Giulia Vettori, Claudio Vezzani, Giuliana Pinto, Lucia Bigozzi, Stephen R. Hooper, Teresa Limpo, Young-Suk Grace Kim, Virginia Berninger, Robert D. Abbott, and Dorit Aram), with one of those authors, Lucia Bigozzi, publishing five articles.

FUNDING OF LONGITUDINAL STUDIES OF WRITING

Finally, we coded studies for whether they reported funding or not, and if so, the agency which funded the study. We cross-referenced this information to the other coding categories in our study (type of study-qualitative/quantitative or mixed methods; number of participants; points of measure; study duration; study setting) in order to better describe the state of funding in longitudinal writing research. Our results show that 37.5 percent of studies in our corpus were funded, with 35.7 percent explicitly denoting that they received no funding and 26.8 percent of studies not reporting one way or the other. Reported sources of funding appear to be associated most frequently with quantitative studies. Further, most of the funding appeared to be linked to a single project and came from governmental agencies. And, while we identified a number of private funding agencies which granted funding for longitudinal studies, only the Spencer Foundation funded more than one study in our corpus.

Funded studies averaged more participants and longer periods of study, but not more points of measure. Funded studies averaged 163.05 participants per study, while non-funded studies averaged 107.593 participants per study. Among the single-subject studies only one was funded (Compton-Lilly, 2014). Funded studies averaged 3.33 years while unfunded and unclear studies averaged 3.23 years per study. Both funded and unfunded studies took place mostly in public school settings, and there was only one funded study of longitudinal writing in a private school setting.

DISCUSSION

Our study shows that during the past twenty-one years of L1 longitudinal writing research, a wide variety of both qualitative and quantitative studies have been carried out. In addition to the wide range of research methods (in both data collection and analysis) being employed in longitudinal writing research, the number of L1 longitudinal studies of writing appears to be increasing, as represented in our systematic review: from 2000-2009 we identified 17 longitudinal writing research studies; while from 2010-2020 there were 39 studies (1.7 articles per year for the first decade of this century; 3.9 articles per year in the most recent decade). Our review, of course, is not exhaustive, as our inclusion criteria did not include books or book chapters, which we did not consider as peer reviewed, even though these studies, such as Gere (2019) and Krogh and Jakobsen (2019) make valuable contributions to longitudinal writing research and our understanding of writing development. Nonetheless, the increase in L1 studies of longitudinal writing suggests that now is an opportune time to continue developing capacity to carry out longitudinal writing research at scale.

Within the broader range of studies, from large *n* quantitative to single subject qualitative case studies, we also see a number of studies with similar research methodologies and methods forming into common lines of inquiry; for example, studies that use observations and interviews to gauge parental influence on early literacy development, studies that use literacy measures to predict student performance in and across the early grades, studies that use textual analysis of curricular writing to investigate how collegiate students acclimate to the demands of higher education, etc. Although currently these studies do not lend themselves towards aggregation or replication, these clusters of work do point towards the advancement of longitudinal writing research methods and the refinement of research designs.

From another viewpoint however, the overall number of L1 longitudinal studies of writing remains relatively small, and the field itself is still at an early stage in which replication studies and even aggregation (let alone meta-analyses), remain

out of reach. In this regard, we feel strongly that while we need more depth within these clusters of work, (i.e., we need more longitudinal studies of writing that build upon previous work), we also need to encourage an even wider range of research methods. As Bazerman noted, “there is still too much to discover about our multidimensional subject to limit what we are looking for and the way we might be looking” (Bazerman, personal communication, May 1st, 2020).

One important step to take in regard to furthering longitudinal studies of writing is to draw out with greater intentionality the nature of each of these research lines so that future researchers can more easily identify scales and measures, units of analysis, objects of study, and tools for data collection and analysis. Further, it would be beneficial to other researchers to provide more methodological transparency regarding the logistics, barriers, false steps, and the nature of the collaborative activities involved in order to continue to build our collective research capacity and to carry out more sophisticated longitudinal studies of writing that can support the empirical grounding of lifespan approaches to writing.

The wide range of L1 longitudinal writing research reflects the strong interdisciplinary nature of the research communities which study writing. Thus, in addition to a greater awareness of the range of research methods being used, we also see an opportunity in this interdisciplinarity for more mixed methods studies, i.e., the integration of quantitative and qualitative studies of writing. In our view, substantive advances in lifespan perspectives of writing development and longitudinal writing research will depend on these kinds of cross-disciplinary, multidisciplinary, and ultimately transdisciplinary studies of writing being carried out in international contexts around the world.

While longitudinal studies of writing offer a great deal to the work on lifespan writing development, lifespan approaches to writing development also have strong potential to move longitudinal writing research forward across several important dimensions. First of all, given the complexity of writing and the wide breadth of research methods being employed, lifespan perspectives can serve to bring together these disparate lines of inquiry into more coherent and productive models that take into account a variety of writing related influences and outcomes across longer time sequences and especially across major transition points (e.g., preschool to school, elementary to secondary, high school to college, and college to the workplace). While more will still need to be done to provide a full lifespan perspective, this weaving together of insights from across studies can move the entire field forward and away from the myopia of repeatedly focusing on particular areas (such as first year composition in the US) to see writing development across the fuller lifespan and to consider the problems and questions associated with teaching, learning, curriculum, assessment, professional development, etc. in ways that take into account a broader set of socio-cultural experiences and cognitive processes.

CHALLENGES

We encountered three main difficulties while conducting our review. The first challenge was identifying studies in which writing was the central object of inquiry. Determining if writing was the focus of a study proved especially difficult in longitudinal studies in the early grades because the components of literacy are so tightly woven together and there are so many studies of literacy that include elements and sub elements of both reading and writing. Additionally, there exists a wide variety of research studies that use writing as a data source or comparison point among other literacy measures rather than being a study of writing alone (Abbott et al., 2010, p.281).

The second challenge we faced was addressing the potential for conflation in measuring writing development with outcomes of specific writing curricula. No writing curriculum can be created without taking development into account as writing often develops in accordance with curriculum. However, following Bazerman (2018), we found that distinguishing between what counts as writing development and specific curricular outcomes in research is critical because the knowledge, skills, dispositions, and abilities associated with writing extend beyond any particular intervention to include the fullness of a person's literate life (see, for example, Dyson, 2003). This meant looking very carefully at each study's outcomes and the types of curricular interventions that were used (if any) to ensure that the researchers were looking beyond an immediate learning gain to broader issues of writing development over time.

The final challenge we worked through, (which is related to the issue of conflation above) was precisely defining the qualities of longitudinal writing research, and more specifically defining how long a study needed to be to count as longitudinal. We built our screening guide using two points of measurement as our baseline definition of longitudinal. However, we kept in mind the importance of the relationship between development and time, recognizing that development can happen in short periods of time during periods of transition (Bazerman, 2018). However, because the term "periods of transition" is subjective in many cases, we also had long discussions on what can be considered as a longitudinal study in cases of studies lasting for shorter periods of time.

The values which shaped the selection of studies for our corpus, though explicitly and transparently reported in this review, led to a somewhat strict definition of what counted as longitudinal writing research that was contingent on our own specific goals for this inquiry. Much discussion remains to be had by the field regarding the wide ranging and varied conceptions of what constitutes or might constitute longitudinal writing research. There exists a vast spectrum of work which is, or could be considered to be, longitudinal writing research well

beyond what is represented in our corpus. Where to draw the line as to what counts has serious implications for reviews like the one reported on here and the generalizations and principles that can be drawn from systematic research reviews. Future research should make clear the specific definition/s of “longitudinal writing” on which the research is predicated.

Throughout our screening and coding processes, a vast majority of exclusions were made on the basis of these distinctions (between curricular outcomes and writing development and writing as the central focus). Further, our review focused explicitly on L1 studies in a K-University context. As a great deal of work on longitudinal writing research takes place outside of school contexts, and as an even larger contingent of this work takes place in L2 contexts, this study reports only a partial view of the work taking place in the broader field of longitudinal research on writing. Future reviews should seek to expand reviews of longitudinal writing research to increasingly broader contexts, and in doing so allow for more fulsome, sophisticated, and nuanced perspectives on longitudinal research of writing and its place in and implications for our understanding of lifespan development.

IMPROVING LONGITUDINAL WRITING RESEARCH

Our review demonstrates that researchers have an array of research methods and designs to choose from when designing a longitudinal study of writing. However, the wide range of methodological options available to researchers in the field does not necessarily translate into those options of research being equally available to everyone. Disciplinary training and epistemology (especially along the qualitative and quantitative divide) dictate many research design decisions from the beginning of a study, as issues of analytic method and measure often shape the direction which a longitudinal study of writing will take. Thus, one’s disciplinary training, though a strength, can also be a limiting factor in addressing nuanced research problems and questions. These considerations of epistemology and disciplinarity and their accompanying methodological choices provide another warrant for encouraging mixed-methods research and cross-disciplinary collaboration in longitudinal writing research. At the field level it will be difficult to build out a robust model of lifespan writing development without a great deal of cross disciplinary collaboration and the richer repertoire of methodological capacities which such collaboration can bring.

For example, our review shows that lines of inquiry can be framed in a variety of ways, (beyond a single research question or even without a research question altogether). An awareness of this wider range of methodological choices at all levels of research is important for making the best possible research design decisions, especially in ways that establish a goodness of fit between problems

under investigation, theoretical frameworks, the existing literature, data sources and collection methods, and methods of analysis, and which can best address relevant gaps in the knowledge base.

In order for future researchers to more fully take into account previous work on longitudinal development in writing, future studies will need to advance the clarity and comprehensiveness of their data reporting. Currently, our results show a lack of reporting in some key areas that limit the ability of writing researchers to build on each other's work. For example, every study in our corpus limited gender reporting to male and female. In a world that increasingly values inclusivity and gender fluidity, limiting gender reporting to strictly male and female runs the risk of washing out key differences in participant populations which might further inform longitudinal writing research across the lifespan in important ways.

Reporting issues present an even greater problem when it comes to age ranges, which has serious implications for those interested in constructing lifespan perspectives on writing development. Specifically, many of the studies in our corpus did not report ages or age ranges, often deferring to grade level instead (or not reporting on age or grade at all). Further compounding these reporting issues is the incompatibility of grade levels across contexts of international schooling. Not identifying ages is problematic because longitudinal writing research is time-based; thus, the reporting of participant age ranges is critical for both the tracing of changes across time and the generalizability (and replicability) of results. We strongly encourage all longitudinal writing researchers to seek to identify and report both grade levels and age ranges in their studies.

Our focus in this review was on L1 longitudinal studies of writing in the school years. In relationship to lifespan perspectives, this leaves out important areas. Our larger study, however, not reported on here, includes L2 studies, and studies that extend before and beyond school, as a lifespan view of writing demands widening our horizons beyond the schooling years. However, we acknowledge that the development of writing ability in school age children certainly deserves the attention it has received. In this review, for example, the oldest participant in any of our studies was 30 years old, which does not yet reflect the full potential of lifespan development research. In order to fully build out the lifespan view, researchers will need to account for writing development well beyond the schooling years. In doing so, longitudinal writing research will contribute to a more sophisticated and nuanced understanding of lifespan writing development and cultivate a more accurate knowledge base from which to think about the types of interventions and pedagogies that can advance learning and writing at different ages and stages throughout the lifespan.

Lifespan perspectives on writing development would also benefit greatly from a longitudinal writing research base which produced aggregable results drawn

from replicable studies. Due to the great variation in human conditions, interactions, settings, personalities, and other complex contributing variables, it is especially important for longitudinal writing researchers to attend to concretizing methodologies with enough centrifugal force to compare results drawn from a wide range of texts which all aim to convey “a unique message between a unique writer to a particular unique audience” (Bazerman, personal communication, May 1st, 2020). It is through the replication of methods in different settings that comparative results can be found that will be able to illustrate the variations and changes among phenomena that will illuminate differences among participants’ development. However, given the relatively young state of lifespan perspectives on writing and longitudinal writing research, at this time, more attention should probably be given to expanding writing research programs and lines rather than narrowing our field of view.

Looking forward, we must continue to generate hypotheses concerning the many interconnected variables and factors which contribute to and shape writing development so that “they can be tested for correlations, their relative importance in contributing to development, as well as the varying degrees in which individual elements contribute to the varying dimensions of writing development” (Rogers, 2009). Finally, we must engage more deeply in model building and hypothesis testing in order to further articulate the impact of specific factors and direction of pedagogical interventions related to the development of writers and writing abilities for people at all points across the lifespan.

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>
- Ahmed, Y., Wagner, R. K., & Lopez, D. (2014). Developmental relations between reading and writing at the word, sentence, and text levels: A latent change score analysis. *Journal of Educational Psychology, 106*(2), 419–434. <https://doi.org/10.1037/a0035692>
- Almalki, S. (2016). Integrating quantitative and qualitative data in mixed methods research—Challenges and benefits. *Journal of Education and Learning, 5*(3), 288. <https://doi.org/10.5539/jel.v5n3p288>
- Aram, D., & Levin, I. (2004). The role of maternal mediation of writing to kindergartners in promoting literacy in school: A longitudinal perspective. *Reading and Writing, 17*(4), 387–409. <https://doi.org/10.1023/B:READ.0000032665.14437.e0>
- Aram, D., Korat, O., & Hassunah-Arafat, S. (2013). *The contribution of early home literacy activities to first grade reading and writing achievements in Arabic*. Springer. <https://doi.org/10.1007/s11145-013-9430-y>

- Bazerman, C. (2018). Lifespan longitudinal studies of writing development: A heuristic for an impossible dream. In C. Bazerman, A. N. Applebee, V. W. 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* (pp. 326-365). National Council of Teachers of English. <https://wac.colostate.edu/books/ncte/lifespan-writing/>
- Beaufort, A. (2004). Developmental gains of a history major: A case for building a theory of disciplinary writing expertise. *Research in the Teaching of English*, 39(2), 136–185.
- Beers, S. F., & Nagy, W. E. (2011). Writing development in four genres from grades three to seven: Syntactic complexity and genre differentiation. *Reading and Writing*, 24(2), 183–202. <https://doi.org/10.1007/s11145-010-9264-9>
- Berninger, V., Mizokawa, D., & Bragg, R. (1991). Theory-based diagnosis and remediation of writing disabilities. *Journal of School Psychology*, 29, 57–79. [https://doi.org/10.1016/0022-4405\(91\)90016-K](https://doi.org/10.1016/0022-4405(91)90016-K)
- Bigozzi, L., & Vettori, G. (2016). To tell a story, to write it: Developmental patterns of narrative skills from preschool to first grade. *European Journal of Psychology of Education*, 31(4), 461–477. <https://doi.org/10.1007/s10212-015-0273-6>
- Cardoso-Martins, C., Corrêa, M. F., Lemos, L. S., & Napoleão, R. F. (2006). Is there a syllabic stage in spelling development? Evidence from Portuguese-speaking children. *Journal of Educational Psychology*, 98(3), 628–641. <https://doi.org/10.1037/0022-0663.98.3.628>
- Caruana, E. J., Roman, M., Hernández-Sánchez, J., & Solli, P. (2015). Longitudinal studies. *Journal of Thoracic Disease*, 7(11), E537–E540. <https://doi.org/10.3978/j.issn.2072-1439.2015.10.63>
- Chapman, M. (2002). A longitudinal case study of curriculum genres, K-3. *Canadian Journal of Education / Revue Canadienne de l'éducation*, 27(1), 21–44. <https://doi.org/10.2307/1602186>
- Chen, V. (2018). Seeing the world through words: A student writer's journey toward developing her own voice. *Journal for Learning through the Arts: A Research Journal on Arts Integration in Schools and Communities*, 13(1). <https://doi.org/10.21977/D913135150>
- Coker, D. (2006). Impact of first-grade factors on the growth and outcomes of urban schoolchildren's primary-grade writing. *Journal of Educational Psychology*, 98(3), 471–488. <https://doi.org/10.1037/0022-0663.98.3.471>
- Compton-Lilly, C. (2014). The development of writing habitus: A ten-year case study of a young writer. *Written Communication*, 31(4), 371–403. <https://doi.org/10.1177/0741088314549539>
- Cordeiro, C., Limpo, T., Olive, T., & Castro, S. L. (2020). Do executive functions contribute to writing quality in beginning writers? A longitudinal study with second graders. *Reading and Writing*, 33(4), 813–833. <https://doi.org/10.1007/s11145-019-09963-6>
- Dippre, R. J., & Phillips, T. (Eds.). (2020). *Approaches to lifespan writing research: Generating an actionable coherence*. The WAC Clearinghouse; University Press of Colorado. <https://doi.org/10.37514/PER-B.2020.1053>

- Doyle, L., Brady, A.-M., & Byrne, G. (2009). An overview of mixed methods research. *Journal of Research in Nursing, 14*(2), 175–185. <https://doi.org/10.1177/1744987108093962>
- Drijbooms, E., Groen, M. A., & Verhoeven, L. (2017). How executive functions predict development in syntactic complexity of narrative writing in the upper elementary grades. *Reading and Writing, 30*(1), 209–231. <https://doi.org/10.1007/s11145-016-9670-8>
- Driscoll, D. L., & Powell, R. (2016). States, traits, and dispositions: The impact of emotion on writing development and writing transfer across college courses and beyond. *Composition Forum, 34*. <https://eric.ed.gov/?id=EJ1113424>
- Dunsmuir, S., & Blatchford, P. (2004). Predictors of writing competence in 4- to 7-year-old children. *British Journal of Educational Psychology, 74*(3), 461–483. <https://doi.org/10.1348/0007099041552323>
- Dyson, A. H. (2003). *The brothers and sisters learn to write: Popular literacies in childhood and school culture*. Teachers College Press.
- Elf, N. (2017). Taught by bitter experience: A timescales analysis of Amalie's development of writer identity. In T. Cremin & T. Locke (Eds.), *Writer Identity and the Teaching and Learning of Writing*. Routledge. <https://doi.org/10.4324/9781315669373>
- Firestone, W. (1987). Meaning in method: The rhetoric of quantitative and qualitative research. *Educational Researcher, 16*(7), 16–21. <https://doi.org/10.3102/0013189X016007016>
- Fischer, J. P., & Koch, A. M. (2016). Mirror writing in typically developing children: A first longitudinal study. *Cognitive Development, 38*, 114–124. <https://doi.org/10.1016/j.cogdev.2016.02.005>
- Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2012). *Applied longitudinal analysis*. John Wiley & Sons. <https://www.doi.org/10.1002/9781119513469>
- Frey, B. B. (Ed.). (2018). *The Sage encyclopedia of educational research, measurement, and evaluation*. Sage Publications. <https://doi.org/10.4135/9781506326139>
- Frost, J. (2001). Phonemic awareness, spontaneous writing, and reading and spelling development from a preventive perspective. *Reading and Writing, 14*(5), 487–513. <https://doi.org/10.1023/A:1011143002068>
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255–274. <https://doi.org/10.3102/01623737011003255>
- Hamilton, E. W., Nolen, S. B., & Abbott, R. D. (2013). Developing measures of motivational orientation to read and write: A longitudinal study. *Learning and Individual Differences, 28*. <https://doi.org/10.1016/j.lindif.2013.04.007>
- Hartley, J., & Chesworth, K. (2000). Qualitative and quantitative methods in research on essay writing: No one way. *Journal of Further and Higher Education, 24*(1), 15–24. <https://doi.org/10.1080/030987700112282>
- Haswell, R. H. (2000). Documenting improvement in college writing: A longitudinal approach. *Written Communication, 17*(3), 307–352. <https://doi.org/10.1177/0741088300017003001>

- Hedeker, D., & Gibbons, R. D. (2006). *Longitudinal data analysis*. Wiley-Interscience.
- Hooper, S. R., Costa, L. J., McBee, M., Anderson, K. L., Yerby, D. C., Knuth, S. B., & Childress, A. (2011). Concurrent and longitudinal neuropsychological contributors to written language expression in first and second grade students. *Reading and Writing, 24*(2), 221–252. <https://doi.org/10.1007/s11145-010-9263-x>
- Hooper, S. R., Roberts, J. E., Nelson, L., Zeisel, S., & Fannin, D. K. (2010). Preschool predictors of narrative writing skills in elementary school children. *School Psychology Quarterly, 25*(1), 1–12. <https://doi.org/10.1037/a0018329>
- Johnson, J. P., & Krase, E. (2012). Coming to learn: From first-year composition to writing in the disciplines. *Across the Disciplines, 9*(2), 1–22. <https://doi.org/10.37514/ATD-J.2012.9.2.02>
- Kim, Y. S. G., & Park, S. H. (2019). Unpacking pathways using the direct and indirect effects model of writing (DIEW) and the contributions of higher order cognitive skills to writing. *Reading and Writing, 32*(5), 1319–1343. <https://doi.org/10.1007/s11145-018-9913-y>
- Kim, Y. S., Al Otaiba, S., & Wanzek, J. (2015). Kindergarten predictors of third grade writing. *Learning and Individual Differences, 37*, 27–37. <https://doi.org/10.1016/j.lindif.2014.11.009>
- Krogh, E., & Jakobsen, K. S. (Eds.). (2019). *Understanding young people's writing development: identity, disciplinarity, and education*. Routledge.
- Kuzeva, O. A., Romanova, A. A., Korneev, A. A., & Akhutina, T. V. (2015). The dynamics of programming, control and the movement sequential organization skills as the basic components of writing. *Russian Education & Society, 57*(9), 757–782. <https://doi.org/10.1080/10609393.2015.1125707>
- Lammers, J. C., & Marsh, V. L. (2018). “A writer more than . . . a child”: A longitudinal study examining adolescent writer identity. *Written Communication, 35*(1), 89–114. <https://doi.org/10.1177/0741088317735835>
- Levin, I., Ravid, D., & Rapaport, S. (2001). Morphology and spelling among Hebrew-speaking children: From kindergarten to first grade. *Journal of Child Language, 28*(3), 741–772. <https://doi.org/10.1017/S0305000901004834>
- Lewis-Beck, M., Bryman, A. E., & Liao, T. F. (2004). *The Sage Encyclopedia of Social Science Research Methods*. Sage. <https://doi.org/10.4135/9781412950589>
- 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), 401–413. <https://doi.org/10.1037/a0031391>
- Lunsford, A. A., Fishman, J., & Liew, W. M. (2013). College writing, identification, and the production of intellectual property: Voices from the Stanford study of writing. *College English, 75*(5), 470–492.
- Mäki, 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), 643–672. <https://doi.org/10.1023/A:1012071514719>
- Myhill, D., & Jones, S. (2007). More than just error correction: Students' perspectives on their revision processes during writing. *Written Communication, 24*(4), 323–343. <https://doi.org/10.1177/0741088307305976>

- Niedo, J., Abbott, R. D., & Berninger, V. W. (2014). Predicting levels of reading and writing achievement in typically developing, English-speaking 2nd and 5th graders. *Learning and Individual Differences, 32*, 54–68. <https://doi.org/10.1016/j.lindif.2014.03.013>
- Oppenheimer, D., Zaromb, F., Pomerantz, J. R., Williams, J. C., & Park, Y. S. (2017). Improvement of writing skills during college: A multi-year cross-sectional and longitudinal study of undergraduate writing performance. *Assessing Writing, 32*, 12–27. <https://doi.org/10.1016/j.asw.2016.11.001>
- Pinto, G., Bigozzi, L., Gamannossi, B. A., & Vezzani, C. (2009). Emergent literacy and learning to write: A predictive model for the Italian language. *European Journal of Psychology of Education, 24*(1), 61. <https://doi.org/10.1007/BF03173475>
- Pinto, G., Bigozzi, L., Gamannossi, B. A., & Vezzani, C. (2012). Emergent literacy and early writing skills. *The Journal of Genetic Psychology, 173*(3), 330–354. <https://doi.org/10.1080/00221325.2011.609848>
- Pinto, G., Tarchi, C., & Bigozzi, L. (2015). The relationship between oral and written narratives: A three-year longitudinal study of narrative cohesion, coherence, and structure. *The British Journal of Educational Psychology, 85*(4), 551–569. <https://doi.org/10.1111/bjep.12091>
- Rogers, P. (2009). The contributions of North American longitudinal studies of writing in higher education to our understanding of writing development. In C. Bazerman, R. Krut, K. Lunsford, S. McLeod, S. Null, P. Rogers, & A. Stansell (Eds.), *Traditions of writing research* (pp. 337–389). Routledge.
- Roozen, K. (2008). Journalism, poetry, stand-up comedy, and academic literacy: Mapping the interplay of curricular and extracurricular literate activities. *Journal of Basic Writing, 27*(1), 5–34. <https://doi.org/10.37514/JPBW-J.2008.27.1.02>
- Rosário, P., Högemann, J., Núñez, J. C., Vallejo, G., Cunha, J., Oliveira, V., Fuentes, S., & Rodrigues, C. (2017). Writing week-journals to improve the writing quality of fourth-graders' compositions. *Reading and Writing, 30*(5), 1009–1032. <https://doi.org/10.1007/s11145-016-9710-4>
- Shulruf, B., Hattie, J., & Dixon, R. (2007). Development of a new measurement tool for individualism and collectivism. *Journal of Psychoeducational Assessment, 25*(4), 385–401. <https://doi.org/10.1177/0734282906298992>
- Skibbe, L. E., Bindman, S. W., Hindman, A. H., Aram, D., & Morrison, F. J. (2013). Longitudinal relations between parental writing support and preschoolers' language and literacy skills. *Reading Research Quarterly, 48*(4), 387–401. <https://doi.org/10.1002/rrq.55>
- Smith, A., & Prior, P. (2020). A Flat CHAT perspective on transliteracies development. *Learning, Culture and Social Interaction, 24*. <https://doi.org/10.1016/j.lcsi.2019.01.001>
- Sommers, N., & Saltz, L. (2004). The novice as expert: Writing the freshman year. *College Composition and Communication, 56*(1), 124–149. <https://doi.org/10.2307/4140684>
- Sullivan, G. M., & Sargeant, J. (2011). Qualities of qualitative research: Part I. *Journal of Graduate Medical Education, 3*(4), 449–452. <https://doi.org/10.4300/JGME-D-11-00221.1>

- Svensson, B. (2018). The development of figurative competence in narrative writing: A longitudinal case study. *International Journal of Language Studies*, 12, 75–102.
- Tong, X., & McBride, C. (2016). Reading comprehension mediates the relationship between syntactic awareness and writing composition in children: A longitudinal study. *Journal of Psycholinguistic Research*, 45(6), 1265–1285. <https://doi.org/10.1007/s10936-015-9401-3>
- Torrance, M., Thomas, G. V., & Robinson, E. J. (2000). Individual differences in undergraduate essay-writing strategies: A longitudinal study. *Higher Education*, 39(2), 181–200.
- Treiman, R., Pollo, T. C., Cardoso-Martins, C., & Kessler, B. (2013). Do young children spell words syllabically? Evidence from learners of Brazilian Portuguese. *Journal of Experimental Child Psychology*, 116(4), 873–890. <https://doi.org/10.1016/j.jecp.2013.08.002>
- Turnbull, K., Deacon, S. H., & Kay-Raining Bird, E. (2011). Mastering inflectional suffixes: A longitudinal study of beginning writers' spellings. *Journal of Child Language*, 38(3), 533–553. <https://doi.org/10.1017/S030500091000022X>
- Woodward-Kron, R. (2009). “This means that . . .”: A linguistic perspective of writing and learning in a discipline. *Journal of English for Academic Purposes*, 8(3), 165–179. <https://doi.org/10.1016/j.jeap.2009.07.002>
- Yan, C. M. W., McBride-Chang, C., Wagner, R. K., Zhang, J., Wong, A. M. Y., & Shu, H. (2012). Writing quality in Chinese children: Speed and fluency matter. *Reading and Writing*, 25(7), 1499–1521. <https://doi.org/10.1007/s11145-011-9330-y>
- Yeung, P., Ho, C. S., Chan, D. W., & Chung, K. K. (2013a). Modeling the relationships between cognitive-linguistic skills and writing in Chinese among elementary grades students. *Reading and Writing*, 26(7), 1195–1221. <https://doi.org/10.1007/s11145-012-9411-6>
- Yeung, P., Ho, C. S., Chan, D. W., & Chung, K. K. (2020). Longitudinal relationships between syntactic skills and Chinese written composition in Grades 3 to 6. *Journal of Research in Reading*, 43(2), 201–228. <https://doi.org/10.1111/1467-9817.12298>
- Yeung, P. S., Ho, C. S. H., Wong, Y. K., Chan, D. W. O., Chung, K. K. H., & Lo, L. Y. (2013b). Longitudinal predictors of Chinese word reading and spelling among elementary grade students. *Applied Psycholinguistics*, 34(6), 1245–1277. <https://doi.org/10.1017/S0142716412000239>
- Zeanah, C. H., Boris, N. W., & Larrieu, J. A. (1997). Infant development and developmental risk: A review of the past 10 years. *Journal of the American Academy of Child & Adolescent Psychiatry*, 36(2), 165–178. <https://doi.org/10.1097/00004583-199702000-00007>

APPENDIX A. SCALES IN LONGITUDINAL STUDIES

All Scales Used by the Studies in Chapters 9 and 10	
PAL-II	ORF 1
CTOPP	ORF2
PPVT-4	Peabody Picture Vocabulary Test III
CREVT-2	HKT-SpLD
VIGIL	WUR
WJ-III	Raven's Colored Progressive Matrices
WRAML-2	Big-Small Stroop-like task
WISC-IV	WISC-III
WIAT-II	BANC
WorkingMem z-score	ERRNI
ReceptiveLang z-score	Systematic Screen of Handwriting Difficulties
Phonemes z-score	PI-dictee
ERA	PPVT-III-NL
Reading Interest Orientation	Tea-Ch
Reading Mastery Orientation	Sky Search
Writing Avoidance	LDST
Ego Orientation	D-KEFS-Letter
TOWRE PDE	LDST
WJ Passage Comp	D-KEFS-Letter Fluency
D-KEFS-TMT	GRE Issue Task
PPST	TOEFL
Kaufman IQ Test	Test of Word Reading Efficacy
SWAN	CTOPP
PAL	UW
WPPSI	BAS
CELF3	KTI