

Final Report:
A Comparison of Audio Observation and In-Person Observation of Teacher Behavior
Submitted by Dr. Lisa Didion, Assistant Professor of Special Education
Department of Teaching and Learning
University of Iowa

The purpose of this project was to test the validity of an audio observation tool by comparing audio and in-person observations of teacher behavior. However, due to the COVID-19 pandemic and related due to school closures, I was unable to carry out the project as outlined in the grant. With permission by H. Dee Hoover, I worked with the graduate research assistant (GA) to complete the data analysis of the pilot study of the audio observation measurement tool and begin a new project examining how teacher knowledge, skills, practices, and beliefs are measured in research related to teacher professional development. In the following sections, I will briefly describe these projects, the GA's responsibilities, and their results.

Audio Observation Measurement Tool Pilot Study. The purpose of this study was to determine the most reliable and efficient method of sampling teacher behaviors through audio recordings. I addressed two research questions: (1) To what extent is the use of audio recordings a reliable method to observe instructional behavior?; and (2) Do rates of instructional behaviors in the first 20 min of a lesson predict the rates of instructional behaviors for the full lesson?

The audio observation tool measured six evidence-based instructional practices (Archer & Hughes, 2011). The six behaviors included: (a) behavior specific praise statements; (b) praise statements; (c) behavior specific corrective feedback; (d) opportunity to respond to instruction; (e) background knowledge or connection to previously learned concepts; and (f) real-word or interdisciplinary connections. To represent the frequency of each behavior, partial interval recording methods were used. Each audio was observed in 1 min units. If a behavior was observed at any time within the 1-min unit, trained coders indicated it was heard by putting the

code in the corresponding time cell in a Microsoft Excel workbook. If a behavior was heard more than one time in a minute, it was only coded once per cell. More than one behavior code could be given within each 1-min cell when multiple teacher behaviors were observed. Once the coder listened to the entire audio file, coders counted the total number of each behavior throughout the audio file for a frequency score. The GA entered the data into a master excel spreadsheet.

Fleiss's kappa, an extension of Cohen's kappa, measured interrater reliability (Fleiss, 1971). We concluded that the tool had moderate to good reliability with kappas ranging from 0.45 to 0.80. In addition, to determine the efficiency of the tool, we explored whether 20 min of audio were predictive and related to the full audio sample. Based on regression analyses controlling for entire lesson length and teacher, it was reasoned that the first 20 min of a lesson is predictive of the rates of behaviors observed in a full lesson. Further, there were large, positive relationships between rates of the behaviors in the first 20 min and the full lesson, with correlations ranging between 0.51 to 0.88. Taken together, we concluded that using direct observation methods of audio recorded lessons holds promise to be both a reliable and efficient method to observe teacher's use of behaviors related to explicit instruction. This manuscript is submitted for publication at a peer-reviewed journal related to measurement. This data was presented at two virtual conferences in the Spring of 2021: the Pacific Coast Research Conference and the Council for Exceptional Children.

The Impact of Professional Development on Teacher Outcomes: A Meta-Analysis.

Professional development (PD) is the most effective route to educate teachers on current research-based practices. Research is evolving and teachers that were sufficiently trained previously may not have the knowledge to implement current, effective, research-based methods

to lessen achievement gaps (Connor et al., 2014). With estimates suggesting 5-10% of teachers' time is spent in PD (Gulamhussein, 2013), it is problematic that there is no conclusive evidence regarding what makes PD effective for teacher learning.

There is a breadth of research on PD, yet questions remain. Meta-analyses indicate student outcomes are significantly improved after teachers participate in PD but it is undecided what variables related to PD elicit the change (e.g., Blank & de las Alas, 2009; Didion et al. 2020; Kennedy, 1998; Yoon et al. 2007). Intensity, relevance, and collaborative learning processes (i.e., active and collective participation) are identified features of high-quality PD (National Center for Education Statistics, 200; Desimone, 2009). Heterogeneity of PD design complicate analyses of the degree these features result in teacher and student learning.

Alternatively, other literature reviews have focused on specific components to draw conclusions related to effective PD. For instance, some research synthesized literature related to a specific PD format (i.e., coaching; see Kraft et al., 2018; professional learning communities, see Vangrieken et al., 2017), content area (e.g., data literacy, see Filderman et al., 2020; science, see van Direfl et al., 2012), or population (e.g., early childhood, see Zaslow et al., 2010). Conclusions suggest PD effectively strengthens teacher outcomes and impacts student achievement, but this finding is limited by study parameters. There are still questions about what factors are meaningful for efficacious PD, significantly impacting teacher behavior and, subsequently, student achievement. A new lens is required to comprehensively review PD literature, such as examining theories of teacher learning and how they are related to outcomes measured.

A comprehensive search of peer-reviewed research published between the earliest possible start date (1974) and 2021 resulted in 128 studies that met inclusion criteria. Study features, effect sizes, and quality indicators are currently being coded. The effect size index that

will be used for all outcome measures is Hedges' g (Hedges, 1983) and will be corrected for sample size bias. All eligible, independent effect sizes will be included from each study, resulting in some studies contributing multiple effect sizes when several reading outcomes are reported. To account for the statistical dependencies of these correlated effects, random effects robust standard estimation will be used (Hedges et al., 2010). Findings may shed new light on complex issues about what makes PD effective for various teacher outcomes and how they are measured. Discussion will be based around moderating effects of PD design and mediating effects of teacher outcomes on student outcomes. The GA's role was to screen potential abstracts for inclusion and code included articles. I anticipate the article will be submitted for publication in the spring of 2022. This data is accepted for presentation at the Pacific Coast Research Conference held in the spring of 2022.

References

- Archer, A., & Hughes, C. (2011). *Explicit Instruction: Effective and efficient teaching*. Guilford Publications.
- Blank, R., & de las Alas, N. (2009). Effects of teacher professional development on gains in student achievement. Report prepared for the Council of Chief State School Officers.
- Connor, C. M., Alberto, P. A., Compton, D. L., & O'Connor, R. E. (2014). Improving Reading Outcomes for Students with or at Risk for Reading Disabilities: A Synthesis of the Contributions from the Institute of Education Sciences Research Centers. *National Center for Special Education Research* (NCSER 2014-3000). U.S. Department of Education.

- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181-199.
- Didion, L., Toste J. R., Filderman, M. J. (2019). Teacher professional development and student reading achievement: A meta-analytic review of effects. *Journal of Research of Educational Effectiveness*, 13(1), 29-66.
- Didion, L. (2022, February). *The Impact of Professional Development on Teacher Outcomes: A Meta-Analysis*. Poster session presented at the Pacific Coast Research Conference, PCRC, San Diego, CA.
- Didion, L. & Filderman, M. J. (2021, February). *Preservice Teachers' Use of Explicit Instruction Practices: Evidence from Audio Observation*. Poster session presented at the Pacific Coast Research Conference, PCRC, virtual event
- Didion, L. & Filderman, M. J. (2021, March). *Efficiently and Accurately Using Audio Observation to Sample Teacher Behavior*. Presentation with Q&A at the Council for Exceptional Children, CEC, virtual event.
- Filderman, M. J., Toste, J. R., Didion, L. A., Peng, P., & Clemens, N. H. (2018). Data-based decision making in reading interventions: A synthesis and meta-analysis of the effects for struggling readers. *Journal of Special Education*, 52(3), 174-187.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378. <https://doi.org/10.1037/h0031619>
- Gulamhussein, A. (2013). Teaching the teachers: effective professional development in an era of high stakes accountability. <http://www.centerforpubliceducation.org/teachingtheteachers>
- Hedges, L. V. (1983). A random effects model for effect sizes. *Psychological Bulletin*, 93(2), 388.

- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research synthesis methods*, 1(1), 39-65.
- Kennedy, M. (1998). Form and Substance in Inservice Teacher Education. Research Monograph. National Institute for Science Education (No-13). U.S. Department of Special Education
- Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, 88(4), 547-588.
- Van Driel, J. H., Meirink, J. A., van Veen, K., & Zwart, R. C. (2012). Current trends and missing links in studies on teacher professional development in science education: A review of design features and quality of research. *Studies in Science Education*, 48(2), 129-160.
- Vangrieken, K., Meredith, C., Packer, T., & Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teaching and Teacher Education*, 61, 47-59.
- Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). Reviewing the Evidence on How Teacher Professional Development Affects Student Achievement. Issues & Answers. Regional Educational Laboratory Southwest (REL 2007-No. 033). U.S. Department of Special Education
- Zaslow, M., Tout, K., Halle, T., Whittaker, J. V., & Lavelle, B. (2010). Toward the Identification of Features of Effective Professional Development for Early Childhood Educators. Literature Review. *Office of Planning, Evaluation and Policy Development, US Department of Education*.