



Universiteit
Leiden
The Netherlands

Teachers' assessment of "implementation leadership" during a new social emotional learning initiative

Shapiro, V.B.; Ziemer, K.L.; Accomazzo, S.; Elizabeth Kim, B. K.

Citation

Shapiro, V. B., Ziemer, K. L., Accomazzo, S., & Elizabeth Kim, B. K. (2019). Teachers' assessment of "implementation leadership" during a new social emotional learning initiative. *Contemporary School Psychology*, 24, 174-180. doi:10.1007/s40688-019-00230-7

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3748507>

Note: To cite this publication please use the final published version (if applicable).



Teachers' Assessment of "Implementation Leadership" during a new Social Emotional Learning Initiative

Valerie B. Shapiro¹ · Kelly L. Ziemer¹ · Sarah Accomazzo¹ · B. K. Elizabeth Kim²

Published online: 28 May 2019

© California Association of School Psychologists 2019

Abstract

School leadership has been identified as a factor related to successful implementation of school-based interventions. Brief, pragmatic tools to monitor implementation leadership may help school psychologists facilitate new initiatives. This pilot study adapted the Implementation Leadership Scale (ILS) to report on teacher-perceived change in implementation leadership across an academic year. Participants were 45 teachers from four elementary schools, implementing a social and emotional learning intervention. Cronbach's alphas were adequate. The proactive leadership subscale was most differentiated between schools. Levels of implementation leadership were significantly higher in the fall than the spring. The ILS shows some promise for assessing implementation leadership in schools.

Keywords School leadership · Implementation · Implementation Leadership Scale (ILS) · Social and emotional learning (SEL)

Introduction

School leadership is "the work of mobilizing and influencing others to articulate and achieve the school's shared intentions and goals" (Leithwood and Riehl 2005, p. 14). High-quality leadership has been conceptualized as critical to successful schools (Astor et al. 2009), and has shown to be positively associated with a variety of positive school outcomes, including increased student engagement and achievement (e.g., Leithwood et al. 2008), collective teacher efficacy (e.g., Ross et al. 2003), and positive school climate (Astor and Benbenishty 2018).

Implementation Leadership for Successful Implementation in Schools

The field of school psychology is committed to supporting the implementation of school-wide initiatives, especially as implicated with the increasing adoption of multi-tiered systems of

support (Forman et al. 2013; Forman and Crystal 2015). Implementation is the process through which an initiative is "brought into and established" within a particular service system (Elias 2010, p. 18). Research has identified school leadership as a predictor of successful implementation of school-wide interventions (Forman et al. 2009) and stressed the importance of leadership in the implementation of school mental health programs (Owens et al. 2013). The behaviors of school leaders may obstruct or facilitate implementation (Forman et al. 2009), which, in turn, can affect intervention effectiveness (Durlak and Dupre 2008). When leaders engage early and explicitly in implementation planning, interventions or initiatives have a greater likelihood of success (Aarons et al. 2016).

Transformational leadership, one of the most commonly utilized leadership models in education, has been theorized to impact the implementation of interventions by broadly transforming stakeholder attitudes and motivations through the articulation of an alternative vision (e.g., Gumus et al. 2018; Leithwood and Sun 2012). Transformational leadership (as opposed to transactional leadership where leaders rely on directives to drive change) encourages multiple leaders, rather than just supervisors, to take action (Gumus et al. 2018). While principals are generally considered school leaders by virtue of their position, teachers, students, parents, pupil service personnel, and other professional school staff, among others, may also provide transformational leadership. In other words, although school principals might have a special role realigning "structures and relationships to achieve genuine

✉ Valerie B. Shapiro
vshapiro@berkeley.edu

¹ School of Social Welfare, University of California Berkeley, 120 Haviland Hall, Berkeley, CA 94720, USA

² USC Suzanne Dworak-Peck School of Social Work, 669 West 34th Street SWC 218, Los Angeles, CA 90089, USA

and sustainable change” (Elias et al. 2006, p.11), transformational leadership requires the contribution of a variety of stakeholder groups, sometimes formally assembled as members of a leadership team (Locke et al. 2018). Elias et al. (2006) suggest that enacting effective transformational leadership requires a clear unified vision, courage for change, attention to building skills and capacity, infusion of innovations into the routine activities of school, and a sustained commitment of 2–5 years.

Measuring Implementation Leadership in Schools

Few psychometrically sound tools exist to assess leadership in schools (The Wallace Foundation 2009). Those that do exist do not focus on *implementation leadership* explicitly (Lyon et al. 2018), but rather on the assessment of broader leadership models or leadership domains, such as the domain of school performance (e.g., Vanderbilt Assessment in Leadership in Education; Porter et al. 2010). Although leadership has been conceptualized as a critical factor in implementation, there are few, if any, brief, pragmatic tools to monitor implementation leadership behavior for school-based interventions that could be used by school psychologists, as members of implementation leadership teams, to facilitate the monitoring and nurturing of effective implementation leadership (The Wallace Foundation 2009).

Implementation Leadership Scale The Implementation Leadership Scale (Aarons et al. 2014) is the first empirically validated, brief measure developed to assess the strategic actions of organizational leaders that influence successful program implementation. The Implementation Leadership Scale (ILS) is designed to assess four dimensions of Implementation Leadership (Aarons et al. 2014): Proactive Leadership (i.e., anticipates and eliminates challenges), Knowledgeable Leadership (i.e., thoroughly understands the practice), Supportive Leadership (i.e., enables the adoption and utilization of the practice), and Perseverant Leadership (i.e., continuously responds to emergent issues). The ILS has demonstrated acceptable evidence of reliability and validity in a variety of non-educational settings, including drug and alcohol treatment facilities (Aarons et al. 2016), child welfare agencies (Finn et al. 2016), and mental health clinics (Aarons et al. 2014). To our knowledge, only two studies (Locke et al. 2018; Lyon et al. 2018) have used the ILS to consider implementation leadership in an educational setting. One used the ILS to sample school-based behavioral health consultants engaged in a Tier 3 initiative at a single time point (Lyon et al. 2018). The other asked open-ended questions to central administrators, principals, and teachers in focus groups about the four ILS constructs to conceptualize implementation leadership in a Tier 1 initiative (Locke et al. 2018). No study to date, according to these authors’ knowledge, has been conducted

using the ILS with teachers as the primary informant of school leadership during the implementation of a school-wide, Tier 1, intervention across multiple time points.

The Current Study

This paper describes findings from a pilot study using the ILS at two time points to assess teacher perceptions of implementation leadership. Specifically, we explore (1) to what extent is the ILS reliable, as completed by teachers, in a school context? (2) Is the ILS sensitive to change across a school year? In other words, can we detect change from fall to spring?

Method

Study Design

Four elementary schools in a Northern California school district participating in the TOOLBOX Implementation Research Project (TIRP) implemented TOOLBOX (Collin 2015) over the 2015–2016 school year. TOOLBOX is a school-wide social and emotional learning (SEL) program. TOOLBOX developers provided an Administrator’s Checklist that directed each site to form an Implementation Leadership Team, inclusive of administrators, teachers, and classified staff, and to create and circulate an implementation plan to all teachers and staff.

Data for this study were collected through the Social Emotional Learning-Implementation Survey (SEL-IS; under review) at two time points: the fall (one month into instruction) and the spring (one month before the end of instruction). Individual responses were primarily collected through a computerized group administration during a staff meeting (or, secondarily, on isolated classroom computers). All teachers in these schools were asked by the district to participate in data collection for the purpose of continuous quality improvement, but the use of the data for research purposes required active consent. Of the 101 classroom teachers in schools implementing TOOLBOX, 95 teachers (94%) consented to be in the study. Informed consent was obtained, and this study was conducted, in compliance with the Institutional Review Board at the University of California, Berkeley.

Measure

Implementation Leadership Scale-TOOLBOX The Implementation Leadership Scale (ILS; Aarons et al. 2014) has four subscales, each composed of three items: Proactive Leadership (e.g., “Supervisor has removed obstacles to the implementation of [evidence-based practice (EBP)]”), Knowledgeable Leadership (e.g., “Supervisor is able to answer my question about [EBP]”), Supportive Leadership

(e.g., “Supervisor supports efforts to use [EBP]”), and Perseverant Leadership (e.g., “Supervisor carries on through the challenges of implementing the [EBP]”).

In the current study, teachers completed the Implementation Leadership Scale-TOOLBOX edition (ILS-TB), adapted from the ILS (Aarons et al. 2014). Items included some modifications to make the ILS items applicable to the current study setting and needs. Specifically, modifications to the questions were made so that the named EBP was “TOOLBOX” and “supervisor” was replaced with “school leadership.” This broader term was used intentionally to be inclusive of any and all personnel providing implementation leadership, since according to the transformational leadership model, implementation leadership teams in schools are likely to include leaders in non-supervisory roles, including teachers and staff. For example, the ILS item “Supervisor is able to answer my question about [EBP]” was modified to “Is your school leadership able to answer questions about TOOLBOX?” For consistency, the ILS-TB used the same response options and scoring procedure as the ILS. Each item was measured from 0 (not at all) to 4 (to a very great extent). Each subscale is the average of 3 items assigned to the scale. The total ILS-TB score is the average of the four subscales.

Implementation Leaders

Although not included in this study, school personnel who were *not* classroom teachers were also invited and encouraged to complete the SEL-IS. In doing so, all responding classroom teachers and other school personnel were asked if they served on a school-wide implementation committee and/or had any formal leadership responsibilities for TOOLBOX implementation. Self-identified leaders, subsequently assessed by teacher non-leaders through the ILS-TB, included classroom teachers, specialty teachers, administrators, other certified professionals (e.g., school psychologists), and other non-certified professionals (e.g., lunch period supervisors). At baseline, self-identified leaders ($n = 17$) described themselves as 87% female, 53% European-American/White, 6% Hispanic/Latinx, 12% African-American/Black, 24% Multi-Race, and 6% Other. Eight (47%) of these self-identified leaders had worked in the district for more than a decade. Nine (53%) of these self-identified leaders continued to identify themselves as having formal leadership responsibilities for TOOLBOX when assessed in the spring, with two additional people self-identifying themselves as leaders in the spring. Parents and youth were not surveyed, nor were they formally incorporated into Implementation Leadership Teams for this SEL Initiative.

Analysis Sample

Teachers who self-identified as implementation leaders at baseline ($n = 10$, or 11% of the 95 teachers who consented to

research) were eliminated from this analysis to maintain the third-party perspective. Nearly three quarters of consenting teachers completed all 12 ILS-TB items in the fall ($n = 63$), enabling the computation of a fall total ILS-TB score. Of these, 45 teachers (71%) also completed all 12 ILS-TB items in the spring. Thus, 45 teachers were included in the analysis sample.

There were no significant differences in gender, race, or fall ILS-TB scores between teachers with ILS-TB data in both fall and spring (i.e., those included in analysis sample) relative to those with only fall data (i.e., excluded from analysis sample). Teachers in the analysis sample identified as 96% female, 62% European-American/White, 11% Asian/Asian-American, 9% Hispanic/Latinx, 7% African-American/Black, 7%, Multi-Race, and 4% Other. About half (49%) had worked in the district for more than a decade.

Data Analyses

Analyses were conducted using SPSS (v.24) and Stata13 (Statacorp 2013). We used Cronbach’s alpha to assess internal reliability of the ILS-TB (Cronbach 1951); a score of .70 or above was considered acceptable as a threshold for judging internal reliability for intended uses (Nunnally 1978). We calculated intraclass correlations (ICCs) to explore variation in teacher-reported implementation leadership by school using multilevel models (Rabe-Hesketh and Skrondal 2012). We used paired sample t tests with Cohen’s d metrics to explore the statistical significance and scope of changes in implementation leadership between fall and spring. Cohen (1988) suggests interpreting d ratios above .2 as small, .5 as medium, and .8 as large. Two-tailed t tests with p values $< .05$ are reported as statistically significant.

Results

Internal Reliability of ILS-TB

Cronbach’s alpha for the fall total ILS-TB score was .94, with the four subscales ranging from .78 (Supportive Leadership) to .88 (Perseverant Leadership). The spring total ILS-TB score was also .94, with the four subscales ranging from .68 (Proactive Leadership) to .91 (Knowledgeable Leadership) (see Table 1). Item-level analyses suggested that inter-item correlations did not improve by removing individual items.

Variance Between Schools

The fall total ILS-TB score and subscale ICCs were all less than .001, with the exception of Proactive Leadership (ICC = .12). In other words, 12% of the variation in teachers’ report of proactive implementation leadership was accounted

for at the school level, but the other leadership constructs did not vary among schools. The spring total ILS-TB score and subscales were, again, all less than .001, with the exception of Supportive Leadership (ICC = .01) and Proactive Leadership (ICC = .12) (see Table 1).

Change in Implementation Leadership

The extent to which the Implementation Leadership Teams were active appears to have evolved from fall to spring. In the fall, only 36% (16) teachers in the analysis sample reported that the SEL leadership committee at their school was active. In the spring, 42% (19) of the teachers in the analysis sample reported that the SEL committee at their school was active. The quality of implementation leadership also evolved from fall to spring. In the fall, teachers perceived implementation leadership to be “moderate” to “great” (total ILS-TB score $M = 2.42$; $SD = .73$) and subscales ranging from the lowest levels in Proactive Leadership ($M = 2.10$; $SD = .92$) to the highest levels in Supportive Leadership ($M = 2.70$, $SD = .79$). In the spring, teachers reported statistically significant lower quality implementation leadership (total ILS-TB score $M = 2.19$; $SD = .66$; $p = .01$), with subscales ranging from Proactive Leadership ($M = 2.03$; $SD = .06$) to Knowledgeable Leadership ($M = 2.41$, $SD = .75$). The decline in the total ILS-TB score appears to be driven by a statistically significant decline (maintained with a Bonferroni correction) in Supportive Leadership ($p < .001$). The magnitude of these effects was medium (total $d = .33$) and large (supportive $d = .69$) (see Table 1).

Discussion

In our study, the ILS-TB performed reasonably well, showing adequate internal reliability across nearly all scales when completed by teachers in school settings. The total score had excellent reliability ($\alpha = .94$) for the intended use of assessing the actions of school implementation leadership teams. The ILS-TB alpha reliability scores, however, were lower than those generated by the ILS when completed by clinicians about the leadership behaviors of their supervisors/administrators; Aarons et al. (2014) and Lyon et al. (2018) observed scale-level alpha reliabilities that were all .95 or higher. This is not unexpected when the assessment target is a diverse team of actors rather than a single person. In practice, including the names of the members of the leadership team may improve the internal reliability of the subscale scores, if they are to be used formatively for continuous quality improvement (CQI) efforts. Alternatively, ILS users may consider disaggregating the assessment of various members such that different individuals receive differential feedback for CQI. Yet, perhaps the even more important implication for

practitioners comes from the finding that the majority of non-leader respondents in this study, conducted under routine practice conditions, did not perceive the implementation leadership teams for this initiative to be active. Practitioners may want to ensure that the leadership team members and their activities are seen, understood, and felt among all teachers and staff as part of achieving implementation success.

Little variance in ILS-TB scores was explained at the school level. The ICCs in this study were quite small (most $< .001$), suggesting minimal variation in leadership among schools within the same district. Data from clinicians completing the ILS in mental health clinics (Aarons et al. 2014) generated bigger ICCs (.22–.29), indicating greater variation across groups relative to within groups. The only other quantitative study using the ILS in a school setting (Lyon et al. 2018) did not analyze individual responses in clusters. Reasons for small variation across schools could include relatively high variance between respondents within schools, joint trainings and leadership planning meetings across the schools, and shared environmental factors within the same district. Future studies with larger samples may consider exploring factors that contribute to variation in perceptions of implementation leadership, including individual respondent and setting-level characteristics.

In our study, on average, teachers reported moderate levels of implementation leadership in the fall ($M = 2.42$; $SD = 1.12$), similar to those reported by clinicians in mental health clinics ($M = 2.42$; $SD = .73$; Aarons et al. 2014). We suspect the range in standard deviations across studies is largely a function of sample size and sampling strategy. Our scores cannot directly be compared to those observed by Lyon and colleagues in their study of a school setting, as Lyons reported factor scores rather than replicating the ILS scoring procedures. We believe the ILS scoring procedures are more pragmatic for routine practice.

To our knowledge, this is the first study to use the ILS to look at implementation leadership over time in an education setting. Teachers reported significantly lower levels of supportive leadership (and total implementation leadership) in the spring compared to the fall. These reductions in perceived implementation leadership over the course of the year could be due to actual changes in leadership behavior or to escalating needs that render the same behavior less adequate or less apparent. For example, Locke et al. (2018) found that teacher and administrator fatigue played a role in their assessment of supportive leadership. They found that to overcome such fatigue, leaders were expected to be empathetic to implementation challenges, provide regular check-ins, and troubleshoot issues (Locke et al. 2018). This is consistent with the emerging role of school psychologists as consultants, collaborators, coaches, and facilitators of “the implementation problem-solving process” (Forman and Crystal 2015, p.279; Rosenfield 2013; Eagle et al. 2015). A fifth construct,

Table 1 Implementation Leadership Scale-TOOLBOX edition (ILS-TB) Scale Statistics

	Fall			Spring			Change between fall and spring		
	Mean	Standard deviation	Alpha ICC	Mean	Standard deviation	Alpha ICC	Significance (<i>p</i> value)	Effect size (Cohen's <i>d</i>)	Confidence interval
Proactive leadership	2.10	0.92	0.82	2.03	0.68	0.68	0.12	0.50	– .33, .50
School leaders provide written plans for TOOLBOX implementation	2.07	1.12		2.22	0.93				
School leaders remove obstacles to implementation	2.09	0.97		1.93	0.81				
School leaders provide clear standards for implementation	2.16	1.11		1.93	0.86				
Knowledgeable leadership	2.55	0.76	0.82	2.41	0.75	0.91	< .001	0.16	– .23, .60
School leaders know what they are talking about when it comes to TOOLBOX	2.67	0.95		2.58	0.87				
School leaders are knowledgeable about TOOLBOX	2.47	0.94		2.42	0.78				
School leaders are able to answer questions about TOOLBOX	2.51	0.82		2.22	0.79				
Supportive leadership	2.70	0.79	0.78	2.17	0.74	0.83	0.01	< .001*	.26, 1.12
School leaders support staff to use TOOLBOX	2.96	0.90		2.38	0.86				
School leaders support staff to learn more about TOOLBOX	2.53	0.99		2.04	0.85				
School leaders recognize staff efforts for successful implementation	2.62	0.96		2.09	0.87				
Perseverant leadership	2.34	0.81	0.88	2.16	0.80	0.89	< .001	0.18	– .19, .64
School leaders perseveres through ups and downs of implementation	2.33	0.83		2.18	0.91				
School leaders carry on through challenges implementing TOOLBOX	2.44	0.84		2.24	0.83				
School leaders openly and effectively addresses problems implementing TOOLBOX	2.24	1.00		2.07	0.91				
Total ILS-TB score	2.42	0.73	0.94	2.19	0.66	0.94	< .001	0.01*	– .09, .75

* $p < .05$ Interpreted as statistically significant

availability leadership, has been suggested in two prior papers as important for implementation and sustainability (Ehrhart et al. 2018; Locke et al. 2018). School psychologists serving on implementation leadership teams who make themselves available to other school personnel may therefore be able to improve implementation success.

Of note, the Proactive Leadership scale performed somewhat differently than the other subscales. Our analyses suggested some deterioration in the internal reliability of the Proactive Leadership scale between fall ($\alpha = .82$) and spring ($\alpha = .68$). Interestingly, the proactive leadership scale is also the only scale showing some variance between schools. We wonder if the behaviors captured in the Proactive Leadership scale might be behaviors that typically occur earlier in the school year, in advance of implementation, and may therefore be more subject to recall bias or other sources of measurement variance. Principal and other personnel changes that occurred over the summer, in some locations, may have played a role in the variance between schools. This pilot study could help make a case for using the ILS as part of readiness assessment and planning processes prior to the initiation of a new initiative.

The small number of teachers and schools involved in the current pilot study limits the generalizability of its findings. As noted previously, future studies may consider clarifying and/or isolating implementation leaders so that the assessment target is less diffuse to teachers. The problem of diffusion in the assessment of leadership is a common issue in schools aspiring for transformational leadership (e.g., Leithwood et al. 2004; Reichenpfader et al. 2015) as distinct from other settings that may emphasize transactional leadership. Future studies with larger samples may consider confirming the ILS factor structure with teacher informants in educational settings. Additionally, it may be useful to move beyond the analysis of scale structures to the exploration of validity by relating direct observations of leaders or artifacts (e.g., meeting minutes) to ILS scores. One may also consider whether ILS scores act in accordance with theory. For example, initial analyses have revealed a positive relationship between initial proactive school leadership and the number of TOOLBOX lessons teachers ultimately delivered during their first year of implementation (Lee et al. 2018).

While leadership is commonly understood to be a critical component of successful implementation of interventions in schools, we know of no other brief, psychometrically sound assessment tools for formatively monitoring implementation leadership to facilitate implementation. The current pilot study, described in this brief report, provides preliminary support (with minor cautions) for school psychologists on Implementation Leadership Teams to use the ILS to inform their efforts to continuously improve implementation efforts of school-wide programs in school settings.

Acknowledgements This project, and the associated data collection and analyses, was enabled through the support of the Stuart Foundation. The authors wish to acknowledge all of the teachers, staff, and administrators who participated in this project, and specifically Fred Brill, Ammar Saheli, Barbara DeBarger, Katarin Jurich, Neal Bloch, Moises Hernandez, Kathleen LaCome, Denise Landry, Kimberly Yearns, Wendy Garner, Gail Drake, Ruben Olivares, and Elisa Garcia-Rojas. The authors also wish to thank Brittany Schmitt and colleagues at the Center for Prevention Research in Social Welfare (CPRSW) who played essential roles in the completion of this work. The authors are appreciative for consultation from Mark Collin and Chuck Fisher on the TOOLBOX program.

Author's note Since completing this work, Sarah Accomazzo has changed her institutional affiliation. She now conducts research at Seneca Family of Agencies.

Compliance with Ethical Standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

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

Disclaimer The opinions expressed in this article do not necessarily reflect the views of any funding agency, resource developer, or community partner.

References

- Aarons, G. A., Ehrhart, M. G., & Farahnak, L. R. (2014). The implementation leadership scale (ILS): development of a brief measure of unit level implementation leadership. *Implementation Science*, 9(1), 45.
- Aarons, G. A., Ehrhart, M. G., Torres, E. M., Finn, N. K., & Roesch, S. C. (2016). Validation of the implementation leadership scale (ILS) in substance use disorder treatment organizations. *Journal of Substance Abuse Treatment*, 68, 31–35.
- Astor, R. A., & Benbenishty, R. (2018). *Mapping and monitoring bullying and violence: building a safe school climate*. New York: Oxford University Press.
- Astor, R. A., Benbenishty, R., & Estrada, J. (2009). School violence and theoretically atypical schools: the principal's centrality in orchestrating safe schools. *American Educational Research Journal*, 46(2), 423–461.
- Cohen J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: Routledge Academic.
- Collin, M. A. (2015). *The TOOLBOX Curriculum Guide (Primer; Lesson Plans K-3, & Lesson Plans 4-6)*. Sebastopol: Dovetail Learning, Inc..
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- 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>.

- Eagle, J. W., Dowd-Eagle, S. E., Snyder, A., & Holtzman, E. G. (2015). Implementing a Multi-Tiered System of Support (MTSS): collaboration between school psychologists and administrators to promote systems-level change. *Journal of Educational and Psychological Consultation*, 25(2–3), 160–177. <https://doi.org/10.1080/10474412.2014.929960>.
- Ehrhart, M. G., Torres, E. M., Green, A. E., Trott, E. M., Willging, C. E., Moullin, J. C., & Aarons, G. A. (2018). Leading for the long haul: a mixed-method evaluation of the Sustainment Leadership Scale (SLS). *Implementation Science*, 13(1), 17. <https://doi.org/10.1186/s13012-018-0710-4>.
- Elias, M. J. (2010). Sustainability of social-emotional learning and related programs: lessons from a field study. *International Journal of Emotional Education*, 2(1), 17–33.
- Elias, M. J., O'Brien, M. U., & Weissberg, R. P. (2006). Transformative leadership for social-emotional learning. *Principal Leadership*, 7(4), 10–13.
- Finn, N. K., Torres, E. M., Ehrhart, M. G., Roesch, S. C., & Aarons, G. A. (2016). Cross-validation of the implementation leadership scale (ILS) in child welfare service organizations. *Child Maltreatment*, 21(3), 250–255.
- Forman, S. G., & Crystal, C. D. (2015). Systems consultation for multi-tiered systems of supports (MTSS): implementation issues. *Journal of Educational and Psychological Consultation*, 25, 276–285. <https://doi.org/10.1080/10474412.2014.963226>.
- Forman, S. G., Olin, S. S., Hoagwood, K. E., Crowe, M., & Saka, N. (2009). Evidence-based interventions in schools: developers' views of implementation barriers and facilitators. *School Mental Health*, 1(1), 26–36.
- Forman, S. G., Shapiro, E. S., Coddling, R. S., Gonzales, J. E., Reddy, L. A., Rosenfield, S. A., ... Stoiber, K. C. (2013). Implementation science and school psychology. *School Psychology Quarterly*, 28(2), 77–100. <https://doi.org/10.1037/spq0000019>.
- Gumus, S., Bellibas, M. S., Esen, M., & Gumus, E. (2018). A systematic review of studies on leadership models in educational research from 1980 to 2014. *Educational Management Administration & Leadership*, 46(1), 25–48. <https://doi.org/10.1177/1741143216659296>.
- Lee, J., Shapiro, V. B., Kim, B. K. E., & Yoo, J. P. (2018). Multilevel structural equation modeling (MSEM) for social work researchers: an introduction and an application to healthy youth development. *Journal of the Society for Social Work and Research*, 9(4), 689–719.
- Leithwood, K. A., & Riehl, C. (2005). What do we already know about educational leadership? In W. A. Riehl & C. Riehl (Eds.), *A New Agenda for Research in Educational Leadership* (pp. 12–27). New York: Teachers College Press.
- Leithwood, K., & Sun, J. (2012). The nature and effects of transformational school leadership: a meta-analytic review of unpublished research. *Educational Administration Quarterly*, 48(3), 387–423. <https://doi.org/10.1177/0013161X11436268>.
- Leithwood, K., Seashore Louis, K., Anderson, S., & Wahlstrom, K. (2004). *How leadership influences student learning*. New York: Wallace Foundation.
- Leithwood, K., Harris, A., & Hopkins, D. (2008). Seven strong claims about successful school leadership. *School Leadership & Management*, 28(1), 27–42. <https://doi.org/10.1080/13632430701800060>.
- Locke, J., Lee, K., Cook, C. R., Frederick, L., Vázquez-Colón, C., Ehrhart, M. G., ... Lyon, A. R. (2018). Understanding the organizational implementation context of schools: a qualitative study of school district administrators, principals, and teachers. *School Mental Health. Advance Online Access*. <https://doi.org/10.1007/s12310-018-9292-1>
- Lyon, A. R., Cook, C. R., Brown, E. C., Locke, J., Davis, C., Ehrhart, M., & Aarons, G. A. (2018). Assessing organizational implementation context in the education sector: confirmatory factor analysis of measures of implementation leadership, climate, and citizenship. *Implementation Science*, 13(1), 5. <https://doi.org/10.1186/s13012-017-0705-6>.
- Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: McGraw-Hill.
- Owens, J. S., Lyon, A. R., Brandt, N. E., Warner, C. M., Nadeem, E., Spiel, C., & Wagner, M. (2013). Implementation science in school mental health: key constructs in a developing research agenda. *School Mental Health*, 6(2), 99–111. <https://doi.org/10.1007/s12310-013-9115-3>.
- Porter, A. C., Polikoff, M. S., Goldring, E., Murphy, J., Elliott, S. N., & May, H. (2010). Developing a psychometrically sound assessment of school leadership: the VAL-ED as a case study. *Educational Administration Quarterly*, 46(2), 135–173. <https://doi.org/10.1177/1094670510361747>.
- Rabe-Hesketh, S., & Skrondal, A. (2012). *Multilevel and Longitudinal Modeling using Stata*. Stata Press.
- Reichenpfader, U., Carljford, S., & Nilsen, P. (2015). Leadership in evidence-based practice: a systematic review. *Leadership in Health Services*, 28(4), 298–316. <https://doi.org/10.1108/LHS-08-2014-0061>.
- Rosenfield, S. (2013). Consultation in the schools—Are we there yet? *Consulting Psychology Journal: Practice and Research*, 65(4), 303–308. <https://doi.org/10.1037/a0035652>.
- Ross, J. A., Hogaboam-Gray, A., & Gray, P. (2003). The contribution of prior student achievement and school processes to collective teacher efficacy in elementary schools. Retrieved February 6, 2018 from <https://eric.ed.gov/?id=ED479719>
- StataCorp. (2013). *Stata Statistical Software: Release 13*. College Station: StataCorp LP.
- The Wallace Foundation. (2009). Assessing the effectiveness of school leaders: new directions and new processes. Retrieved April 18, 2018 from <http://www.wallacefoundation.org/knowledge-center/pages/assessingthe-effectiveness-of-school-leaders.aspx>.

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

Valerie B. Shapiro PhD, is faculty chair of the Doctoral Program in Social Welfare and the Co-Director of the Center for Prevention Research in Social Welfare at Berkeley. She was the recipient of the 2014 Prytanean Faculty Prize, recognized for scholarly achievement and distinguished teaching at UC Berkeley. Dr. Shapiro researches the prevention of mental, emotional, and behavioral problems in children and youth through the adoption, implementation, and sustainability of effective prevention practices. She serves on the steering committee of the Coalition for the Promotion of Behavioral Health that produced the National Academy of Medicine Paper entitled “Unleashing the Power of Prevention” and serves on a statewide taskforce to improve infrastructure for prevention service delivery in schools.