



Exploring Relationships Between CARE Program Fidelity, Quality, Participant Responsiveness, and Uptake of Mindful Practices

Sebrina L. Doyle¹ · Patricia A. Jennings² · Joshua L. Brown³ · Damira Rasheed³ · Anna DeWeese³ · Jennifer L. Frank¹ · Christa Turksma¹ · Mark T. Greenberg¹

Published online: 29 October 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract

As the demand for mindfulness-based interventions grows, so does the need to monitor implementation to ensure effective delivery. Research has shown that the fidelity and quality of program implementation can impact intervention outcomes. This study reports on the development and use of implementation monitoring measures for the Cultivating Awareness and Resilience in Education (CARE) professional development program. Three facilitator-focused observer-rated measures of fidelity and quality of intervention delivery and three participant-focused measures that examined responsiveness and attendance to the program were used. Data were evaluated from six intervention implementations including 166 teachers from 36 schools in New York; 80% of participants provided follow-up data on engagement in mindful practices. We report on regression analyses that examined relationships between implementation monitoring measures and follow-up teacher outcomes related to uptake and sustainment of mindful awareness practices. Results indicated significant positive relationships with an alternative measure of adherence focused on completion of participant objectives and facilitation quality with follow-up teacher reports of continued practice (mindfulness practices used weekly, length of practices, extent that practices helped with stress). Further, participants' perceived usefulness of the material demonstrated a strong positive relationship with length of practice and whether practices helped to reduce stress. Finally, interactions were found between attendance and adherence, and between facilitation quality and participant knowledge on practice outcomes. We explore the way these measures were used to support program delivery and facilitator training, and discuss the utility of an alternative measure of adherence focused on completion of participant learning objectives.

Keywords Mindful awareness practices · Fidelity · Implementation · Participant responsiveness · Education · Professional development · Facilitation quality · Mindfulness-based interventions

Mindfulness-based interventions (MBIs) have enjoyed a recent surge in popularity in the secular world. As the demand for such programs grows, so can the tension between increased program reach and high program fidelity. Program fidelity refers to the extent to which an MBI is delivered in

the way originally intended by the developers. When developers are also facilitators, there is less concern about the program being delivered as intended, therefore it is common to see only basic implementation monitoring tools used. The focus is often on examining participant satisfaction or knowledge gained through end-of-program evaluations, or measuring program fidelity by recording whether specific program content was delivered.

As programs mature, the demand for implementation can exceed the developers' capacity for delivery; when this occurs, "train the facilitator" models are often employed. As program delivery expands in scope and in geography, the developers' ability to supervise implementation becomes increasingly difficult. During this growth process, there is increased risk of drift or dilution of the core essence of the program; therefore, increased monitoring of

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s12671-018-1034-9>) contains supplementary material, which is available to authorized users.

✉ Sebrina L. Doyle
sld40@psu.edu

¹ Pennsylvania State University, University Park, PA, USA

² University of Virginia, Charlottesville, VA, USA

³ Fordham University, New York, NY, USA

implementation is critical for successful scaling efforts (Milat et al. 2015).

Similar to program growth seen by other MBIs, in recent years the Cultivating Awareness and Resilience in Education (CARE) professional development program has moved from being offered yearly by the developers in a week-long retreat setting, to being delivered several times a year across the country. CARE was originally conceptualized to address the growing problems of teacher stress and burnout contributing to high turnover rates in educational settings; with current decreases in educational funding combined with increases in school testing, the demand for programs that address stress management is growing (Greenberg et al. 2016). CARE is a mindfulness-based stress reduction program that encompasses mindful awareness practices, caring and compassion practices, and emotion skills instruction (for a more detailed description of the program model and curriculum elements, please see Jennings 2015 and Jennings et al. 2011). The program has demonstrated improvements in teachers' well-being through reductions in psychological distress and increases in adaptive emotion regulation, efficacy, and mindfulness; improvements have also been observed at the classroom level in emotionally supportive and organization-related interactions between teachers and students (Jennings et al. 2013, 2017).

In 2009, funding was received to refine the program and test its efficacy for improving teacher, classroom, and student outcomes, in two pilot studies (Jennings et al. 2011, 2013) and more recently in a large cluster randomized controlled trial (RCT) (Jennings et al. 2017). Implementation assessment in the pilot studies included measurement of fidelity through delivery of specified curriculum components (adherence), and participant responsiveness through end-of-program satisfaction surveys. With the growth in demand for CARE, and the need for more implementations than the developers could facilitate or supervise, more detailed, observable measures of program implementation were needed for documenting fidelity and quality for research purposes, and for supporting facilitator training and program improvement.

For many years, program implementation monitoring received little attention in the social sciences. Programs that demonstrated significant impacts on desired outcomes were assumed to have been delivered as intended. In the late 1980s, increased attention to the importance of treatment fidelity began to emerge in intervention research. In their review on the topic, Moncher and Prinz (1991) wrote, "Treatment fidelity is a relatively recent methodological consideration..." (p. 247), linking the increased focus on treatment fidelity with the demand for accountability in social science research. Several reviews published since have addressed the link between program implementation and study outcomes (Dane and Schneider 1998; Domitrovich and Greenberg 2000; Durlak and Wells 1997; Dusenbury et al. 2003). Durlak and DuPre

(2008) found that, in 76% of studies they reviewed, program implementation was significantly positively correlated with outcomes; further, programs that monitored their intervention showed effect sizes for their programs three times larger than those that did not.

Though some MBIs designed for clinical populations have been quick to adopt implementation fidelity measurement (e.g., Segal et al. 2002), programs delivered to a broader, non-clinical audience show less evidence of monitoring for implementation effectiveness. The tension between program adherence and facilitators' freedom to adapt the program "in the present moment" according to the facilitators' perceptions of participants' needs has been responsible for some resistance (Dobkin et al. 2014). The lack of implementation monitoring is especially evident in MBIs designed for use in educational settings. In a recent review, Feagans Gould et al. (2016) found that fewer than 20% of mindfulness and yoga interventions designed for delivery in educational settings assessed the fidelity of their implementation beyond simple dosage.

There is a growing trend toward being explicit about how one is monitoring implementation fidelity in education. In a recent study, Schonert-Reichl and Lawlor (2010) explained their measurement of implementation fidelity of the Mindfulness Education program for students; fidelity was measured as adherence to the curriculum through completion of core components, and teacher ratings of student outcomes were also captured to measure participant responsiveness. Feagans Gould et al. (2014), in their study of a school-based yoga program, detailed the process of developing a fidelity measurement strategy they termed "CORE Process," including conceptualizing core components, operationalization and measurement of those components, analysis and review of implementation findings, and enhancement and refinement of those measurements.

Several in-depth reviews of the literature on program implementation have provided an accounting of measurement strategies used to monitor implementation in social science research; four elements are common across reviews: adherence to the program curriculum, dosage of program, quality of program delivery, and participant responsiveness (Dane and Schneider 1998; Durlak and DuPre 2008; Dusenbury et al. 2003). Dane and Schneider (1998, p. 45) defined adherence as "the extent to which specified program components were delivered as prescribed in program manuals"; quality of delivery was defined as "a measure of qualitative aspects of program delivery that are not directly related to the implementation of prescribed content." The term *adherence* is often used interchangeably with *fidelity* in the literature; however, fidelity has also, at times, been defined in a way that encompasses aspects of quality (e.g., fidelity to process; Mowbrey et al. 2003). Finally, Dane and Schneider defined participant responsiveness as "a measure of participant response to program sessions...such as levels of participation and

enthusiasm” (p. 45). Dosage, broadly defined, is the amount of intervention provided. Helpful tables are provided in the articles mentioned above, as well as in Feagans Gould et al. (2016), that illuminate how implementation monitoring elements were measured across several different studies.

Recently, Berkel et al. (2011) posited a theoretical framework that attempted to explain the relationships between elements of implementation monitoring. They proposed facilitator behaviors (e.g., adherence, quality of delivery) as separate from, but interactive with participant responsiveness (e.g., attendance, participation, satisfaction) which, in turn, affected outcomes. In this “integrated theoretical model of program implementation,” Berkel et al. offer several hypotheses regarding how these elements interact. First, they hypothesized that facilitation quality interacts with implementation fidelity (adherence) to influence program outcomes. Second, they proposed that participant responsiveness serves as a mediator for facilitation quality and as a moderator for implementation fidelity on outcomes. This model was proposed, in part, to account for what may be seen as a lack of direct effects of implementation fidelity on hypothesized outcomes.

The present study has two aims. The first aim is to provide a detailed description of the comprehensive measurement strategy employed to monitor six implementations of the CARE program provided during the course of an RCT when developer-trained facilitators delivered the program; we examine how these implementation monitoring tools serve as a guide for new facilitators and for overall program improvement. The second aim is to elucidate the relationships between measures of adherence to the curriculum and facilitation quality, participant responsiveness, and outcomes related to the uptake of mindfulness practices. In addition to exploring direct relationships, we also test Berkel et al.’s (2011) mediation and moderation hypotheses.

Method

Participants

The present study examined data collected during six program implementations provided during an RCT. The parent study sample consisted of 224 teachers recruited from 36 schools who expressed interest in attending the CARE program. Fifty-three teachers were recruited in 2012–2013, and 171 teachers were recruited in 2013–2014. Two program implementations were offered for teachers enrolled in the study in 2012–2013, and four program implementations were offered for teachers who enrolled in 2013–2014 (see Table 1 for timelines for implementation groups). Teachers were randomly assigned within schools to receive CARE either in fall/winter of the current school year or in the spring of the following school year (information on recruitment and randomization of parent

study can be found in Jennings et al. 2017). All 224 teachers (intervention and wait-list control) were offered the opportunity to participate in a program; 166 (74%) attended at least one day. Reasons for non-attendance included the following: 48%, no longer interested/other commitments; 10%, no longer in teaching profession when program was offered; 7%, administrative refusal prevented participation; 3%, illness or death in the family; 5%, withdrew from the study; and 26%, no response to invitation.

Teachers who attended at least one day were considered part of the sample for this study. Of those teachers, 93% were female ($n = 154$); 34% of teachers identified as White, 31% as Hispanic, 25% as African American/Black, 4% as Asian, and 5% as mixed racial background. Teachers ranged from 22 to 73 years old. Eighty percent of participating teachers ($n = 134$) also completed a follow-up assessment. Of teachers who did not complete the follow-up, 18% indicated a job change (e.g., retirement, not teaching); the remainder either said they did not have time to participate or did not respond. No significant demographic differences were found between those who did or did not participate in the program, or for those who did and did not complete the follow-up (age, gender, race, education, experience, or class grade level).

A post-hoc power analysis was conducted using G*power version 3.1 (Erdfelder et al. 2009) using the follow-up sample size ($n = 134$) and an α error probability of 0.05; accounting for ten predictors, our study had 80% power to detect change in a regression coefficient with a Cohen’s f^2 effect size of .05. This is above the cutoff for a small effect (.02) but much lower than the cutoff (.15) for a medium effect (Cohen 1992).

Procedures

The Development of Observational Measures The first step in the development of implementation monitoring tools was to define what fidelity to the manual would look like to an observer. The process began with identification of all teaching components in the CARE facilitator’s manual (Jennings et al. 2016); 46 components were identified over the entire 30-hour program. This included didactic (e.g., introduction to mindfulness, working with scripts), practice (e.g., intention setting, basic breath awareness), and community building components (e.g., establishing group agreements, sustainability planning). Next, facilitation activities required to teach each component were delineated (e.g., defined explanations, guided participant experiences, debriefing topics). Each component contained between one and seven facilitation activities, depending on the complexity and whether it was taught in earlier sessions; some practice components were repeated over program days. Identification of facilitation activities allowed researchers to examine traditional adherence to curriculum and also provided an outline for facilitators during program implementation to ensure coverage of key points. We also identified

Table 1 Data collection and implementation timeline for implementation groups

Recruited	Group	N	2012		2013		2014			2015		
			Fall	WI	SP	Fall	WI	SP	Fall	WI	SP	Fall
2012–2013	1	25	B, I	I		FU						
2012–2013	2	14				B		I				FU
2013–2014	1	40				B, I	I		FU			
2013–2014	2	47				B, I	I		FU			
2013–2014	3	16							B		I	FU
2013–2014	4	24							B		I	FU

WI winter, SP spring, B baseline, I intervention, FU follow-up

participant learning objectives, capturing what should be understood or experienced by participants during each component; this was used as a second, alternative measure of adherence. Each component contained between one and three participant learning objectives. This process resulted in the creation of the CARE Daily Session Rating Forms (Doyle et al. 2014). To measure quality of delivery, we identified CARE specific behavioral indicators for observed facilitation skills listed on a preexisting measure (Iowa State University Extension and Outreach 2010). The original version of this measure was used previously to capture quality of facilitation for the mindfulness-based adaptation of the Strengthening Families Program (Coatsworth et al. 2015). Facilitators and observational coders were provided with all forms prior to the training (see online supplemental Table 1, online supplemental Table 2, and online supplemental Document A for more detailed examples of the materials).

CARE Program Implementation Curriculum for the six program implementations included a menu of mindful awareness practices, emotional awareness skills, caring and compassion topics, and community building activities. Participants were encouraged to develop a self-care plan by selecting and practicing activities in their life and work that they found helpful for managing stress and improving teaching. Practices were introduced sequentially, starting with the most basic (e.g., taking three deep breaths) and moving to longer, more complex practices (e.g., 10-minute body scan); practices accounted for ~25% of program time. Emotional awareness skills (e.g., understanding our anger) were taught through didactic presentation, activities, and discussion and encompassed ~30% of the session time; some practices (e.g., experiencing joy) also helped participants identify somatic experiences associated with emotions. Caring and compassion topics and practices represented ~25% of the curriculum; this included didactic material and discussion on empathy and compassion as well as practices such as “caring practice” (*metta*) and mindful listening. Participants also engaged in community building activities (e.g., facilitated introductions) ~20% of the time to foster comradery and growth of a CARE network for program

sustainability. Participants received a workbook and pre-recorded mindfulness practices, and were encouraged to use the resources at home between session days. For more details and the rationale behind program components, see Jennings et al. (2011).

A team of three facilitators delivered each program implementation; eight facilitators assisted across the six program implementations. All facilitators met basic criteria of having a graduate-level degree, at least two years of personal mindfulness practice, and participation in a prior CARE program. The requirement of a personal mindfulness practice is consistent with other MBIs (Dobkin et al. 2014; Prowse et al. 2015).

Implementation was monitored by a team of coders who had previously attended a CARE program. They also attended at least four hours of coder training; during training, coders were shown videos of the program developers leading sessions to use as a benchmark for “exemplary” facilitation behaviors. Non-exemplary examples were also presented for contrast. Coders also received and reviewed a coding manual (available from the first author upon request). At each daily session, two trained coders independently rated implementation. With the exception of the first implementation when researchers who were developing the tools conducted the coding, coders were randomized across session days to minimize coder bias. Codes were checked for reliability and disagreements were resolved by consensus with support from the coding supervisor. Inter-rater reliability was calculated using the independent coder scores from the first three CARE program implementations; for the final three program implementations, only combined consensus codes were recorded. The final consensus codes for all six program implementations were used for data analysis.

Measures

Three observer-rated facilitator-focused measures of program implementation were used: (i) traditional adherence-completion of facilitation activities, (ii) alternative adherence-completion of participant learning objectives and (iii) facilitation quality. Additionally, three measures of

participant responsiveness were used: teacher self-reports of (i) understanding of material and (ii) usefulness of material, and (iii) recorded attendance to the program. Finally, we used three outcome measures related to continued use of CARE practices assessed approximately nine months following the training: (i) number of CARE practices used weekly, (ii) length of longest practice, and (iii) teachers' perceptions of whether practices reduced their stress level. Baseline assessments of previous use of mindful practices and teacher age were used as covariates in all statistical models (see online supplemental Table 3 for more information).

Traditional Adherence—Completion of Facilitation Activities

Using the previously described CARE Daily Session Rating Forms, each expected facilitation activity was rated on a 0 = absent/1 = present scale. Cohen's kappa was used to assess inter-rater reliability on the raw individual coder binary response scores from the first three implementations; inter-rater reliability on completion of facilitation activities fell in the acceptable range ($\kappa = .67$) (Cohen 1960). For all six implementations, coders' final consensus ratings were used to create daily program-level scores that were calculated for each session day and represented the average percentage of manualized facilitator activities completed across all components for that day. For analysis, each teacher received a teacher-level score that represented the average of the daily program-level scores for all days the teacher attended. Across all teachers, the average completion of facilitation activities during the days attended was 90% (Range = 78–95%).

Alternative Adherence—Completion of Participant Learning Objectives

Achievement of participant learning objectives for each component on the CARE Daily Session Rating Form was rated on a 0–4 scale (0 = activity/skill did not meet any participant objectives; 4 = all participant objectives were met at an exemplary level). Because this measure was rated on a continuous scale, rater agreement was assessed using inter-class correlations (Hallgren 2012); the level of agreement (ICC of .75) was considered excellent (Cicchetti 1994). Daily program-level scores for “completion of participant learning objectives” were calculated by averaging final consensus scores across components for each program day. For analysis, a teacher-level “completion of participant learning objectives” score was computed for each teacher that represented the mean of the daily program-level scores for all days the teacher attended. Average scores reflected an adequate to exemplary level of participant learning objectives completed ($M = 3.38$, Range = 2.89–3.70).

Facilitation Quality To measure quality of delivery, coders provided ratings each day based on behavioral indicators identified for the modified Iowa Strengthening Families Facilitator Rating Form for each facilitator on ten positive (clarity when

explaining concepts, acceptance and friendliness, level of preparation, managing group dynamics, clarity when explaining activities, comfort level with material, pace of delivery, ability to engage participants, redirecting participants as needed, co-facilitation) and six negative facilitation skills (losing track of time, ineffective answers to questions, being critical of participants, acting distant, reading from manual, providing therapeutic advice/problem solving to individual participants during session). Positive skills were rated on a 0–4 scale where 0 = Facilitator performed very few or none of the behavioral indicators and 4 = Facilitator exemplified all behavioral indicators at all or almost all opportunities. Negative facilitation behaviors were rated based on frequency of occurrence on the same scale in a reverse order (0 = Very frequently to 4 = Never). Interclass correlation ratings for the first three implementations for facilitation skill were excellent (.79) (Cicchetti 1994).

Because CARE is delivered by a facilitation team, to arrive at a program-level facilitation quality score for a given program day, each facilitator's final consensus score was weighted by the amount of content s/he delivered that day (e.g., a facilitator delivering 50% of the daily content would have their facilitation skills score account for 50% of the daily program-level score). Teacher-level scores were calculated for each teacher by taking the average of the daily positive facilitation skills program-level scores and the average of the daily negative facilitation skills program-level scores for all days of the program attended. Across all teachers, scores indicated a high level of positive ($M = 3.74$, Range = 3.47–3.97) and low level of negative facilitation behaviors experienced ($M = 3.88$, Range = 3.59–4.00). Scores for negative and positive facilitation skills were highly correlated (.88); for analysis, teacher-level scores were averaged to form an overall facilitation quality rating ($M = 3.79$, Range = 3.52–3.97).

Participant Responsiveness Measures In addition to collecting attendance records, participant self-assessments were created for the program. These measures assessed teachers' understanding of the material and how useful they felt the information was. Teachers were not compensated for any measures completed during the training; teachers were aware that the self-assessments were reviewed by the facilitators before the next session.

Participant Responsiveness—Understanding of Material

Brief daily participant self-assessments consisted of a series of five to nine questions that measured basic understanding through a series of true/false questions about the topics and activities presented that day (e.g., “I know how to recognize the bodily sensations that arise when I experience unpleasant emotions”). For each program day, a percentage of accurately endorsed questions was computed; for analysis, daily scores were averaged to form a single score. Teachers indicated a high level of understanding of material ($M = 91\%$, $SD = .06$).

Participant Responsiveness—Usefulness of Material A single question assessed the usefulness of the material covered each session day. The question “Overall, I feel like I will use the concepts and activities taught today in my life and my work as a teacher” was rated on a 1 to 5 scale (1 = Strongly Disagree to 5 = Strongly Agree). Teacher daily ratings of the usefulness of the material were averaged to calculate an overall program usefulness score; average usefulness ratings were high ($M = 4.31$, $SD = .46$).

Participant Responsiveness—Attendance (Dosage) Attendance records were collected each program day. Four of the CARE program implementations were provided over five days (6 hours each); two program implementations provided to the wait-list control group were compacted into four days (7.5 hours) in response to time constraints. Dosage was calculated for each teacher as a percentage of days of the program attended relative to the amount provided for a given implementation. This method afforded an assessment of the relative amount of material to which teachers were exposed. On average, teachers were present for 85% of the material ($SD = .20$, Range = 20–100%).

Baseline and Follow-up Participant Self-Report Surveys Teachers completed self-report surveys on average three months before beginning the program and again on average 9.5 months after (see Table 1 for schedule); they were compensated for survey completion during after-school hours equivalent to the district pay rate of \$42/hour.

Baseline Engagement in Mindful Practices Before starting the CARE program, participants were asked to report on how frequently they engaged in meditation (e.g., mindfulness practices, vipassana, zen concentration, guided visualization) or meditative movement practices (e.g., yoga, qi gong, tai chi, martial arts) in the last year (0 = Not currently engaged, to 7 = Engaged multiple times per day). These questions were asked for surveillance only, not as a participant selection mechanism. The two questions were averaged and used as a covariate for analyses. Most teachers (81%) reported no current engagement in mindful practices.

Uptake of CARE Practices—Number of Practices At follow-up, participants were asked to report on their uptake and engagement in mindful awareness practices taught during the program. Teachers responded to the following question “How often do you do this?” on a scale ranging from “0 = I do not do this” to “5 = I do this once a day or more.” There were nine CARE practices queried: three deep breaths, intention setting, breath awareness, caring practice, body scan, mindful walking, mindful eating, using an emotion log, and using guided meditation. A binary score was created for each practice; if the practice was performed on a weekly or greater basis, a value of

one was recorded; for analysis, a sum score was computed for each teacher for all activities performed on a weekly or greater basis. On average, teachers who attended the program engaged in four practices weekly or more at follow-up ($M = 3.87$, $SD = 2.76$).

Uptake of CARE Practices—Length of Practice Participants responded to the question “When you did the following [practice], on average how long did you do it for?”; the length of each practice was endorsed on a scale from 0 = No engagement to 7 = 30 minutes or more; for analysis, the length of each teacher’s longest weekly practice was identified. On average, teachers’ longest practice was reported in the 6–10 minute range ($M = 3.18$, $SD = 2.15$).

Uptake of CARE Practices—Reducing Stress Teachers who reported engaging in at least one practice ($n = 119$) were asked “Are these practices helping with your stress level?”; teachers answered on a five-point Likert scale from 1 = No definitely not to 5 = Yes, definitely. Overall, teachers who were engaged in at least one practice reported that the practices were helping somewhat ($M = 4.03$, $SD = .77$; Range = 3–5).

Data Analysis

We examined relationships between all observer-rated facilitator implementation variables and participant responsiveness variables, first through correlation analysis, and then through hierarchical regression analyses. Finally, we examined direct relationships and interactions hypothesized by Berkel et al. (2011) between adherence, quality, and participant responsiveness on outcomes first through hierarchical regression, and then using the PROCESS software for SPSS (Hayes 2012) when significant interactions were found.

Data Availability Data used in this paper are available from the first author by request.

Results

Correlational Analyses

Pearson correlations were examined among implementation monitoring measures (see Table 2). Correlations among the three observer-rated facilitator implementation variables revealed that facilitation quality and completion of facilitation activities were both significantly positively related to completion of participant learning objectives ($r = .54$ and $r = .57$, respectively); however, they were not significantly related to each other. Exploring participant-focused measures, understanding of the material was significantly positively correlated with completion of

Table 2 Descriptive statistics and correlations for implementation and participant responsiveness

	<i>N</i>	Mean	SD	1	2	3	4	5
Participant responsiveness								
1. Attendance (dosage)	166	.85	.19	1	*	*	*	*
2. Understanding of material	164	.91	.06	.00	1	*	*	*
3. Usefulness of material	164	4.31	.46	.16*	.09	1	*	*
Implementation								
4. Adherence- Facilitation activities	166	.90	.04	.13	.18*	-.15	1	*
5. Adherence- Participant learning objectives	166	3.38	.21	.10	.17*	.00	.57**	1
6. Facilitation quality	166	3.79	.11	-.05	.01	.11	-.11	.54**

p < .01 **, *p* < .05*

facilitation activities (*r* = .18) and completion of participant objectives (*r* = .17), but not with facilitation quality. Teacher endorsement of usefulness of material and program attendance were not significantly correlated with any of the facilitator implementation variables, but were significantly positively correlated with each other (*r* = .16).

Relationships Between Implementation and Participant Responsiveness

A series of hierarchical linear regressions were used to further examine the relationships between facilitation quality and adherence to the curriculum with participant responsiveness variables (see Table 3). Teacher age and baseline mindful practices were used as covariates in all analyses. To address issues with multicollinearity, each facilitator-focused measure was first z-scored, then examined separately with participant responsiveness variables. After controlling for covariates, completion of facilitation activities was positively related to participant self-reported understanding of the material ($\beta = .17, p < .05$); there was a significant R^2 change (.03) for the second model when the facilitation activities completed variable was added, $F(1,160) = 4.91, p < .05$, adjusted $R^2 = .03$ for the overall model. Similarly, completion of participant learning objectives was positively related to participant understanding of the

material ($\beta = .17, p < .05$); there was a significant R^2 change = .03 for the second model, $F(1,160) = 4.76, p < .05$, adjusted $R^2 = .03$ for overall model. No significant relationships were found for adherence measures with attendance or participant-reported usefulness of the material. No significant direct relationships were found between facilitation quality and any participant responsiveness variables.

Relationships Between Implementation and Responsiveness on Uptake of CARE Practices

Table 4 shows the results of a series of hierarchical linear regressions used to examine the effects of each of the facilitator implementation variables on outcomes after accounting for participant responsiveness; each facilitator-focused implementation variable was examined separately to address issues of multicollinearity. We used teacher age, baseline mindful practice, and time elapsed between program implementation and follow-up as covariates for all analyses. First, all covariates were added. Second, all participant responsiveness variables were added. Facilitator implementation variables were added separately in the third step; finally, interaction terms for all of the participant responsiveness variables with the facilitator implementation variable were added in the final step.

Table 3 Relationships between adherence and facilitator quality with participant responsiveness

Model	Variables added	Attendance			Understanding of material				Usefulness of material				
		ΔR^2	B	SE	β	ΔR^2	B	SE	β	ΔR^2	B	SE	β
1	All covariates only	.02				.02				.06**			
	Teacher age		.00	.00	.10	.00	.00	-.14		.01	.00	.22**	
	Baseline mindfulness practice		.02	.02	.07	.00	.01	-.03		.05	.04	.09	
2	Adherence- Facilitation activities	.02 ^t	.03	.02	.14 ^t	.03*	.01	.01	.17*	.02 ^t	-.06	.04	-.14 ^t
2	Adherence- Participant learning objectives	.01	.02	.02	.10	.03*	.01	.01	.17*	.00	.00	.04	-.01
2	Quality of delivery- Facilitation quality	.00	-.01	.02	-.05	.00	.00	.01	.03	.01	.05	.04	.10

p < .01 **, *p* < .05*, *p* < .10^t, $\Delta R^2 = R^2$ change

Table 4 Relationships between implementation, responsiveness, and practice uptake and usefulness

Model	Variables added	# Weekly practices				Length of practice				Practice helping stress			
		ΔR^2	B	SE	β	ΔR^2	B	SE	β	ΔR^2	B	SE	β
1	All covariates only	.22**				.08*				.07*			
	Teacher age		.10	.02	.39**		.04	.02	.21*		.01	.01	.18*
	Baseline mindfulness practice		.72	.24	.23**		.40	.21	.16 [†]		.15	.08	.17 [†]
	Time since training		-.01	.00	-.14 [†]		.00	.00	-.08		.00	.00	-.04
2	All participant responsiveness variables	.05*				.04				.20**			
	Attendance		.84	1.23	.05		.72	1.06	.06		.54	.41	.11
	Understanding of material		2.73	3.92	.05		3.88	3.38	.10		.73	1.26	.05
	Usefulness of material		1.26	.49	.21*		.65	.42	.14		.79	.16	.44**
3	Main effects- Adherence- facilitation activities	.01	.29	.29	.10	.00	.06	.25	.03	.01	.10	.09	.12
4	Interactions- Adherence- facilitation activities	.01				.02				.03			
	X Attendance		.04	.29	.01		.01	.25	.00		-.11	.09	-.12
	X Understanding of material		-.16	.28	-.05		.08	.24	.03		.04	.08	.04
	X Usefulness of material		.26	.23	.09		.30	.20	.13		.10	.07	.13
3	Main effects- Adherence- participant learning objectives	.03*	.52	.24	.18*	.03*	.44	.21	.19*	.02 [†]	.12	.07	.15 [†]
4	Interactions- Adherence- participant learning objectives	.01				.02				.06*			
	X Attendance		.25	.28	.07		-.10	.24	-.04		-.27	.09	-.27**
	X Understanding of material		.05	.23	.02		.27	.20	.12		-.07	.07	-.08
	X Usefulness of material		.11	.23	.04		.04	.20	.02		.00	.07	.01
3	Main effects- Quality of delivery- facilitation quality	.02 [†]	.38	.22	.14 [†]	.03*	.38	.19	.17*	.03*	.16	.07	.20*
4	Interactions- Quality of delivery- facilitation quality	.01				.03				.02			
	X Attendance		.09	.22	.03		.00	.19	.00		-.11	.09	-.12
	X Understanding of material		.21	.24	.07		.43	.21	.18*		-.10	.08	-.11
	X Usefulness of material		.20	.23	.07		-.02	.20	-.01		.01	.08	.01

$p < .01^{**}$, $p < .05^*$, $p < .10^{\dagger}$, $\Delta R^2 = R^2$ change

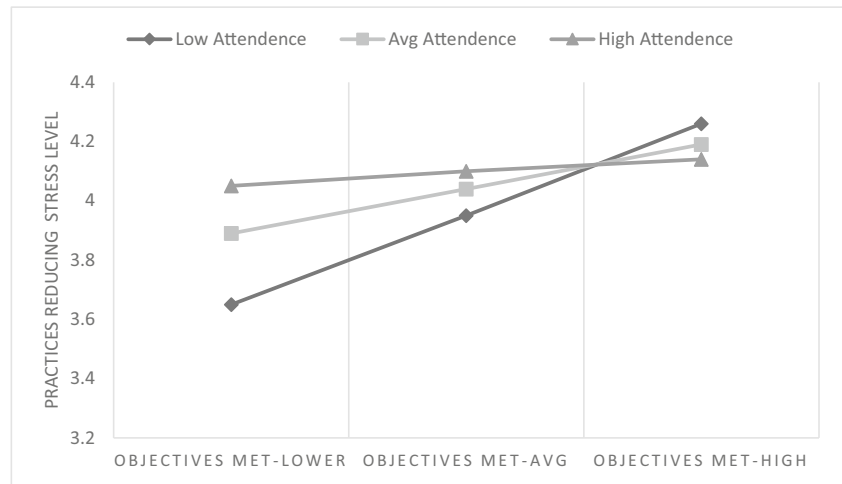
The covariate only model was significant for the number of weekly practices (Adjusted $R^2 = .20$, $F(3,130) = 12.06$, $p < .01$), practice length (Adjusted $R^2 = .055$, $F(3,130) = 3.57$, $p < .05$), and for whether practices reduced teachers' stress level (Adjusted $R^2 = .042$, $F(3,113) = 2.71$, $p < .05$). Teacher age showed a significant positive effect on all outcomes indicating that older teachers were using more practices ($\beta = .385$, $p < .01$), practicing for longer ($\beta = .21$, $p < .05$), and finding their practices more helpful for their stress level ($\beta = .182$, $p < .05$). Teachers' baseline level of mindful practice had a significant positive effect on the number of CARE practices used weekly ($\beta = .231$, $p < .01$). When examining participant responsiveness variables in the second step with covariates, teacher ratings of usefulness of the material showed a significant relationship with the number of weekly practices reported ($\beta = .211$, $p < .05$) and whether practices were helping with stress ($\beta = .443$, $p < .01$).

In the third step, each facilitator implementation variable was examined in a separate model with covariates and all participant responsiveness variables. No significant relationships were found for the traditional measure of

adherence, facilitation activities completed; however, the alternative measure of adherence, completion of participant learning objectives, showed a significant positive relationship with the number of practices teachers were engaging in ($\beta = .179$, $p < .05$), and the length of the practice ($\beta = .194$, $p < .05$). Facilitation quality had a significant positive relationship with the length of practice ($\beta = .172$, $p < .05$) and the extent to which practices were helping with stress ($\beta = .195$, $p < .05$).

In the final step of the model, we examined interactions between each participant responsiveness variable and the targeted facilitator-focused implementation variable; the PROCESS model (Hayes 2012) in SPSS was employed to investigate specific relationships when there was a significant interaction. Although no interaction effects were found for completion of facilitation activities, there was a significant interaction found between the completion of participant learning objectives and participant attendance in relation to whether the CARE practices were helping reduce stress ($\beta = -.26$, $p < .01$). As shown in Fig. 1, when participants had high attendance to the program, no significant relationships were found

Fig. 1 Interaction between attendance and completion of participant learning objectives on the extent to which practices help reduce stress



for completion of learning objectives with teachers’ report of how much practices were helping reduce stress. However, in cases when participants had average or lower levels of attendance, completion of learning objectives was significantly positively related to how much practices helped reduce stress.

Berkel et al. (2011) hypothesized that the impact of facilitation quality on outcomes was mediated by participant responsiveness. Baron and Kenny (1986) state that the initial variable must be correlated with the potential mediator variable to be a candidate for mediation. Because no relationships were found in our first set of analyses (see Table 3), the mediation hypothesis was not explored. These data, however, did show moderation effects of participant understanding of the material for facilitation quality on practice length reported at follow-up ($\beta = .175, p < .05$). Teachers who reported lower understanding of the material showed no difference in their length of practice related to facilitation quality. However, teachers who reported average or higher understanding of the material reported longer practice times when facilitation quality was at an average or higher level (see Fig. 2).

Finally, we tested Berkel et al.’s (2011) hypothesis regarding interactions between adherence and quality on outcomes. We ran analyses with facilitation quality and each of the adherence variables separately for all participant outcomes related to uptake and usefulness of CARE practices; no interaction effects were found. Figure 3 provides a visual model for relationships found between adherence, quality, participant responsiveness, and teacher practice outcomes.

Discussion

The aim of the present study was to elucidate the process of creating implementation monitoring tools for the CARE program, and to explore the relationships between intervention adherence and quality, participant responsiveness, and continued use of CARE practices at follow-up. We found mixed support for the hypotheses presented in the theoretical model proposed by Berkel et al. (2011). Our findings suggest that participant responsiveness plays a role in the relationship

Fig. 2 Interaction between participant understanding of material and facilitation quality on practice length

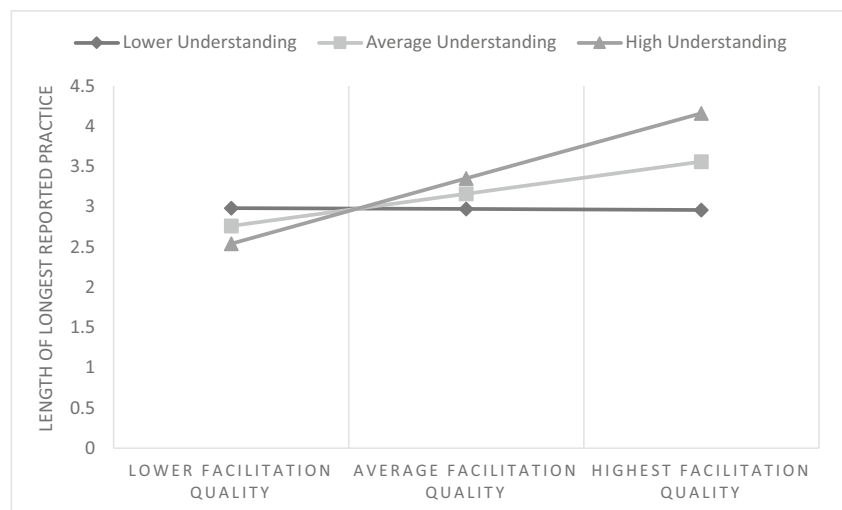
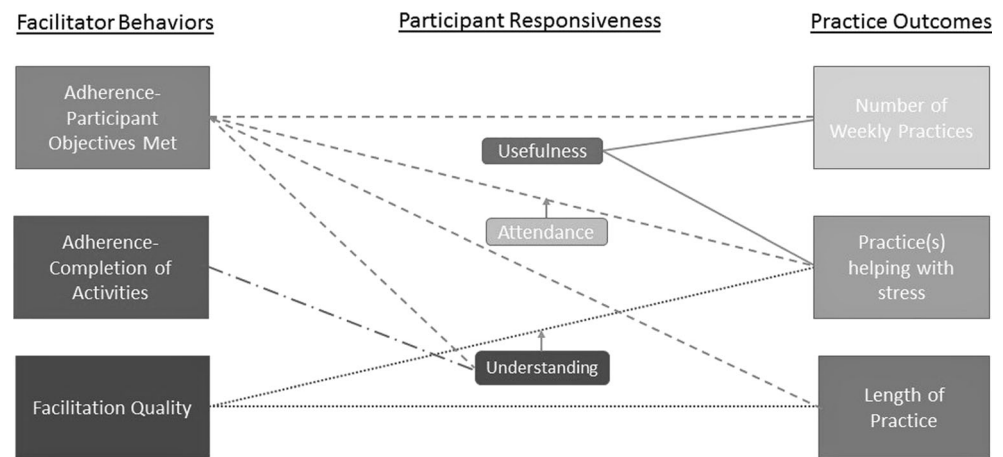


Fig. 3 Relationships between facilitator behaviors, participant responsiveness, and practice outcomes



between facilitation of the program and continued practice; however, these aspects also have independent effects on practice outcomes. Further, the ability to identify relationships between facilitation behaviors and participant responsiveness may depend on measurement.

A unique aspect of this study was the inclusion of “completion of participant learning objectives” as a measure of adherence. Although the completion of facilitation activities did show a significant relationship with participant understanding of material, only completion of participant learning objectives had significant positive relationships with teacher follow-up outcomes, specifically, the number of CARE practices used and how much practices helped reduce teacher stress. These findings suggest that it is the completion of learning objectives rather than completion of facilitation activities directly that more effectively stimulates greater use of practices by participants. This type of measurement may be especially important because programs that allow skilled facilitators to adapt to the needs of participants may complete participant learning objectives without completing 100% of the facilitation activities listed in a manual (e.g., a more advanced audience may not require review of basic definitions). Indeed, the two measures of adherence were only modestly correlated in this study ($r = .57$). That said, it should be noted that our *lowest level* of facilitation activities completed over the six implementations was 78% ($M = 90\%$); this level of completion may have been high enough, and included too little variation to show significant associations with continued practice. When deciding on implementation monitoring strategies needed to promote and measure adherence, it is important to consider who will be doing the facilitating; similar to programs like MBSR, CARE programs were implemented by facilitators with an established mindfulness practice and an advanced degree who had already been through the CARE program. In some newer programs being delivered to youth in schools, teachers with little or no personal practice may be delivering the curriculum. The less foundational knowledge a facilitator has, the more detailed and explicit tools may be needed for successful implementation.

In support of Berkel et al.’s (2011) hypothesis, participant attendance showed an important moderating effect on the extent to which participant learning objectives met was related to practice-driven reductions in teacher stress. For teachers with average or lower attendance, completion of participant learning objectives was significant. The lack of findings for completion of learning objectives for high attenders may be explained because, in the CARE program, information is often repeated over program days. Facilitators also were provided information in between session days about which components were not covered adequately so they could adjust when the topic arose in later sessions. This finding speaks both to the importance of completion of learning objectives and the importance of tracking such information during implementation to assure that content is fully covered.

Instead of mediation as hypothesized by Berkel et al. (2011), our data suggested that participant understanding of the material moderated the relationship between facilitation quality and length of practice. When participants reported lower levels of understanding, facilitation quality was unrelated to practice length; however, at average or higher levels of understanding, an increasing positive relationship between practice length and facilitation quality was found. These findings indicate that participants need to be able to grasp and apply concepts on a basic level before facilitation quality effects practice outcomes. One explanation for this finding may be offered anecdotally from the trainings; given that CARE was offered as an alternative to in-classroom work on some of the days, some teachers may have opted to attend even if they did not plan to truly engage and learn the material. As some school-based mindfulness programs begin to consider whole school approaches, these findings are important to keep in mind.

In addition to relationships between facilitator implementation and teacher practice outcomes, we also found significant positive relationships directly with participant responsiveness. Teacher perceptions of usefulness of practices and activities taught was significantly related to the number of

weekly practices they were engaging in at follow-up and how much practices were helping reduce their stress level; neither adherence nor quality showed a significant relationship with teachers' perceptions of usefulness. It may be that some people, because of their exposure, background, or general orientation, are more likely to see the applicability of information, activities, and practices provided. We did find that increased age was significantly related to higher ratings of usefulness. This area is ripe for future study for MBIs. Overall, these findings highlight the importance of using several different implementation monitoring tools as each offer a unique aspect of the picture for uptake and use of practices.

Using Fidelity Measures for Continued Training and Supervision of Facilitators

Structured implementation monitoring tools, while tedious to create, can be extremely useful for program developers, especially during program expansion. These tools can serve both as a way to measure implementation and as a guide for newer facilitators. Anecdotal information provided by facilitators during the RCT indicated that having information in advance on the “how” and the “why” for each component helped them focus on what was important and implement more in line with the original intention of the developer. Recent research has indicated that facilitators' understanding of implementation fidelity may vary, so having tools that answer both questions can be helpful (Cutbush et al. 2017). Further, having observer ratings of facilitation quality and adherence, and participant responses related to understanding and usefulness of material allowed facilitators to become aware of areas of strength and weakness in their facilitation and to adjust accordingly during implementation. This may explain some of the findings from Durlak and Dupre (2008) showing that just the act of implementation monitoring can improve outcomes. If important activities or concepts were missed or under-emphasized, facilitators could incorporate the missing information into future program days.

For overall program improvement, implementation monitoring data can be shared with program developers in order to provide more extensive feedback to facilitators. Furthermore, by looking over multiple implementations, it is possible to identify common areas of weakness. For example, adherence data was used by program developers to identify areas where multiple facilitators struggled to provide certain topics and/or activities. This information has since been used to make clarifying changes to the facilitator manual and materials and enhance facilitator training.

The CARE tools have been offered as an example for programs that may be new to implementation monitoring. Neither facilitation behaviors (adherence, quality), nor participant responsiveness alone explained teacher uptake and use of

practices. Rather, a complex picture emerged indicating that facilitators play a role in these practice outcomes, but the participants also bring their full selves into the room, with all of their distractions, previous experiences, and motivations. Although some of what facilitators do helps to influence this, there is more left to understand about how to help participants engage with these practices. For established programs, the first step is to have a clear picture of what is being provided during program implementations.

Limitations

Though we report on the development and use of these implementation monitoring tools, they did not undergo rigorous evaluation before use. These types of tools are difficult to evaluate in advance because evaluation can only occur within the context of use; this is a common limitation for such instruments (Schoenwald 2011).

Additionally, our participant self-assessment of understanding used a true/false response for brevity; because of the way some questions were worded (e.g., I know how to apply the 4Rs in working with my emotional triggers), endorsement of the “incorrect” answer could indicate a lack of understanding, or a lack of confidence in their ability to apply it. Further, participant responses were reviewed by facilitators and concerns or misunderstandings were often addressed in later sessions; therefore, scores may not be an accurate reflection of participant understanding by the end of the program; participants also may have altered responses knowing that facilitators were reviewing the information provided. Finally, because we did not deliver a pre-test, it is possible that some teachers already knew the information.

In relation to our measurement of attendance, two of the six program implementations were presented over four days rather than five. This meant that a session day for the 4-day implementations received a heavier weighting than those in the longer program implementations. Additionally, teachers were paid to attend one program day; their schools were compensated for substitutes for two days. This compensation may have affected attendance.

Finally, for the facilitation measures collected, the levels of adherence and quality for the majority of days across all implementations were in the adequate to high range. Results for these analyses may have been different if there were more variability in the program implementation.

Acknowledgements We would like to thank the teachers and schools who participated in this research. Without you, we could not do this work.

Author Contributions SLD: designed and executed the study, performed data analyses, and wrote the paper. PAJ and JLB: assisted with design and execution of current study, assisted with editing of the paper, and designed and executed parent study. DR: assisted with revisions of paper and data analysis. AW, JLF, and CT: assisted with creation of the

measures used in the study. MTG: provided feedback on drafts and helped with design and execution of parent study.

Funding The project described was supported by Grant Award R305A120180 from the Institute of Educational Sciences (IES). The content is solely the responsibility of the authors and does not necessarily represent the official views of the IES or the U.S. Department of Education.

Compliance with Ethical Standards

Ethical Approval All procedures performed involving human participants were in accordance with the ethical standards of the University of Virginia, Pennsylvania State University, and the New York City Department of Education Institutional Review Boards and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all participants in the study.

Conflict of Interest In accordance with ethical obligations, PAJ and CT report that they are the developers of the CARE program and, as such, may benefit from publication of this research. SLJ, DR, AW, JLF, and MG have no conflicts of interest to declare.

References

- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Berkel, C., Mauricio, A. M., Schoenfelder, E., & Sandler, I. N. (2011). Putting the pieces together: an integrated model of program implementation. *Prevention Science, 12*(1), 23–33. <https://doi.org/10.1007/s11121-010-0186-1>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290. <https://doi.org/10.1037/1040-3590.6.4.284>
- Coatsworth, J. D., Duncan, L. G., Nix, R. L., Greenberg, M. T., Gayles, J. G., Bamberger, K. T., et al. (2015). Integrating mindfulness with parent training: effects of the mindfulness-enhanced strengthening families program. *Developmental Psychology, 51*(1), 26–35. <https://doi.org/10.1037/a0038212>
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement, 20*, 37–46. <https://doi.org/10.1177/001316446002000104>
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159 Retrieved from <http://www.bwgriffin.com/workshop/SamplingACohentables.pdf>
- Cutbush, S., Gibbs, D., Krieger, K., Clinton-Sherrod, M., & Miller, S. (2017). Implementers' perspectives on fidelity of implementation: "Teach every single part" or "Be right with the curriculum"? *Health Promotion Practice, 18*(2), 275–282. <https://doi.org/10.1177/1524839916672815>
- Dane, A. V., & Schneider, B. H. (1998). Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review, 18*(1), 23–45. [https://doi.org/10.1016/S0272-7358\(97\)00043-3](https://doi.org/10.1016/S0272-7358(97)00043-3)
- Dobkin, P. L., Hickman, S., & Monshat, K. (2014). Holding the heart of mindfulness-based stress reduction: balancing fidelity and imagination when adapting MBSR. *Mindfulness, 5*(6), 710–718. <https://doi.org/10.1007/s12671-013-0225-7>
- Domitrovich, C., & Greenberg, M. T. (2000). The study of implementation: current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation, 11*, 193–221. https://doi.org/10.1207/S1532768XJEPC1102_04
- Doyle, S. L., Jennings, P. A., DeWeese, A., & Frank, J. (2014). Evaluating the fidelity of the Cultivating Awareness and Resilience in Education (CARE) professional development program. In *22nd Annual Society for Prevention Research Conference*. Washington, D.C.
- 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>
- Durlak, J. A., & Wells, A. M. (1997). Primary prevention mental health programs for children and adolescents: a meta-analytic review. *Annals of Clinical Psychiatry, Wiley-Blackwell*. <https://doi.org/10.1023/A:1024654026646>
- Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health Education Research, 18*(2), 237–256. <https://doi.org/10.1093/her/18.2.237>
- Erdfelder, E., Faul, F., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Feagans Gould, L., Mendelson, T., Dariotis, J. K., Ancona, M., Smith, A. S. R., Gonzalez, A. A., et al. (2014). Assessing fidelity of core components in a mindfulness and yoga intervention for urban youth: applying the CORE Process. *New Directions for Youth Development, 2014*(142), 59–81. <https://doi.org/10.1002/yd.20097>
- Feagans Gould, L., Dariotis, J. K., Greenberg, M. T., & Mendelson, T. (2016). Assessing fidelity of implementation (FOI) for school-based mindfulness and yoga interventions: a systematic review. *Mindfulness, 7*(1), 5–33. <https://doi.org/10.1007/s12671-015-0395-6>
- Greenberg, M. T., Brown, J. L., & Abenavoli, R. M. (2016). Teacher stress and health: effects on teachers, students, and schools. *Edna Bennett Pierce Prevention Research Center, The Pennsylvania State University*, 1–12. Retrieved from <http://www.rwjf.org/en/library/research/2016/07/teacher-stress-and-health.html>
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: an overview and tutorial. *Tutorial in Quantitative Methods for Psychology, 8*(1), 23–34. <https://doi.org/10.20982/tqmp.08.1.p023>
- Hayes, A. F. (2012). PROCESS: a versatile computational tool for observed variable mediation, moderation, and conditional process modeling. *White paper*.
- Iowa State University Extension and Outreach. (2010). Strengthening Families Program for parents & youth 10-14: fidelity observation form. Retrieved from <http://www.episcenter.psu.edu/sites/default/files/SFPParent7ObservationUPDATED08.29.16.pdf>
- Jennings, P. A. (2015). Early childhood teachers' well-being, mindfulness, and self-compassion in relation to classroom quality and attitudes towards challenging students. *Mindfulness, 6*(4), 732–743. <https://doi.org/10.1007/s12671-014-0312-4>
- Jennings, P. A., Snowberg, K. E., Coccia, M. A., & Greenberg, M. T. (2011). Improving classroom learning environments by Cultivating Awareness and Resilience in Education (CARE): results of two pilot studies. *The Journal of Classroom Interaction, 46*(1), 37–48 Retrieved from <http://www.jstor.org/stable/23870550>

- Jennings, P. A., Frank, J. L., Snowberg, K. E., Coccia, M. A., & Greenberg, M. T. (2013). Improving classroom learning environments by Cultivating Awareness and Resilience in Education (CARE): results of a randomized controlled trial. *School Psychology Quarterly*, 28(4), 374–390. <https://doi.org/10.1037/spq0000035>.
- Jennings, P. A., Turksma, C., & Brown, R. (2016). *Cultivating Awareness and Resilience in Education (CARE) professional development program manual*. Garrison: Garrison Institute.
- Jennings, P. A., Brown, J. L., Frank, J. L., Doyle, S., Oh, Y., Davis, R., et al. (2017). Impacts of the CARE for teachers program on teachers' social and emotional competence and classroom interactions. *Journal of Educational Psychology*, 1–19. <https://doi.org/10.1037/edu0000187>.
- Milat, A. J., Bauman, A., & Redman, S. (2015). Narrative review of models and success factors for scaling up public health interventions. *Implementation Science*, 10(1), 113. <https://doi.org/10.1186/s13012-015-0301-6>.
- Moncher, F. J., & Prinz, R. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review*, 11(3), 247–266. [https://doi.org/10.1016/0272-7358\(91\)90103-2](https://doi.org/10.1016/0272-7358(91)90103-2).
- Mowbrey, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: development, measurement, and validation. *American Journal of Evaluation*, 24(3), 315–340. <https://doi.org/10.1177/109821400302400303>.
- Prowse, P. T., Meadows, G., & Enticott, J. (2015). An exploratory study into the effectiveness of fidelity scales in the delivery of Mindfulness-Based Cognitive Therapy. *Mindfulness*, 6(6), 1401–1410. <https://doi.org/10.1007/s12671-015-0412-9>.
- Schoenwald, S. K. (2011). It's a bird, it's a plane, it's...fidelity measurement in the real world. *Clinical Psychology*, 18(2), 142–147. <https://doi.org/10.1111/j.1468-2850.2011.01245.x>.
- Schonert-Reichl, K. A., & Lawlor, M. S. (2010). The effects of a mindfulness-based education program on pre- and early adolescents' well-being and social and emotional competence. *Mindfulness*, 1(3), 137–151. <https://doi.org/10.1007/s12671-010-0011-8>.
- Segal, Z. V., Teasdale, J. D., Williams, J. M., & Gemar, M. C. (2002). The mindfulness-based cognitive therapy adherence scale: inter-rater reliability, adherence to protocol and treatment distinctiveness. *Clinical Psychology & Psychotherapy*, 9(2), 131–138. <https://doi.org/10.1002/cpp.320>.