An Investigation of Motivational Interviewing Proficiency Among Natural Implementers of the Classroom Check-Up

Wendy M. Reinke, K. C. Herman, A. J. Frey & S. Owens

Published online: 26 Nov 2025.

Submit your article to this journal ⧉

Article views: 15

View related articles ⧉

View Crossmark data ⧉

Routledge
Taylor & Francis Group

Check for updates

An Investigation of Motivational Interviewing Proficiency Among Natural Implementers of the Classroom Check-Up

Wendy M. Reinke ⓘ, K. C. Herman, A. J. Frey, and S. Owens

University of Missouri

ABSTRACT

Educational researchers and school-based practitioners are increasingly using Motivational Interviewing (MI) skills as an implementation tool within coaching applications. To date, however, only a few researchers have measured MI skills rigorously, and among those that do comparisons across applications have been difficult to make. In this manuscript, we present MI proficiency data from 114 audio-recorded conversations with 44 unique teachers and 8 natural implementors (i.e. school-based mental health providers), who employed the Classroom Check-up intervention. Mean summary scores and the percentage of the sample that reached proficiency thresholds at the overall (i.e. all recorded sessions), coach-level and session-level (e.g. first interview session, second session) are provided. Results indicate the *Fair* proficiency threshold was reached in most sessions, and by most coaches; and that the *Good* proficiency threshold was reached less frequently. Further, results suggest that MI proficiency varied by session with a higher percentage of coaches reaching proficiency in the first session than subsequent sessions. Implications for practice and future research are discussed.

Inspired by Fixsen and colleagues (2005) ground-breaking publication documenting the challenge of adopting and effectively implementing interventions with fidelity in nearly every discipline and service delivery setting, implementation science is now widely recognized as a multi-disciplinary field of study. Since 2005, challenges adopting and effectively implementing interventions in educational settings have been well documented (Spoth et al., 2013). Implementation science focuses on many factors related to translation, most notably dissemination, adoption, high-fidelity implementation, and sustained use (National Research Council and Institute of Medicine, 2009; Pas et al., 2015). From an implementation science perspective, successful translation of evidence-based practices is facilitated by competency drivers (e.g., strategies and methods for selecting staff, training, coaching, and monitoring fidelity), organizational drivers (e.g., institutional supports such as policies, procedures, data systems, and feedback loops), and leadership drivers (e.g.,

technical and adaptive skills; Fixsen et al., 2005). Consultation and coaching are promising implementation strategies in educational settings, particularly for supporting teachers to implement effective behavioral and instructional practices (Ennis et al., 2020; Owens et al., 2014, Pas et al., 2023; Stormont et al., 2015).

Motivational interviewing (MI) is "a particular way of talking with people about change and growth to strengthen their own motivation and commitment" (Miller & Rollnick, 2023, p. 3). Further, Lyon et al. (2024) identified MI as the predominant method within the consultation and coaching literature to enhance teacher motivation to modify their behavioral and instructional practices. While coaching is considered an implementation strategy, MI is considered an implementation technique that can be infused into different coaching models and frameworks to target a variety of behavioral and instructional practices (Frey, Herman, & Reinke, this issue 2025). Despite the promise and flexibility of MI as an implementation technique, relatively little is known about how well MI is implemented, or how proficient the coach is in using it, in school-based intervention.

A recent scoping review Small et al. (2025) provides a detailed analysis of studies that utilize MI in school settings. Of the 62 school-based studies reviewed, less than half (43.5%) reported data on MI proficiency. Additionally, fewer studies reported on MI proficiency in authentic settings during the intervention. Further, less than 20% ($N = 11$) of those studies used valid and reliable tools to examine how proficiently the coach used MI. Even when MI proficiency has been rigorously measured, the use of different measures, or different indicators or standards within the same measure, make comparisons across school-based applications difficult to ascertain.

Of the 11 studies that used the gold standard measures, 10 are limited in the extent to which they serve as a comparator for future studies designed to compare the MI skill in a new study to the extant literature. For example, two studies (Barnett et al., 2012; Rochat, 2019) utilized the Motivational Interviewing Skills Code (Houck et al., 2011) and one study (Pas et al., 2021) utilized the Motivational Interviewing Sequential Code for Observing Process Exchanges (Hannöver et al., 2013), which do not contain recommended proficiency thresholds or summary scores that facilitate comparison across studies. Seven of the eight studies that utilized the Motivational Interviewing Treatment Integrity Code (MITI) use version 3.0 (Moyers et al., 2010) rather than 4.0 (Moyers et al., 2014). While both versions contain summary scores with associated proficiency thresholds, they only share two summary scores with associated proficiency thresholds (i.e., Percent Complex Reflections and Ratio of Reflections-to-Questions [R:Q]). Three of the ten studies that feature natural implementers, or coaches who were employed by the school district, use the MITI 3.0 and involved interventions that worked directly with middle or high school students. For example, Lyons et al., (2017) evaluated natural

implementors (i.e., coaches) who had been extensively trained to use MI, however they do not report the summary scores or the associated MI proficiency thresholds. Robbins et al., (2012) included two natural implementers (i.e., nurses); on average, their percent complex reflections scores did not reach the *Fair* threshold, and their R:Q ratio reached the *Good* proficiency threshold. Finally, the purpose of Simon and Ward (2014) was to train natural implementors (i.e., academic advisors), rather than assess implementation in authentic practice settings; nevertheless, on average, their percent complex reflections scores reached the *Good* proficiency threshold, and their R:Q ratio did not reach the *Fair* threshold. These findings point to a critical gap in understanding how MI functions in day-to-day school-based coaching environments. While the data is limited regarding the use of similar MI quality metrics and sample size, the available evidence suggests natural implementers, even when extensively trained, struggle to consistently reach even the *Fair* proficiency threshold.

In the most comprehensive analysis of MI proficiency in a school-based application to date, Small et al. (2021) used the MITI (Moyers et al., 2016) to examine between and within coach variability. While Small et al. (2021) included research team members, rather than natural implementors, as coaches and included an application with parents rather than teachers, the study is unique in its contribution to the literature for two reasons. First, it is the only study in the scoping review that utilized the most recent version of the MITI (Moyers et al., 2014). Thus, it contains the four current summary scores (i.e., technical global, relational global, percent complex reflections, and R:Q ratio) and associated MI proficiency thresholds. Additionally, their analysis provided MI proficiency data at the overall, coach-level, and session-level. Thus, the Small et al. (2021) analysis provides a roadmap for the analysis in the current study and is uniquely positioned as a referent. Coaches in Small et al. included 20 professionals who were trained and supervised to implement the homeBase intervention. The coaches were trained using the Motivational Interviewing Training and Assessment system, which included 12 hours of workshop sessions, three individual coaching/supervision sessions with an experienced MI coach, and weekly group supervision meetings (see Frey et al., 2017). Sixty percent of the coaches had master's degrees and forty percent were students enrolled in a master's level Social Work program at the time of the study. Eighty percent of all coaches were school- or community-based social workers, and the rest were teachers. Prior to being trained, less than half had previously participated in MI training; none had experience implementing the homeBase intervention.

At the overall level, across the total 245 sessions, mean scores on the MITI technical global summary score scale exceeded the *Fair* proficiency threshold (e.g., ≥ 3.0) with 97% of the overall sessions meeting the *Fair* proficiency threshold and 60% meeting the *Good* proficiency threshold. In relation to

relational global summary scores, 78% met the Fair proficiency threshold while 59% met the Good threshold. For complex reflections and reflections-to-questions summary scores, both mean scores exceeded the *Good* proficiency threshold and 87% and 60% of sessions, respectively, met *Fair* proficiency thresholds.

At the coach-level, 100% and 89% of the coaches mean scores exceeded the *Fair* proficiency threshold on the technical and relational global summary threshold, respectively, and 28% of the coaches' scores met the *Good* proficiency threshold for both the technical and relational global summary scores. The *Fair* proficiency threshold for complex reflections by coach ranged from 33% to 100%. The reflections-to-questions ratio ranged from a low of 0.1 (e.g., one reflection for every 10 questions) to a high of 3.4 (e.g., 3.4 reflections to each question), with 67% and 61% of the coaches exceeding the *Fair* and *Good* proficiency thresholds, respectively.

Classroom Check-up

The Classroom Check-Up (CCU) is a consultation approach designed to provide support at the classroom level while avoiding common challenges with implementation fidelity in school-based consultation. It is a short-term intervention that uses MI techniques to encourage teachers to engage in the change process. The CCU aims to: (a) strengthen teachers' motivation to continue practices that are crucial for student achievement, (b) decrease teacher–student interactions that may worsen behavioral issues, and (c) enhance the use of effective classroom management and teaching practices that support student competence and success. The CCU is modeled after the Family Check-Up (FCU), a well-established assessment-based intervention used successfully with families of children exhibiting behavioral problems (Dishion & Kavanagh, 2003; Dishion & Stormshak, 2007). Similar to the FCU, the CCU emphasizes the connection between assessment, intervention, and behavioral change. It is grounded in evidence-based theory and guided by the principles and strategies of MI (Miller & Rollnick, 2023).

The CCU has been shown to increase teacher implementation of classroom management strategies, including increased use of both general and behavior-specific praise, and decreased use of reprimands (Reinke et al., 2008). Multiple studies have assessed the effectiveness of the CCU using rigorous single-case research designs and daily, real-time tracking of key variables. In one study, Reinke et al. (2008) used a multiple baseline design across four elementary classrooms to examine changes in teacher and student behaviors. The CCU led to increased use of classroom management strategies by teachers, including more frequent and behavior-specific praise, and fewer reprimands. These improvements in teacher behavior were linked to reductions in student disruptive behavior, and the effects were

sustained over time. Another study by Reinke et al. (2007) used a similar design with three general education teachers and six students with behavior challenges. The study found consistent increases in praise directed at targeted students. Notably, the teachers also extended the use of praise to other students, suggesting that the intervention had positive, class-wide collateral effects. Similarly, Mesa et al. (2005) found that the CCU significantly increased teachers' use of behavior-specific praise, which led to a decrease in disruptive behavior across entire classrooms. In another underpowered randomized control trial with 39 teachers and 617 students, Reinke et al. (2023), found that CCU increased teacher use of behavior-specific praise compared to a control group and showed promise for reducing student disruptive behaviors. In all four of these studies, teachers rated the CCU as highly important, effective, and practical, emphasizing its minimal demands on time, resources, and effort. The CCU has also shown to be adaptable based on the specific intervention implemented or teacher behavior change sought. It has been adapted to improve teacher delivery of social emotional curriculum to students (Reinke et al., 2012) and culturally responsive practices (Bradshaw et al., 2018) as well as to intervene in bullying behavior (Pas et al., 2019). While no formal studies of the CCU with secondary teachers have been conducted to date, *Double Check*, an adapted version of the CCU to support teachers in using culturally responsive classroom practices, has been used extensively in secondary settings with evidence that teachers who received coaching had more proactive behavior management and fewer disruptive behaviors relative to teachers who did not receive coaching (Bradshaw et al., 2018). Based on the promising findings from these studies, the CCU and its online support systems were featured prominently as a recommended coaching practice in a recent IES Practice Guide for Promoting Social and Behavioral Success for Learning in Elementary Schools (Nisar et al., 2022). The Practice Guide highlighted the CCU's integration of MI with a focus on evidence-based dimensions of classroom management in a feasible, effective, and accessible format.

CCU implementation supports

The CCU includes a host of implementation supports designed to ensure high fidelity implementation of the full model including MI techniques. These supports include a structured interview guide that prompts open-ended questions and reminds coaches to summarize and reflect teacher comments and structured feedback and goal-setting tools that prompt MI-related questions, the creation of a menu of options, and frequent summaries, reflections, and affirmations. Access to these CCU supports is currently free to anyone via the CCU website (www.classroomcheckup.org).

The role of MI in the CCU

Despite the growing body of evidence supporting the CCU in improving teacher practices and student behaviors, little is known about the role of MI in CCU implementation. As noted, MI is a core component of CCU coaching. However, other active ingredients of the CCU include relationship building strategies, structured agendas and goal setting, and performance feedback (see Reinke et al., 2011). Documenting CCU coach use of MI skills during their consultation meetings would advance knowledge about this aspect of CCU implementation.

Additionally, as noted previously, the limited knowledge base about MI implementation in schools has been in the context of research studies that deployed highly trained external coaches, rather than natural implementers (i.e., school practitioners rather than trained researchers). Missing from the current literature is a description of MI proficiency of coaches employed in school settings. Demonstrating that natural implementers can be trained to deliver MI proficiently could add to the promise of widescale implementation of MI and the CCU.

Current study

The purpose of this study was to examine mean MITI summary scores and the number and percentage of the sample that reached *Fair* and *Good* MI proficiency thresholds at the overall- (i.e., all recorded sessions), coach-level and session-level (e.g., first interview session, second session) among natural CCU implementers. Given the extensive CCU implementation supports, some training in MI techniques, and substantial experience implementing the CCU, we hypothesized that most of the sessions and coaches (i.e., natural implementers) would meet the *Fair* MI proficiency threshold for all summary scores.

Methods

Procedures

Natural implementer coaches were school-based mental health providers working across six school districts implementing a comprehensive school mental health model in one county in the mid-West. Five of the school districts were located in rural settings and one was located in a suburban setting. These individuals were school psychologists, social workers, and counselors working in school buildings and providing the CCU to teachers. The CCU is commonly used by these natural implementers and it is part of their role in the comprehensive school mental health model. Teachers worked in these school districts and received the CCU. The

study was approved by the Institutional Review Board at the University. Prior to beginning the CCU process, both the coach and teacher consented to have CCU sessions recorded and received compensation for participating in the study. The coach was responsible for recording the session via Zoom. Each recorded session was uploaded to a secure folder on SharePoint. The study occurred across two school years post the COVID-19 pandemic.

Classroom Check-up

The CCU consists of multiple sessions. During the first session, the coach conducts a "getting to know you interview," asking about prior experiences with consultation and discussion of the teacher's classroom management style. The coach also conducts a values card sort activity in which the teacher identifies three values that they feel are the most important to their teaching. Following the first interview session, the coach conducts at least two classroom observations, gathering data on the teacher's use of opportunities to respond, praise, and reprimands and student disruptive and aggressive behavior. The coach also observes the physical layout of the classroom, looking to see if there are posted rules, a consistent schedule, and whether expectations for behaviors and routines are clear. Following these observations the coach summarizes the data using the CCU feedback form that guides the feedback conversation, which occurs during session 2. In some cases, the coach and teacher create an action plan in session 2, but depending on the time allotted, often action planning occurs in session 3. Action planning involves the teacher identifying an area of which to focus (e.g., increase use of behavior-specific praise), viewing possible strategies from the CCU menu of options, identifying one or more strategies to implement in the classroom, and planning when and how to do so. In some instances, but with less frequency than other sessions, a fourth session may occur if the coach and teacher are not able to fully complete planning in the third session. Following feedback and action planning, the coach arranges to observe the classroom and provide ongoing support in implementing the selected strategies in the classroom.

CCU training

Each of the coaches received at least three hours of training in the CCU procedures and use of MI techniques. Trainings occurred each school year during a professional training day, in person with a certified CCU trainer. The start of the training covered the overall procedures for conducting the CCU and use of the CCU website. The remaining training focused on conducting the values card sort, the use of MI with modeling and practice opportunities embedded throughout the training (i.e., half of training focused on MI). The natural implementers in this study may have attended more than one CCU

training, but not more than once a year. We did not track whether coaches attended more than one CCU training. No ongoing supervision was provided to coaches.

Participants

Natural implementer CCU coaches ($n = 8$) were working as school-based mental health clinicians in schools. Schools had one CCU coach per building. One coach was a White male, two were Black females, and five were White females. Unfortunately, we did not gather information on their years of experience in schools, but all had been in schools for at least three years at the time of the study. All had a master's degree, one also had a doctorate and two were working toward a doctoral degree. The teachers in the study were all White. Twenty-three percent ($n = 10$) were male, and 77% ($n = 34$) were female. A total of five teachers were from secondary schools (one high school, four middle school) and the other 39 teachers were from elementary schools. At the time of this study, early career elementary teachers were also being recruited to participate in an efficacy trial of the CCU using natural implementers. These teachers and coaches were also eligible, but not required, to participate in this study. Thus, a larger sample of elementary teachers participated.

Measures

We used the MITI 4.2 to assess the MI proficiency of coaches (Moyers et al., 2014, 2016). The MITI enables examination of 4 global dimensions, and 10 behavior counts. The four global dimensions, which are rated on a 5-point scale, include cultivating change talk (CCT), softening sustain talk (SST), partnership, and empathy. High scores on the global ratings indicate higher proficiency than lower scores. For example, the highest anchor for CCT indicates the coach or practitioner "shows a marked and consistent effort to increase the depth, strength, or momentum of the client's language in favor of change" (p. 5). The behavior counts document the frequency of MI skills (e.g., questions, complex reflections, simple reflections, affirmations). The global dimensions and behavior counts are combined to yield four summary scores indicative of MI proficiency. The relational global and technical global summary scores are the mean ratings of the partnership and empathy and CCT and SST dimensions, respectively. The percent of complex reflections (CR) summary score is calculated by dividing CR by total reflections (e.g. simple reflections plus CR) and the ratio of reflections summary score is the ratio of total reflections to the number of questions. Moyers et al. (2014) recommend coders review a randomly generated 20-minute audio segment.

The MITI manual establishes two-levels of proficiency for each of the four summary scores, referred to as *Fair* and *Good* MI proficiency thresholds. For the relational global summary score, greater than or equal to 3.5 indicates *Fair* MI proficiency and scores greater than or equal to 4 indicate *Good* MI proficiency. Thresholds for the technical global summary score are greater than or equal to 3 meet the *Fair* proficiency threshold and scores of 4 or greater meet the *Good* proficiency threshold. For the percent of CR summary score, 40% CRs or greater meets the *Fair* proficiency threshold and 50% or above meets the *Good* proficiency threshold. Finally, for the reflections-to-questions summary score, are a 1:1 ratio for *Fair* proficiency threshold and a 2:1 ratio or higher for the *Good* proficiency threshold (Moyers et al., 2014). Importantly, *Fair* and *Good* proficiency thresholds are based on expert opinion.

Monitoring MI proficiency

Three coders completed a two-day training on the MITI 4 and participated in ongoing group coding until reaching 90% reliability on behavior counts and 100% reliability on global scores. Inter Rater Reliability (IRR) checks were completed on a random sample of 14% of sessions. We assessed IRR via 2-way mixed effects, absolute agreement, average-measures ICCs. We used Koo and Li (2016) benchmarks to categorize the quality of the ICC in which < 0.5 is considered poor reliability, 0.5 – 0.74 is considered moderate reliability, 0.75 – 0.9 is considered good reliability, and > 0.9 is considered excellent reliability. IRR was moderate for CCT (.736), and *empathy* (.692) and good for *partnership* (.861) and *SST* (.873). For technical global scores (.632) and reflections-to-question ratio (.521), reliability was moderate. Relational global scores (.879) IRR were good. For percent CRs (.063) reliability was poor.

Statistical analysis

The overall mean MITI summary scores were calculated for all recorded sessions. Next, the mean MITI summary scores were calculated separately for each coach. Independent sample t-tests were conducted to determine if coaches who met the *Good* proficiency threshold scores across all the MITI summary scores for at least one session differed compared to coaches who did not meet this criterion. Next, mean MITI summary scores were calculated across sessions (e.g., mean scores for session 1, 2, 3, and 4). Independent sample t-tests were conducted to determine if particular sessions differed from all others on MITI summary scores.

Results

In total, 8 coaches delivered 114 sessions with 44 teachers. Coaches delivered between one and four sessions per teacher. Thirty-nine percent of the sessions were the first session in which coaches conducted the "getting to know you interview" ($n = 44$), 37% ($n = 42$) were session two which included feedback to the teacher, 22% ($n = 25$) consisted of session three, and 3% ($n = 3$) were an additional fourth session. Over half of teachers (52%) received three sessions.

Overall MI proficiency

The top row in Table 1 summarizes the mean scores across the 114 sessions for the four MITI summary scores. Across the 114 sessions, 100% met the *Fair* proficiency threshold for the technical global summary score. On average, the relational global summary scores met the *Fair* proficiency threshold, with 94% of sessions meeting this threshold. For CR and reflections-to-questions summary scores, 90% and 62% of sessions, respectively, met the *Fair* MI proficiency thresholds. *Fair* proficiency thresholds were met on all four MITI scores for 62 (54%) of the sessions. With regard to meeting the *Good* proficiency thresholds, 4 (3.5%) sessions met criteria across all four MITI summary scores, 24 (21%) met three out of the four, 49 (43%) met two of the four scores, 29 (25%) met for one of the MITI summary scores, and only 8 (7%) did not meet any of the *Good* proficiency threshold scores.

Coach-level MI fidelity

Mean scores and proficiency threshold status by coach are provided in Table 1. All eight coaches had at least one session (range = 1–14) in which they met the *Fair* MI proficiency threshold across all four summary scores. Four of the 8 coaches (50%) had one session in which they met the *Good* MI proficiency threshold across all four summary scores. There were no statistically significant difference across the four summary scores compared to those who did not.

Session-level MI fidelity

See Table 2 for session-level MI proficiency summary scores. When investigating the MI skills proficiency across sessions, the mean score for technical global summary score proficiency was 3.61 (0.30) for session one, 3.64 (0.32) for session two, 3.70 (0.35) for session three, and 3.67 (0.58) for session four. With regard to the relational global summary score, the mean score for session one was 3.89 (0.43), 3.65 (0.52) for session two, 3.58 (0.53) for session 3, and

Table 1. Distribution of miti proficiency scores, overall and by coach.

Coach	# of teachers	# of sessions	Technical Proficiency Scores M(SD)	Global cutoffs % Fair	% Good	Relational Proficiency Scores M(SD)	Global cutoffs % Fair	% Good	Complex Reflections Scores M(SD)	Global cutoffs % Fair	% Good	R:Q Ratio Scores M(SD)	Global cutoffs % Fair	% Good
All	44	114	3.64 (0.32)	100	38	3.70 (0.54)	94	56	65.34 (20.70)	90	82	1.35 (1.26)	62	13
1	9	26	3.65 (0.31)	100	31	3.63 (0.70)	100	54	67.81 (20.16)	89	81	1.19 (1.13)	50	15
2	10	23	3.80 (0.25)	100	61	3.76 (0.33)	100	52	52.46 (22.89)	78	70	1.15 (0.48)	61	4
3	9	21	3.71 (0.30)	100	48	3.73 (0.46)	100	57	62.21 (19.18)	91	76	1.04 (0.63)	43	5
4	7	15	3.67 (0.31)	100	40	3.87 (0.35)	100	73	71.99 (13.10)	100	93	1.81 (0.65)	93	33
5	6	14	3.46 (0.31)	100	14	3.79 (0.54)	93	79	75.85 (17.37)	100	100	2.00 (2.91)	79	21
6	5	10	3.45 (0.37)	100	20	3.40 (0.61)	80	20	71.81 (21.58)	100	80	1.38 (0.40)	80	0
7	2	3	3.33 (0.29)	100	100	3.67 (0.76)	100	33	63.64 (28.91)	67	67	0.89 (0.43)	33	0
8	1	2	3.50 (0.71)	100	50	3.25 (1.06)	50	50	61.00 (15.56)	100	100	1.54 (1.36)	50	50

Technical proficiency cutoffs: Fair ≥3.0, Good ≥4.0; Relational proficiency cutoffs: Fair ≥3.5, Good ≥4.0; Complex reflections cutoffs: Fair ≥40%, Good ≥50%; R:Q ratio cutoffs: Fair ≥1:1 (e.g., ≥ 1.0), Good ≥2:1 (e.g., ≥ 2.0).

Table 2. Distribution of miti proficiency scores by session.

Session #	Session Type	Total Sessions	Technical Proficiency			Relational Proficiency			Complex Reflections			R:Q Ratio		
			Scores	Global cutoffs		Scores	Global cutoffs		Scores	Global cutoffs		Scores	Global cutoffs	
			M(SD)	% Fair	% Good	M(SD)	% Fair	% Good	M(SD)	% Fair	% Good	M(SD)	% Fair	% Good
1	Interview	44	3.61 (0.30)	100	32	3.89 (0.43)	98	75	71.26 (13.68)	98	93	1.30 (0.65)	68	11
2	Feedback	42	3.64 (0.32)	100	33	3.65 (0.52)	93	48	65.47 (22.84)	93	79	1.14 (0.54)	55	12
3	Planning	25	3.70 (0.35)	100	52	3.58 (0.53)	96	56	56.61 (24.65)	72	72	1.82 (2.41)	64	20
4	Planning	3	3.67 (0.58)	100	67	2.67 (0.76)	33	0	49.44 (9.24)	100	33	1.33 (0.65)	67	0

Technical proficiency cutoffs: Fair ≥3.0, Good ≥4.0; Relational proficiency cutoffs: Fair ≥3.5, Good ≥4.0; Complex reflections cutoffs: Fair ≥40%, Good ≥50%; R:Q ratio cutoffs: Fair ≥1:1 (e.g., ≥ 1.0), Good ≥2:1 (e.g., ≥ 2.0).

2.67 (0.76) for session 4. For CR the mean score was 71.26 (13.68) for session one, 65.47 (22.84) for session two, 56.61 (24.65) for session three, and 49.44 (9.24) for session four. The average reflections-to-questions summary scores were 1.29 (0.65) for session one, 1.13 (0.54) for session two, 1.82 (2.41) for session three, and 1.33 (0.65). The mean relational global summary score for session one was significantly higher when compared to all other sessions, t (112) = −3.02, p = .003, and the use of CR was also significantly higher, $t(112)$ = −2.48, p = .015. No other significant differences were found when comparing sessions scores.

Discussion

Consistent with our hypotheses, we found that natural implementer CCU coaches delivered the intervention with acceptable levels of MI proficiency, as defined by the MITI. For example, the vast majority of sessions met the *Fair* proficiency threshold for the technical (100%) and relational (94%) global thresholds, and CR summary scores (90%); over half (63%) exceed the *Fair* threshold for reflections-to-questions. Additionally, over half (54%) of the sessions exceeded the *Fair* proficiency threshold for all four indicators. The mean summary scores for all four summary indicators met the *Fair* threshold but fell short of the *Good* threshold. Further, all eight coaches met thresholds for *Fair* proficiency in MI technical skills. Nearly all sessions also met *Fair* proficiency thresholds for MI relational skills and over half of the coaches met the *Fair* proficiency threshold for all four summary scores. With the exception of global relational scores and use of CR's which were higher in session 1 than sessions 2–4, the sessions generally met *Fair* proficiency thresholds. The difference in relational global scores and CRs in session 1 is likely due to the nature of the session. It is the first meeting between the teacher and coach, and the session is guided by a "get to know you" interview that prompts coaches to reflect and summarize, making it more likely that these MI skills occur. Alternatively, the novelty of the first session or rating challenges with CRs could have influenced scores.

MI performance of natural implementers in this study was comparable to that observed in Small et al. (2021) using research team implementers. While other school-based MI intervention researchers have used the MITI (Small et al., 2025), they have not used the MITI proficiency thresholds and the four summary scores as the primary criteria for determining skill. Similar to Small et al. (2021), the present study revealed overall-level variability: 54% of the sessions met the *Fair* threshold on all four summary score indicators, versus 46% in the Small study. However, only 4% of the sessions in the present study met the *Good* proficiency threshold across all four global MITI indicators, versus 16% of the sessions in the Small study. As a whole, the present study sample performed similarly in technical and relational MI global scores with

relational being slightly higher (unlike Small et al., (2021), where technical was higher). CRs were the strongest global score in the present study whereas the R:Q ratio was the lowest summary. Compared to Small et al. (2021), the present study revealed far less coach-level variability. While a higher percentage of natural implementer coaches met *Fair* proficiency than was the case for the Small study, fewer met the *Good* proficiency threshold.

It is important to note that in addition to the coach circumstance, natural versus research team implementers, coaches in the present study differed from those in the Small et al. (2021) study in other ways. First, all coaches in the present study had a master's degree and several years of experience in school settings compared to the Small study which included coaches with a master's degree (60%) or enrolled in a master's degree training program (40%). Second, all coaches in the present study had participated in prior MI trainings compared to only 50% of the coaches in the Small study. Third, the coaches in the present study received significantly less direct workshop training (3 vs. 12 hours), individualized supervision (0 vs. 3 sessions) during the course of the study compared the Small coaches. Thus, while it is important that natural implementers were able to deliver MI with *Fair* proficiency with minimal training and support in the present study, prior training and experiences likely supported their successful implementation skills. Additionally, while natural implementers performed as well on average as research team implementers in the Small study, some coaches in the Small study also achieved higher levels of proficiency. This effect may be due to the more extensive training and support during the study offered to these coaches compared to those offered to the natural implementers in the present study.

Limitations

The present study has many strengths to note. First, this was the first study of the MI skills using the CCU and with natural implementers. Second, the study used a rigorous, high-quality measure of MI skills. Third, the use of proficiency thresholds and summary scores allowed for meaningful comparison with the study by Small et al. (2021). Fourth, the study provided a good snapshot of the level of proficiency one might expect when relying on a group of experienced school professionals who receive some training in MI and are implementing an intervention designed to encourage use of MI skills.

However, the study also has limitations. First, we lacked information on the years of experience of the coaches and teachers. We know that the coaches were all master's trained, have substantial school experience, and many have been implementing the CCU and been exposed to multiple, MI trainings – all likely of short duration. Second, neither the MITI nor the literature base provides guidelines for which MITI indicators represent MI proficiency, and

the proficiency thresholds (*Fair* and *Good*) for the summary scores are based on expert opinion and also established in the context of substance abuse therapy. In the absence of a single indicator of MI proficiency, we used the four summary scores provided in in the MITI manual. Third, the study relied on relatively few coaches (8) to fully capture MI implementation variability in natural contexts. The coaches and teachers volunteered and were compensated to participate and record their sessions, making their motivation to participate a limitation when interpreting findings and underscores the importance of studying MI delivery in real-world, non-volunteer contexts. In addition, the number of sessions varied by session number. For instance, session 4 only had three occurrences. Despite this we thought investigating difference between scores on sessions was warranted given the difference in the types of activities occurring during these sessions. Yet, these findings can only be considered preliminary in nature and future research is warranted with a larger sample. Lastly, CRs were subject to lower inter-rater reliability, which limits confidence in session-level comparisons on this subscale. Future studies should consider enhanced coder training or complementary assessment methods.

Implications

The primary implication of the present study is that in the context of the CCU, natural implementers with prior exposure to MI and schoolwork experience can deliver MI with *Fair* proficiency even with minimal training. This suggests that MI is a promising approach for widespread dissemination in school coaching applications. The study also supports the use of the CCU as a conduit to the dissemination of MI in schools. The CCU includes embedded prompts and structures to support MI implementation. These implementation supports may be a key factor in MI fidelity in the context of minimal to modest training and supervision. We suspect natural implementers with similar MI and experience implementing a structured intervention guided by the principles and strategies of MI, would reach similar levels of MI proficiency.

Additionally, findings suggest there are areas for improvement needed to support coach attainment of *Good* proficiency standards. Because most coaches, on average, exceeded the *Fair* threshold on all four summary scores but did not reach the Good proficiency threshold level, there is room for improvement. More intensive training, coaching, and supervision may be needed to attain the highest levels of MI proficiency. One area for improvement noted in the present study was to increase the R:Q ratio of CCU coaches. Still research is needed to determine if these improvements in MI proficiency from *Fair* to *Good* improve teacher or student outcomes. Little is currently known about MI proficiency thresholds in school settings where implementation often looks very different than in traditional therapy context where most prior MI fidelity research has occurred. For instance, MI consultation meetings in school are often brief

compared to the typical therapy session hour; teachers have different motives for seeking support than clients who enter psychotherapy; and teacher–coach relationships also differ in important ways from client-therapist relationships (see Herman et al., 2021). All of these differences may have implications for understanding how MI effectiveness might change in the school context.

Future research

The study offers guidance for future research. It will be important to replicate findings across school-based applications of MI to understand the proficiency levels of school-based practitioners. As noted previously, the CCU incorporates explicit MI structure, prompts, and implementation supports that may enable higher levels of MI proficiency. Replicating the findings with CCU coaches as well as with coaches of other school-based interventions would help disentangle how much these implementation supports increase MI compliance versus interventions without such supports. As more data is collected, we will be able to examine MI proficiency in the context of desired outcomes and eventually learn effects at the session level.

It is imperative to begin understanding how the MITI global indicators, behavior counts, and summary scores are related to process measures such as satisfaction with the coaching procedures, teacher motivation, and coach-teacher alliance. Future research might add in qualitative interviews with teachers and coaches to gain additional insights. It is also necessary to begin to understand how the MITI global indicators, behavior counts, and summary scores are related to teacher outcomes (e.g., teacher commitment language, teacher instructional practices) and student outcomes. Comparative research such as this will help confirm or refine the current threshold standards to match implementation in school contexts.

Conclusion

This study examined MI skills of natural implementers of the CCU intervention, demonstrating coaches with limited MI training and experience implementing the CCU implemented MI with skill, according to established MI proficiency norms. Additional research is needed to better understand the relationship between MI proficiency scores and process measures (e.g., satisfaction with the coaching procedures, teacher motivation, and coach-teacher alliance), teacher outcomes (e.g., teacher commitment language, teacher instructional practices) and student outcomes.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

Notes on contributors

Wendy M. Reinke is a Curators' Distinguish Professor at the University of Missouri. Her primary research interest includes teacher consultation, school mental health, and prevention of disruptive behaviors in youth.

K.C. Herman is a Curators' Distinguish Professor at the University of Missouri. His primary research interests include school mental health screening and intervention, teacher wellbeing, and developmental psychopathology.

A.J. Frey is a professor at the University of Missouri. His primary interest areas include school mental health services, early intervention, and applications of motivational interviewing in school settings.

S. Owens is an Associate Teaching Professor at the University of Missouri. Her primary research interests include teacher implementation of mental health interventions, school mental health, and prevention-based frameworks in schools.

ORCID

Wendy M. Reinke 🔗 http://orcid.org/0000-0001-7235-624X

References

Barnett, E., Sussman, S., Smith, C., Rohrbach, L. A., & Spruijt-Metz, D. (2012). Motivational interviewing for adolescent substance use: A review of the literature. *Addictive Behaviors*, *37* (12), 1325–1334. https://doi.org/10.1016/j.addbeh.2012.07.001

Bradshaw, C. P., Pas, E. T., Bottiani, J. H., Debnam, K. J., Reinke, W. M., Herman, K. C., & Rosenberg, M. S. (2018). Promoting cultural responsivity and student engagement through double check coaching of classroom teachers: An efficacy study. *School Psychology Review*, *47*(2), 118–134. https://doi.org/10.17105/SPR-2017-0119.V47-2

Dishion, T. J., & Kavanagh, K. (2003). *Intervening in adolescent problem behavior: A family-centered approach*. Guilford Press.

Dishion, T. J., & Stormshak, E. A. (2007). *Intervening in children's lives: An ecological, family-centered approach to mental health care*. American Psychological Association. https://doi.org/10.1037/11485-000

Ennis, R. P., Royer, D. J., Lane, K. L., & Dunlap, K. D. (2020). The impact of coaching on teacher-delivered behavior-specific praise in pre-K-12 settings: A systematic review. *Behavioral Disorders*, *45*(3), 148–166. ERIC. https://doi.org/10.1177/0198742919839221

Fixsen, D. L., Naoom, S. F., Blase, K. A., Freidman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. Louis de la Parte Florida Mental health Institute.

Frey, A., Herman, K. C., & Reinke, W. M. (2025). Motivational interviewing applications in educational consultation: Introduction to the special issue. *Journal of Educational and Psychological Consultation*.

Frey, A. J., Lee, J., Small, J. W., Walker, H. M., & Seeley, J. R. (2017). Motivational interviewing training and support for school-based applications. *Emotional & Behavioral Disorders in Youth*, *17*(4), 86–92.

Hannöver, W., Blaut, C., Kniehase, C., Martin, T., & Hannich, H. J. (2013). Interobserver agreement of the German translation of the Motivational Interviewing Sequential Code for Observing Process Exchanges (MI-SCOPE;D). *Psychology of Addictive Behavior*, *27*(4), 1196–1200. https://doi.org/10.1037/a0033041

Herman, K. C., Reinke, W. M., & Frey, A. (2021). *Motivational interviewing in schools: Strategies for engaging parents, teachers, and students* (2nd ed.). Springer.

Houck, J. M., Moyers, T. B., Miller, W. R., Glynn, L. H., & Hallgren, K. A. (2011). *Elicit coding manual: Motivational interviewing skill code (MISC) 2.5*. Center on Alcoholism, Substance Abuse, and Addictions, The University of New Mexico.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lyon, A. R., Connors, E. H., Lawson, G. M., Nadeem, E., & Owens, J. S. (2024). Implementation science in school mental health: A 10-year progress update and development of a new research agenda. *School Mental Health*, *16*(4), 1013–1037. ERIC. https://doi.org/10.1007/s12310-024-09731-0

Lyons, M. D., Jones, S. J., Smith, B. H., McQuillin, S. D., Richardson, G., Reid, E., & McClellan, A. (2017). Motivation coaching training for instructional coaches: A pilot study of motivational interviewing skills training. *Mentoring and Tutoring: Partnership in Learning*, *25*, 548–565. https://doi.org/10.1080/13611267.2017.1415796

Mesa, J., Lewis-Palmer, T., & Reinke, W. (2005). Providing teachers with performance feedback on praise to reduce student problem behavior. *Beyond Behavior*, *15*(1), 3–7.

Miller, W. R., & Rollnick, S. (2023). Motivational interviewing: Helping people change and grow (4th ed.). The Guilford Press.

Miller, W. R., & Rollnick, S. R. (2023). *Motivational interviewing: Helping people change (applications of motivational interviewing)* (4th ed.). The Guilford Press.

Moyers, T. B., Manuel, J. K., & Ernst, D. (2014). *Motivational interviewing treatment integrity coding manual 4.2.1* [Unpublished manual].

Moyers, T. B., Rowell, L. N., Manuel, J. K., Ernst, D., & Houck, J. M. (2016). The motivational interviewing treatment integrity code (MITI 4): Rationale, preliminary reliability and validity. *Journal of Substance Abuse Treatment*, *65*, 36–42. APA PsycInfo. https://doi.org/10.1016/j.jsat.2016.01.001

Moyers, T., Martin, T., Manuel, J., Miller, W., & Ernst, D. (2010). *The Motivational Interviewing Treatment Integrity (MITI) Code Manual: Version 3.0*. University of New Mexico, Center on Alcoholism, Substance Abuse & Addictions. http://casaa.unm.edu/download/miti.pdf

National Research Council and Institute of Medicine. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. National Academies Press.

Nisar, H., Elgin, D., Bradshaw, C., Dolan, V., Frey, A., Horner, R., Owens, J., Perales, K., & Sutherland, K. (2022). *Promoting social and behavioral success for learning in elementary*

schools: Practice recommendations for elementary school educators, school and district admin- istrations, and parents. 2M Research Services. Contract No. 92990019F0319.

Owens, J. S., Lyon, A. R., Brandt, N. E., Warner, C. M., Nadeem, E., Spiel, C., & Wagner, M. (2014). Implementation science in school mental health: Key constructs in a developing research agenda. School Mental Health: A Multidisciplinary Research and Practice Journal, 6 (2), 99–111. APA PsycInfo. https://doi.org/10.1007/s12310-013-9115-3

Pas, E. T., Borden, L., Herman, K., & Bradshaw, C. P. (2021). Leveraging motivational interviewing to coach teachers in the implementation of preventive evidence-based prac- tices: A sequential analysis of the motivational interviewing process. Prevention Science, 22 (6), 786–798. Academic Search Premier. https://doi.org/10.1007/s11121-021-01238-3

Pas, E. T., Bradshaw, C. P., Becker, K. D., Domitrovich, C., Berg, J., Musci, R., & Ialongo, N. S. (2015). Identifying patterns of coaching to support the implementation of the good behavior game: The role of teacher characteristics. School Mental Health, 7(1), 61–73. https://doi.org/ 10.1007/s12310-015-9145-0

Pas, E. T., Kaiser, L., & Owens, J. S. (2023). Innovative Approaches to Coaching Teachers in Implementing Tier 1 and Tier 2 Classroom Interventions. In S. W. Evans, J. S. Owens, C. P. Bradshaw, & M. D. Weist (Eds.), Handbook of School Mental Health. Issues in Clinical Child Psychology. Springer. https://doi.org/10.1007/978-3-031-20006-9_23

Pas, E. T., Waasdorp, T. E., & Bradshaw, C. P. (2019). Coaching teachers to detect, prevent, and respond to bullying using mixed reality simulation: An efficacy study in middle schools. International Journal of Bullying Prevention, 1(1), 58–69. https://doi.org/10.1007/s42380- 018-0003-0

Reinke, W. M., Herman, K. C., & Copeland, C. (2023). Student engagement: The importance of the classroom context. In A. Reschly & S. Christenson (Eds.), The Handbook of Research on Student Engagement (2nd ed., pp. 529–544). New York: Springer.

Reinke, W. M., Herman, K. C., Darney, D., Pitchford, J., Becker, K., Domitrovich, C., & Ialongo, N. (2012). Using the classroom check-up model to support implementation of paths to pax. Advances in School Mental Health Promotion, 5(3), 220–232. https://doi.org/10. 1080/1754730X.2012.707441

Reinke, W. M., Herman, K. C., & Sprick, R. (2011). Motivational interviewing for effective classroom management: The classroom check-up. Guilford Press.

Reinke, W. M., Lewis-Palmer, T., & Martin, E. (2007). The effect of visual performance feedback on teacher use of behavior-specific praise. Behavior Modification, 31(3), 247–263. https://doi.org/10.1177/0145445506288967

Reinke, W. M., Lewis-Palmer, T., & Merrell, K. (2008). The classroom check-up: A classwide consultation model for increasing praise and decreasing disruptive behavior. School Psychology Review, 37(3), 315–332. https://doi.org/10.1080/02796015.2008.12087879

Robbins, L. B., Pfeiffer, K. A., Maier, K. S., LaDrig, S. M., & Berg-Smith, S. M. (2012). Treatment fidelity of motivational interviewing delivered by a school nurse to increase girls' physical activity. The Journal of School Nursing, 28, 70–78.

Rochat, S. (2019). Effects of motivational interviewing training in career counseling: A Pilot study. Journal of Career Development, 46, 280–294.

Simon, P., & Ward, N. L. (2014). An evaluation of training for lay providers in the use of motivational interviewing to promote academic achievement among urban youth. Advances in School Mental Health Promotion, 7, 255–270.

Small, J. W., Frey, A. J., & Lee, J. (2025). Motivational interviewing training and fidelity monitoring in school-based research: A scoping review. School Mental Health, 17, 387– 399. https://doi.org/10.1007/s12310-025-09754-1

Small, J. W., Frey, A., Lee, J., Seeley, J. R., Scott, T. M., & Sibley, M. H. (2021). Fidelity of motivational interviewing in school-based intervention and research. *Prevention Science*, *22* (6), 712–721. https://doi.org/10.1007/s11121-020-01167-7

Spoth, R., Rohrbach, L. A., Greenberg, M., Leaf, P., Brown, C. H., Fagan, A., Catalano, R. F., Pentz, M. A., Sloboda, Z., Hawkins, J. D., & Society for Prevention Research Type 2 Translational Task Force Members and Contributing Authors. (2013). Addressing core challenges for the next generation of type 2 translation research and systems: The translation science to population impact (TSci impact) framework. *Prevention Science*, *14*(4), 319–351. CINAHL. https://doi.org/10.1007/s11121-012-0362-6

Stormont, M., Reinke, W. M., Newcomer, L., Marchese, D., & Lewis, C. (2015). Coaching teachers' use of social behavior interventions to improve children's outcomes: A review of the literature. *Journal of Positive Behavior Interventions*, *17*(2), 69–82. https://doi.org/10.1177/1098300714550657