

University of Denver

Digital Commons @ DU

---

Electronic Theses and Dissertations

Graduate Studies

---

1-1-2015

## Psychometric Investigation of the Check-In, Check-Out Fidelity of Implementation Measure

Erin Ann Sullivan  
*University of Denver*

Follow this and additional works at: <https://digitalcommons.du.edu/etd>



Part of the [Educational Psychology Commons](#), and the [Quantitative Psychology Commons](#)

---

### Recommended Citation

Sullivan, Erin Ann, "Psychometric Investigation of the Check-In, Check-Out Fidelity of Implementation Measure" (2015). *Electronic Theses and Dissertations*. 1052.  
<https://digitalcommons.du.edu/etd/1052>

This Dissertation is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact [jennifer.cox@du.edu](mailto:jennifer.cox@du.edu), [dig-commons@du.edu](mailto:dig-commons@du.edu).

---

# Psychometric Investigation of the Check-In, Check-Out Fidelity of Implementation Measure

## Abstract

Suspension and expulsion are utilized frequently and disproportionality in schools in the United States. Many schools utilize Positive Behavioral Interventions and Supports (PBIS), a tiered framework to prevent problem behavior and reduce the use of discipline practices (Sugai et al., 2000). Check-In, Check-Out (CICO) is a targeted group behavioral intervention that is utilized within this framework in schools to prevent severe problem behavior in students that are beginning to exhibit externalizing and/or internalizing behavioral needs; thus, preventing the use of exclusionary discipline practices (Crone et al., 2010; Hawken & Horner, 2003). As the use of CICO in schools continues to grow, so too does the need for an instrument measuring its fidelity of implementation. The purpose of this study was to investigate the reliability and validity of the Check-In, Check-Out Fidelity of Implementation Measure (Crone et al., 2010), an instrument created to measure the fidelity of implementation of CICO intervention.

This study assessed the psychometric properties of the instrument utilizing an archival data set collected by the statewide PBIS initiative in a western state in the U.S. The results demonstrated promising content validity, construct validity, internal consistency, and interrater reliability. A unidimensional structure was determined to be the best structure for the instrument based on parsimony and the strong results obtained from the item loadings, internal consistency, and interrater reliability. Implications for use and future research are discussed.

## Document Type

Dissertation

## Degree Name

Ph.D.

## Department

Child, Family and School Psychology

## First Advisor

Gloria Miller, Ph.D.

## Second Advisor

Darrin Hicks

## Third Advisor

Kathy Green

## Keywords

Check-in, Check-out, Fidelity of implementation

## Subject Categories

Education | Educational Psychology | Quantitative Psychology

## Publication Statement

Copyright is held by the author. User is responsible for all copyright compliance.

PSYCHOMETRIC INVESTIGATION OF THE CHECK-IN, CHECK-OUT

FIDELITY OF IMPLEMENTATION MEASURE

---

A Dissertation

Presented to

the Faculty of the Morgridge College of Education

University of Denver

---

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

Erin A. Sullivan

August 2015

Advisor: Gloria Miller, Ph.D.

©Copyright by Erin A. Sullivan 2015

All Rights Reserved

Author: Erin Ann Sullivan  
Title: PSYCHOMETRIC INVESTIGATION OF THE CHECK-IN, CHECK-OUT  
FIDELITY OF IMPLEMENTATION MEASURE  
Advisor: Gloria Miller, Ph.D.  
Degree Date: August 2015

### **ABSTRACT**

Suspension and expulsion are utilized frequently and disproportionality in schools in the United States. Many schools utilize Positive Behavioral Interventions and Supports (PBIS), a tiered framework to prevent problem behavior and reduce the use of discipline practices (Sugai et al., 2000). Check-In, Check-Out (CICO) is a targeted group behavioral intervention that is utilized within this framework in schools to prevent severe problem behavior in students that are beginning to exhibit externalizing and/or internalizing behavioral needs; thus, preventing the use of exclusionary discipline practices (Crone et al., 2010; Hawken & Horner, 2003). As the use of CICO in schools continues to grow, so too does the need for an instrument measuring its fidelity of implementation. The purpose of this study was to investigate the reliability and validity of the Check-In, Check-Out Fidelity of Implementation Measure (Crone et al., 2010), an instrument created to measure the fidelity of implementation of CICO intervention.

This study assessed the psychometric properties of the instrument utilizing an archival data set collected by the statewide PBIS initiative in a western state in the U.S. The results demonstrated promising content validity, construct validity, internal consistency, and interrater reliability. A unidimensional structure was determined to be the best structure for the instrument based on parsimony and the strong results obtained from the item loadings, internal consistency, and interrater reliability. Implications for use and future research are discussed.

## **ACKNOWLEDGEMENTS**

This dissertation is the result of the support, guidance, and encouragement of a great number of people. They will never know the extent of the impact of their expertise, generosity, and kindness on this project and my professional growth.

Thank you to the wisdom, leadership, inspiration, and tenacity of my advisor, Dr. Gloria Miller. I am sure that I would not have completed this project without your positive reinforcement, superb editing, and belief in me. Thanks also to the other members of my committee, Drs Kathy Green, Cynthia Hazel, and Darrin Hicks. I learned from your knowledge and experience. I truly appreciate your dedication to my project.

This project was made possible by the hard work of many school professionals dedicated to evidence-based implementation of the interventions adopted at their school as well as the devoted technical assistance providers that shepherded this implementation. Thank you for putting students first and sharing insight into your work with me.

Finally, thanks to my family and friends. Without their love, patience, confidence, and dinners delivered to the house, none of this would be possible.

## TABLE OF CONTENTS

I. Chapter 1: Statement of the Problem.....	1
Exclusionary Discipline in Schools .....	1
Disproportionate use of Exclusionary Discipline in Schools .....	2
Impact of Exclusionary Discipline in Schools .....	4
Summary .....	7
Alternatives to Exclusionary Discipline .....	7
Comprehensive Systemic Approach to Positive Behavior Support.....	8
Targeted Group Interventions within a System of Positive Behavior Support.....	10
Check-in Check-out (CICO) an Evidence-Based Targeted Group Intervention.....	10
Fidelity Assessment and CICO .....	12
CICO Fidelity of Implementation Measure (CICO-FIM) .....	13
Purpose of the Study .....	14
Research Questions.....	15
Definitions Used in Current Study.....	15
II. Chapter 2: Review of Literature.....	17
Tiered Prevention.....	17
Check-In, Check-Out (CICO).....	22
Theoretical Foundations of CICO.....	23
Description of the CICO Intervention.....	26
Staff agreement .....	26
Personnel.....	26
Student identification .....	28
Daily CICO cycle.....	29
Ongoing Implementation and Follow-up.....	30
Evidence of the Effectiveness of CICO .....	32
Summary of CICO .....	35
Implementation Fidelity.....	35
Fidelity Assessment Defined .....	36
The Importance of Fidelity Assessment .....	37
A Comprehensive Approach to Fidelity Assessment .....	40
Fidelity Assessment and CICO.....	44
Summary .....	46
III. Chapter 3: Method .....	47
Study Design.....	47
Participants.....	48
Population .....	48
Research Question 1 Sampling: Content and Construct Validity.....	52
Research Question 2 and 3 Sampling: Factor Structure and Internal Consistency .....	53

Research Question 4 Sampling: Interrater Reliability .....	54
Consent from Participants.....	56
Instrument .....	56
Development of the CICO-FIM.....	56
Description of the CICO-FIM.....	57
Completion of the CICO-FIM .....	58
Procedures.....	58
Research Question 1 Procedures: Content and Construct Validity .....	59
Research Question 2 and 3 Procedures: Factor Structure and Internal Consistency.....	59
Research Question 4 Procedures: Interrater Reliability.....	61
Data Analysis .....	62
Data Preparation, Screening, and Descriptive Statistics.....	62
Research Question 1 Analysis: Content and Construct Validity .....	62
Research Question 2 Analysis: Factor Structure .....	63
Research Question 3 Analysis: Internal Consistency .....	64
Research Question 4 Analysis: Interrater Reliability .....	64
IV. Chapter 4: Results.....	66
Data Preparation and Screening.....	66
Descriptive Statistics.....	67
Research Question 1 Results: Content and Construct Validity .....	68
Content Validity.....	69
Construct Validity.....	70
Research Question 2 Results: Factor Structure.....	74
Test of Assumptions: Normality.....	74
Test of Assumptions: Linearity.....	75
Exploratory Factor Analysis .....	76
Research Question 3 Results: Internal Consistency.....	80
Research Question 4 Results: Interrater Reliability.....	81
Overall Summary of Results.....	82
V. Chapter 5: Discussion .....	84
Purpose and Design of the Study .....	84
Major Findings.....	85
Research Question 1 Major Findings: Content and Construct Validity ....	86
Research Question 2 Major Findings: Factor Structure.....	87
Research Question 3 Major Findings: Internal Consistency.....	87
Research Question 4 Major Findings: Interrater Reliability.....	88
Implications of Results .....	88
Implications of Results in Research.....	89
Implications of Results in Practice .....	90
Limitations .....	92
Instrument .....	92
Sampling .....	94

Dimensions of Fidelity Assessment.....	95
Future Research .....	96
Conclusion .....	97
References.....	102
Appendices.....	120
Appendix A .....	120
Appendix B .....	122
Appendix C .....	128
Appendix D .....	129

## LIST OF TABLES

Table 1. Race/Ethnicity of the Students in the Sample as Compared to the Students of the State.....	49
Table 2. Number of Schools within the Sample by Grade Level and Region .....	49
Table 3. Research Questions, Sample Size, and Data Analysis Means.....	52
Table 4. Criteria Utilized to Identify Content Experts.....	53
Table 5. Race/Ethnicity of the Students in the 27 School Sample as Compared to the Students of the State.....	55
Table 6. Number of Schools within the 27 School Sample by Grade Level State .....	55
Table 7. Item Means, Medians, Mode, and Standard Deviations of the Largest Sample .68	
Table 8. Item Relevance and Clarity Means and Standard Deviations.....	70
Table 9. Item Means and Standard Deviations for the Predictor and Criterion for the Focused Association .....	71
Table 10. Item Means, Standard Deviations, Skewness and Kurtosis.....	75
Table 11. Principal Component Analysis and Parallel Analysis Results of the CICO-FIM .....	78
Table 12. Principal Component Loadings of the CICO-FIM .....	79

## **CHAPTER 1: STATEMENT OF THE PROBLEM**

*Our shared goal of preparing all students for college, careers, and civic life cannot be met without first creating the safe campuses and positive school climates where teaching and learning can thrive.*

—United States Secretary of Education Arne Duncan on January 8, 2014

### **Exclusionary Discipline in Schools**

The above quote comes from the powerful introduction given by Secretary Arne Duncan to introduce the School Discipline Package Guidance from the United States Department of Education (DOE) and Department of Justice (DOJ). This package of guidance was more than a letter to public schools to remind them of their legal obligation to administer discipline fairly, equitably, and without discrimination. The package of guidance represents a joint Supportive School Discipline Initiative between Secretary Duncan and Attorney General Holder (DOE, 2014), and initiative that focused on creating safe and supportive learning environments and reducing the disproportionate representation of students of color and with disabilities in exclusionary discipline practices (DOE, 2014). This clarion call from Secretary Duncan focused on the integral role of prevention and on using a positive school climate to create safe and supportive learning environments (DOE, 2014). These environments must ensure that schools are safe from violence and disruption while also guaranteeing that exclusionary disciplinary practices are used minimally, only for the worst offenses (DOE, 2014).

The exclusionary disciplinary practices of out-of-school suspension and expulsion are the most severe consequences that a school or district can enact on a student for unacceptable behavior (American Academy of Pediatrics [AAP], 2013). However in practice, data signify that removals or exclusionary discipline are utilized frequently, disproportionately, subjectively, inconsistently, and for low-level behavior violations (AAP, 2013).

### **Disproportionate Use of Exclusionary Discipline in Schools**

The March 2014 *Data Snapshot* provided a first glimpse of data from the results of the 2011-2012 school year of the Civil Rights Data Collection (CRDC) study, which was conducted by the United States DOE Office of Civil Rights (OCR; 2014). The preliminary report included data from every public school in United States, and represented information from 49 million students across the country (OCR, 2014). Unfortunately, the results were bleak. The *Snapshot* reported that during the 2011-2012 school year, over three million public students received out-of-school suspensions, and that over 100,000 students were expelled (OCR, 2014). In this same data set, Black students were suspended three times more often than their White peers, with an average of 16.4% of Black students receiving a suspension, compared to a 4.6% average of White students (OCR, 2014). This disproportionality continues along lines of race and gender, with 20% of Black boys and 12% of Black girls receiving out-of-school suspensions (OCR, 2014). Black girls are suspended more than any other group of girls by race, and more than any other subgroup of boys—with the exception of American Indian/Alaskan Native students (OCR, 2014). American Indian/Alaskan Native students were disproportionately suspended or expelled. In the study, indigenous students represented

only 1% of the total student population, but were 3% of the total students expelled and 2% of students with out-of-school suspensions during the 2011-2012 school year (OCR, 2014). Additionally, students with disabilities were two times more likely to receive one or more out-of-school suspensions, when compared to their non-disabled peers (OCR, 2014).

Exclusionary discipline is shown to be consistently utilized for subjective and low-level behavior infractions (AAP, 2013; American Psychological Association [APA], 2008; Skiba, Michael, Nardo, & Peterson, 2002). *Breaking Schools' Rules*, the groundbreaking study of the discipline of Texas students released in 2011, put an exclamation point on this early research (Fabelo, Thompson, Plotkin, Carmichael, Miner, & Booth, 2011). Fabelo et al. (2011) found that roughly six out of ten Texas public school students were suspended at least once during middle school and high school, and only 3% of the exclusionary discipline actions taken were for infractions that were designated as mandated suspensions and expulsions in the Texas state discipline code. In general, state mandates represent serious offenses, like possession and/or use of weapons and drugs (Fabelo et al., 2011). The remainder of the 97% of suspensions and expulsions were found to be for infractions determined at the discretion of local school officials; violations of local school conduct codes were the primary reason cited for behavior violations (Fabelo et al., 2011). Additionally, multivariate analysis of the data found that school-level administrators within the same district, but at different buildings, interpreted the disciplinary system differently and were thus likely to make different decisions on behavior management and the use of exclusionary discipline. This result gives credence to previous research that exclusionary decisions made at the local school level lack

consistency in implementation, and are often subjective decisions for low-level problem behaviors (Fabelo et al., 2011; APA, 2008; Heaveside, Rowand, Williams, & Farris, 1998). Ultimately, *Breaking Schools' Rules* further brought to light the disproportionate exclusion of African American students, demonstrating that African Americans were more likely to be disciplined during their middle and high school tenure than students of other races, but also that African American students were no more likely to commit offenses considered to be state mandates for suspension and expulsion than students of other races (Fabelo et al., 2011). More questions surface than answers with the data and research, but one thing is clear: the use of exclusionary discipline is frequent, disproportionate, subjective, inconsistently implemented, and enlisted for low-level behavior violations.

### **Impact of Exclusionary Discipline in Schools**

In addition to profoundly impacting students and families, utilizing exclusionary discipline impacts schools, districts, and communities. According to the American Psychological Association (APA; 2008), exclusionary discipline inherently disengages students and families from the learning environment. APA reports that when students are required to leave the school or the learning environment as a result of this punishment, students miss out on important academic instruction and other opportunities. Moreover, using exclusionary discipline negatively impacts the emotional health and wellbeing outcomes for students by increasing student isolation, anxiety, and rejection; it disrupts healthy, adult relationships and interactions. Thus, the time a student spends suspended or expelled results in a loss of engagement in academics and hinders the development of students' connections to the school community. Due to the fact that most parents work,

schools release disciplined students into unsupervised settings, which places students back into the environment that contributed to his or her problematic behavior and provides the student with other opportunities to engage in even more inappropriate behavior. Moreover, when schools release students into a problematic environment, students associate with individuals who also engage in maladaptive behaviors, thus further increasing the risk of problem behaviors.

Recent research has demonstrated that the use of suspension leads to academic failure, students dropping out of high school, and even youth involvement in the legal system (Fabelo et al., 2011; APA, 2008). The “*Breaking Schools’ Rules*” study found that 31% of students with one or more suspensions or expulsions repeated their grade level at least once, compared to 5% of students that were held back and received no disciplinary exclusions (Fabelo et al., 2011). The results of this study confirmed previous research that students with higher drop-out rates had at least one suspension or expulsion compared to their peers, who had no disciplinary involvement (Fabelo et al., 2011). These data, in addition to data from previous studies, submits that students involved with juvenile justice systems are more likely to have been suspended or expelled (Fabelo et al., 2011; APA, 2008).

Exclusionary discipline adversely affects the schools and districts in which they are utilized. In a study by the APA (2008), it was found that schools that utilize high rates of suspension and expulsion have lower rates of academic achievement and fewer indicators of maintaining a positive school climate. As a result, these schools obtained lower ratings on school governance surveys and measures. Further, the study confirmed that suspension and expulsion can also result in a loss of funding for schools, due to the

means with which funding is determined by school attendance. This loss of funding speaks to larger issues that impact school policy on the district level with regard to exclusionary discipline, according to APA (2008). If district policy requires or suggests the use of suspensions or expulsions for disciplinary violations, this places the responsibility of the policy and its outcomes on districts. Thus, the district is responsible for school-level and systems-level issues that result from the use of exclusionary discipline. The unfavorable impacts of exclusionary discipline at the school level have been presented previously (APA, 2008). The systems level disadvantages of utilizing exclusionary discipline are concerned with the nature and implementation of the practice. First, if exclusionary discipline is meant to be a severe response to a high-stakes problem (e.g., weapons or drugs in the school), the theory that a student will be absent from the school for a class period or for a number of days, which will remove the problem, provides a superficial response to a complicated need (AAP, 2013). A severe response does not ensure an effective response. Second, clear data and research has demonstrated that there is variation at the school level with respect to implementing policies of exclusionary discipline that are set at the district and state levels (AAP, 2013). The creation and implementation of exclusionary discipline at the district level mandates that the district is responsible for the consistent, fair, and appropriate implementation at the school level; however, a number of studies have established clear, predictable inconsistencies in implementation at the school level (Fabelo et al., 2011; APA, 2006).

Eventually, our society bears the adverse effects of exclusionary discipline in the form of long-term financial and health consequences (AAP, 2013). A student influenced by exclusionary discipline that does not graduate high school faces missed earnings that

average a \$485,000 loss for men, and \$400,000 for women over their lifetime, compared to high school graduates (Shore & Shore, 2009). This loss of earnings also represents lost taxes collected for federal and state governments, totaling in billions of dollars each year. Ultimately, the average high school dropout is unhealthier than the average high school graduate and has a life expectancy that is six to nine years shorter than the average high school graduate (National Institute of Health, 2003).

### **Summary**

Suspensions and expulsions are used frequently, subjectively, inconsistently, and disproportionately with students of color and with disabilities (Skiba, Chung, Trachok, Baker, Sheya, & Hughs, 2014; Skiba, Horner, Chung, Rausch, May, & Tobin, 2011). The use of exclusionary discipline has profound and adverse effects on students, families, schools, districts, and society (Skiba et al., 2014; Skiba et. al, 2011). Further, not only does the use of suspension and expulsion have these aforementioned problematic effects, it also does not work as a consequence because it does not create the intended behavioral changes (Skiba et. al, 2011). In fact, suspended students are more likely to receive future office discipline referrals and suspensions (Skiba & Sprague, 2008). Suspensions act as a reward rather than a punishment for some students; thus, this punishment is does not extinguish the behavior (Tobin, Sugai, & Colvin, 1996). Moreover, schools with higher rates of exclusionary discipline have been shown to be unsafe for students and staff (Skiba et al., 2014; Skiba et al., 2011; Skiba et al., 2002).

### **Alternatives to Exclusionary Discipline**

Our education system must utilize guidance from the federal government and from extensive research in the field to employ evidence-based prevention and

intervention strategies to create schools that are characterized as safe and supportive learning environments where problem behavior is prevented before exclusionary discipline is needed. The Zero Tolerance Task Force of the APA (2008) called on the utilization of a multi-tiered approach to intervention and support as a means to address the issues of violence and antisocial behavior in schools. Walker et al. (1996) began by calling for a national discussion to bring into focus to the problem of school violence and problem behavior in all its forms. Then Walker et al. (1996) stressed the need for school professionals to lead the efforts in reducing violence and antisocial behaviors. Schools have this unique opportunity due to the nature of their environment, as they hold access to youths as well as professional expertise to address these issues head on. Finally, Walker et al. (1996) provide the foundation for addressing violence and problem behavior through a proposed tiered framework of prevention and early intervention to identify students at risk of violence and problem behaviors.

### **Comprehensive Systemic Approach to Positive Behavior Support (PBS)**

The comprehensive system of behavior support described by Walker et al. (1996) was born out of public health research and serves as a whole-school approach for the effective and efficient delivery of evidence-based instruction, support, and strategy (Sugai & Horner, 2002). The systems and practices at each tier serve as an integrated and comprehensive layering of behavior support that is delivered to the student based on identified risk and/or need (Sugai & Horner, 2002; Walker et al, 1996; Walker & Shinn, 2002). Thus, as the need of the student intensifies, so does the support given to the student (Sugai & Horner, 2002; Walker et al, 1996). Additionally, data is utilized across

the framework to measure student need, fit, response to intervention, and integrity of adult treatment (Hawken, MacLeod, & Rawlings, 2007).

The primary tier, or universal level of prevention, utilizes evidence-based strategies designed to prevent problem behaviors from occurring in the first place. Primary prevention is proactive, delivered to the whole school, and builds on existing protective factors (Walker et al., 1996; Walker & Shinn, 2002). Effective and quality implementation of the universal tier of prevention addresses the needs of an estimated 80% of the school's population (Walker et al., 1996; Sugai, Horner, & Todd, 2000). An example of universal behavioral strategies is the presence of clear, consistent, and school-wide behavioral expectations that are taught, reinforced, and corrected (Sugai & Horner, 2002). The purpose of secondary prevention—or targeted group strategies—is to impede or reverse problem behaviors that have begun to occur (Walker & Shinn, 2002). Such individuals are often considered non-responsive to primary prevention in that they have received the benefit of universal strategies but still demonstrate problem behaviors (Walker et al., 1996). Therefore, targeted strategies are group interventions intended to support approximately 15% of the student population; these supports are delivered to supplement universal strategies (Hawken et al., 2007).

Finally, tertiary prevention—or intensive intervention and supports—are evidence-based, individualized strategies that are delivered to reduce the number of incidents of problem behaviors rather than reverse existence of problem behaviors (Walker & Shinn, 2002). Students requiring tertiary support demonstrate that the most severe behavioral need often occurs across multiple life domains (Eber, Hyde, & Suter, 2011; Walker et al., 1996). As with secondary supports, tertiary supports are provided in

addition to universal and other targeted supports. Examples of tertiary behavioral supports include behavior support plans grounded in functional behavioral assessment, wraparound, and person-centered planning (Hawken et al., 2007).

### **Targeted Group Interventions within a System of Positive Behavior Support**

Targeted group interventions have been found to be effective in addressing the needs of students at-risk for severe problems (Anderson, Christenson, Sinclair, & Lehr, 2004; Colvin, Kame‘enui, & Sugai, 1993; Walker et al., 1996). In fact, a wide range of targeted group interventions has been found to reduce problem behaviors and to increase the use of social skills and student engagement (Golly, Sprague, Walker, Beard, & Gorham, 2000; Hawken & Horner, 2003; Sinclair, Christenson, Evelo, & Hurley, 1998). To be successful at school, students who require this level of intervention have demonstrated that universal practices are not sufficient to support their behavior, but they do not yet require intensive individualized support. Instead, their behaviors demonstrate a need for a group intervention (Anderson et al., 2004; Colvin et al., 1993; Walker et al., 1996). Targeted interventions can be administered utilizing a standard protocol across all students within the group (Walker et al., 1996), and this group approach signifies an efficiency that is not present in more individualized interventions (Colvin et al., 1993; Filter, McKenna, Benedict, Horner, Todd, & Watson, 2007).

#### **Check-In, Check-Out (CICO): An Evidence-Based Targeted Group Intervention**

Check-In, Check-Out (CICO) is one example of a targeted group intervention. The CICO approach to targeted group behavior support is based on the simple strategy of increasing positive adult attention and providing immediate feedback for students (Hawken, 2006; Hawken & Horner, 2003), and the fundamental practices of CICO have

been used in schools for some time (Chafouleas, McDougal, Riley-Tillman, Panahon, & Hilt, 2005; Chafouleas, Riley-Tillman, & Sassu, 2006;). Two foundational components of CICO are daily progress reports (DPR) and the practice of checking in and out daily with an assigned adult mentor. These components have a long history of being effective in supporting student behavior in schools (Anderson et al., 2004; Sinclair et al., 1998). Daily progress reports (DPR) are used to provide structure and ongoing feedback, based on clearly defined goals and expectations, for students participating in the intervention (Anderson et al., 2004; Sinclair et al., 1998). Students also receive instruction on expected behavior and appropriate social skills as well as increased corrective feedback; they also receive increased reinforcement and incentives for following expectations through regular check-in with teachers throughout the day in order to complete their point card (Anderson et al., 2004; Sinclair et al., 1998). This natural data collection system can be used to monitor student progress and trends in behavior (Hawken, 2006; Hawken & Horner, 2003).

CICO is also effective in decreasing problem behavior and the need for use of exclusionary discipline (Campbell & Anderson, 2011; Hawken, 2006; Hawken & Horner, 2003; Filter et al., 2007; March & Horner, 2002; Todd, Campbell, Meyer, & Horner, 2008). However, this intervention lacks a psychometrically sound measure of fidelity to determine the efficacy with which the intervention is delivered. Here fidelity is defined as the degree to which a practice, strategy, or intervention adheres to or is implemented as it was intended (Gresham, 1989). This is of particular importance, because measuring fidelity determines the degree to which the independent variable or treatment is responsible for the change in behavior (Moncher & Prinz, 1991). Consequently, an

accurate measurement of the fidelity of an intervention is critical to ensure that changing behaviors are truly a result of the implementation of the intervention as planned, as opposed to the behavior change being the result of chance or another mitigating factor (Moncher & Prinz, 1991; VanDerHeyden, Witt, & Gilbertson, 2007).

### **Fidelity Assessment and CICO**

Fidelity assessment is an essential practice in the implementation of practices and interventions in education today (Fixsen, Blase, Metz, & VanDyke, 2013; Moncher & Prinz, 1991). Measuring the fidelity of CICO is essential to support further implementation of the intervention, both from practical and scientific perspectives. First, measuring the fidelity of interventions addresses a practical need, and centers on effectiveness and improving implementation, replication, and personnel decisions at the school and district level. According to Fixsen et al. (2013), a tool that measures the fidelity of CICO supports schools in measuring, assessing, improving, and implementing the intervention, as the fidelity of implementation is highly correlated with intended outcomes for students. Thus, schools use a fidelity tool to measure the effectiveness of the intervention to ensure that intended outcomes are attained. Second, Fixsen et al. identified that the tool can be used to identify areas of weakness in implementation and to guide the process to strengthen its implementation. Third, districts can utilize data on fidelity in conjunction with outcome data to determine the need to replicate the intervention in schools across the district. Finally, the data can be used to influence personnel decisions across the district (Fixsen et al., 2013). For example, if a district successfully demonstrates an intervention and decides to replicate the intervention as

described previously, it might be determined that more personnel is needed to institutionalize the intervention across the district.

The scientific or research needs of a fidelity measure for CICO center on the fundamental need to determine if a change in the dependent variable is a direct result of an independent variable (Moncher & Prinz, 1991). A reliable and valid tool that measures CICO effectively will allow researchers to discriminate between implementation with and without fidelity, thus being able to isolate and demonstrate if the results are in fact a product of the intervention (Moncher & Prinz, 1991; VanDerHeyden et al., 2007). A valid and reliable fidelity measure of CICO will also allow researchers and evaluators to measure the extent to which schools utilize CICO by serving as an on/off measure of the evidence-based presence of the intervention in a school (Horner, Todd, Lewis-Palmer, Irvin, Sugai, & Boland, 2004). Furthermore, such an instrument would allow researchers and evaluators to measure the impact of training and/or technical assistance of CICO (Moncher & Prinz, 1991; VanDerHeyden et al., 2007). Thus, the development of a fidelity implementation tool for CICO will allow for expanded, research-based measurements of the relationships between CICO implementation and other valued school outcomes. As stated by DeVallis (2003), it is critically important to first quantify the particular phenomena before tackling other research to determine its impact on a valued outcome.

### **CICO Fidelity of Implementation Measure (CICO-FIM)**

An instrument to measure fidelity of implementation of CICO was constructed and published for the second edition of *The Behavior Education Program: A Check-In, Check-Out Intervention for Students at Risk* (Crone, Hawken, & Horner, 2010) as is

called the Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM). A CICO coordinator and/or a school-based CICO team use the instrument to drive the implementation of CICO at the school level. The instrument provides guidance on areas in need of improvement and on areas that are currently working to determine the fidelity of the model; thus, schools can utilize the data to plan accordingly. In the western state utilized in this study, state, district, and/or school-level coaches are directed to utilize the CICO-FIM to support the improvement of school-level efforts in CICO. State and district personnel also utilize the CICO-FIM to investigate the need for further professional development specific to CICO. The CICO-FIM was designed to provide critical fidelity information with respect to measuring CICO implementation by isolating its independent effects (Moncher & Prinz, 1991; Borelli et al., 2005). Thus, further examination of this tool is needed to ensure that it is in fact a valid and reliable method of measuring the implementation of this important targeted group intervention that has an extensive and wide-reaching impact.

### **Purpose of the Study**

The purpose of this research project was to measure the psychometric properties of the new fidelity measure for CICO. A reliable and valid fidelity instrument was needed before a full examination could be made of changes in dependent variables that may be a result of the systematic implementation of any program employed to reduce the use of exclusionary discipline practices in schools (Gresham, 1989). Since CICO is a widely employed intervention within a system-wide approach to PBIS, more work was needed to establish the fidelity of the implementation. In fact, establishing the fidelity of implementation of an intervention is one of the more important aspects of a scientific

examination of behavior change. Without an understanding of the degree to which an intervention is implemented to fidelity, the effectiveness of the intervention in attaining or failing to attain the desired behavior change cannot be measured (Moncher & Prinz, 1991; VanDerHeyden et al., 2007).

The CICO-FIM is an instrument that was developed to measure the implementation by school personnel of this targeted group intervention (Crone et al., 2010). This study investigates the psychometric properties of this instrument.

### **Research Questions**

This study assesses the validity, factor structure, and reliability of the CICO-FIM by utilizing archival data from schools implementing CICO across a western state. The schools in this study have received at least seven hours of professional development in the systems, data, and practices of CICO. The CICO-FIM was completed by a CICO coordinator or CICO team, typically one to three months after training and then annually to ensure ongoing implementation fidelity. This study investigates the following questions regarding the CICO-FIM:

1. Does the CICO-FIM demonstrate adequate content and construct validity?
2. What is the CICO-FIM factor structure?
3. Does the CICO-FIM demonstrate adequate internal consistency?
4. Does the CICO-FIM demonstrate adequate interrater reliability?

### **Definitions Used in Current Study**

***Check-In, Check-Out (CICO) intervention.*** CICO intervention is a targeted group intervention intended to support students at-risk for severe problem behavior by

increasing positive adult attention and feedback on expected behavior (Hawken & Horner, 2003).

***Fidelity assessment.*** For this study, fidelity assessment is defined as the process of measuring the fidelity of implementation of a practice, program, or intervention (Fixsen et al., 2013).

***Fidelity of implementation.*** Fidelity of implementation is the degree to which an intervention is implemented as it was designed (Carroll, Patterson, Wood, Booth, Rick, & Balain, 2007; Fixsen et al., 2013; Moncher & Prinz, 1991; VanDerHeyden et al., 2007).

## **CHAPTER 2: REVIEW OF THE LITERATURE**

The purpose of this study was to investigate the psychometric properties of the Check-In, Check-Out Fidelity of Implementation (CICO-FIM) tool. Check-in, Check-out (CICO) is an evidence-based, targeted group intervention utilized in schools across the United States as part of a multi-tiered approach to prevent problem behavior and increase the use of prosocial behaviors in students. CICO intervention plays an integral role in a whole-school approach to increase the likelihood that behavioral needs are addressed and supported in a manner that decreases the need of utilization of suspension and expulsion. Thus, an instrument that measures the fidelity of CICO implementation is essential in supporting student behaviors and preventing the use of more intensive disciplinary action.

Chapter 2 presents a review of the literature to provide a context for the study and to demonstrate why this study is important and timely in education today. First, foundational, evidence-based literature in support of the use of a multi-tiered approach to prevention and intervention for social, emotional, and behavioral success is reviewed. CICO, as well as research supporting its efficacy, is described within this framework. Finally, the concept and use of fidelity assessments as a critical component of evidence-based practices is covered and tied specifically to the CICO process.

### **Tiered Prevention**

A tiered prevention framework is the necessary foundation of recent school reform efforts to make schools a safer, more predictable environment, one where every

student can be successful. One measure of the success of this prevention framework is an expected decrease in suspension and expulsion as a standard disciplinary practice. The comprehensive implementation of a CICO intervention is one cornerstone of such efforts.

The conceptual foundation of a tiered prevention framework for academics and behavior in schools evolved from the public health model of prevention (Walker et al., 1996; Walker & Shinn, 2002). This study focused on the behavioral aspect of this framework. Under the public health model, evidence-based interventions intend to prevent problem behaviors, and each tier has a unique goal and desired outcome (Walker & Shinn, 2002). This demonstrates an important change in the educational approach towards prevention and its role as part of early intervention for students (Walker & Shinn, 2002). Since the initial inception of this framework, a number of studies have been conducted that demonstrate the value of prevention and early intervention in increasing outcomes for every student, regardless of their current status of achievement. It is noteworthy that interventions employed within this framework, when implemented as designed, must be evidence-based to demonstrate their reliability in achieving intended outcomes for students (Cohen, Kincaid, & Childs, 2007; Horner, Todd, Lewis-Palmer, Irvin, Sugai, & Boland, 2004; Vincent, Spaulding, & Tobin, 2010). Furthermore, in order to ensure comprehensive application of the multi-tiered framework, the practices, interventions, and programs across all three tiers must be coordinated across the school, in homes, and in the community (Walker & Shinn, 2002). The tiered framework is often illustrated as a triangle, with the primary prevention serving as the base, the secondary prevention appears in the middle and towards the top, and tertiary prevention is the very tip of the triangle.

The goal of primary prevention, or the universal tier of support, is to keep problem behaviors from occurring (Walker & Shinn, 2002). Carr et al. (2002), Walker et al. (1996), and Walker and Shinn (2002) agree that primary prevention requires interventions that are implemented across the school in order to reach each student and prevent problem behaviors. This school-wide approach also requires the buy-in and engagement of all staff members to actively participate in the instruction and support of the model. Furthermore, this school-wide approach is characterized by each student receiving the same primary prevention intervention in similar settings by utilizing the same instructional methods for the same amount of time. Research from Carr et al. (2002), Walker et al. (1996), and Walker and Shinn (2002) all put forth that implementation of evidence-based interventions with fidelity at the primary prevention level effectively supports 75% to 85% of students in a school.

The goal of secondary prevention, or the targeted tier of support, is to reverse or obviate further harm (Walker & Shinn, 2002). Students receiving support at this level are typically considered to be at-risk for serious problem behavior (Walker & Shinn, 2002). Targeted supports are often characterized as small group or individualized interventions in order to effectively and efficiently support students that have proven to be non-responsive to school-wide universal supports, but not yet exhibited severe behaviors (Walker et al., 1996; Walker & Shinn, 2002). As the focus of this study, CICO is one example of a targeted group intervention (Crone, Hawken, & Horner, 2010; Hawken, 2002). This intervention was designed to support the behavior of at-risk students within this framework (Crone et al., 2010). Typically, 10% to 15% of students should respond to

interventions at the secondary prevention level when implemented with fidelity (Walker et al., 1996; Walker & Shinn, 2002).

Finally, the goal of tertiary prevention, or the intensive tier of support, is to reduce harm to the student (Walker & Shinn, 2002). Both Walker et al. (1996) and Walker and Shinn (2002) stated that at this point, the student demonstrating this level of need requires more individualized and intensive interventions that occur across the domains of home, school, and community to decrease the severity of the problem behavior. Furthermore, students demonstrating a need for tertiary prevention have not responded to primary or secondary prevention interventions. Finally, when evidence-based interventions at the primary and secondary levels of prevention are implemented, only 3% to 5% of students require tertiary prevention (Walker et al., 1996; Walker & Shinn, 2002).

The ongoing utilization of data across these three tiers of prevention is integral in the application of the framework, and is often referred to as the problem-solving model (Newton, Horner, Algozine, Todd, & Algozine, 2012; Tilly, 2008). The problem-solving model is the practice of utilizing data across the tiers of prevention to measure the effectiveness of the interventions at each level of implementation and ultimately ensure increased outcomes for students (Newton et al., 2012). This model is characterized by four general themes to guide decision making (Tilly, 2008). It is important to note that a number of problem-solving models demonstrate four or fewer questions (Newton et al., 2012). However, these four themes are considered essential, and are found throughout the research in this area, as well as in additional models. These themes require the utilization of data to do the following: define the problem, determine why the problem is happening, determine what can be done to mitigate the problem, and ensure that the selected

intervention achieved the intended goal of mitigating the problem (Newton et al., 2012; Tilly, 2008). Outcome and fidelity assessments are the two sources of data used throughout the problem-solving process.

The process of problem-solving is essential to the prevention framework (Newton et al., 2012). It is the ongoing use of data that supports decision making in a school-wide approach regarding the efficacy of practices, strategies, and programs used, as well as the successes or needs of students across the school (Anderson & Borgmeier, 2010; Sugai & Horner, 2002). This approach is referred to as Response to Intervention (RTI) logic (Batsche et al., 2005; Fairbanks, Sugai, Guardino, & Lathrop, 2007; Cheney, Flower, & Templeton, 2008). Exercising this logic, the data is used to assess how well students are achieving at each tier of support and to influence decisions regarding current and future needs and supports. It is essential that these decisions use data regarding student outcomes and adult implementation, as student progress must be weighed against adult implementation and a school cannot expect a student to be successful if the adults of the school do not implement an intervention as it was designed (Anderson & Borgmeier, 2010; Fairbanks et al., 2007).

Positive Behavior Support (PBS) is the widely used application of a multi-tiered prevention framework (Horner, Sugai, & Anderson, 2010). According to Horner et al. (2010), PBS is a set of systems and practices across the universal, targeted, and intensive tiers that are characterized by a focus on prevention, fidelity of implementation, ongoing data-based decision making, and job-embedded professional development and coaching. The conceptual model of PBS is built upon the use of applied behavior analysis, the prevention framework, universal screening, accurate and efficient data-based decision

making, the integration of academics and behavior, and replication through the use of implementation science. The systems and practices of PBS build local leadership capacity to adopt and implement effective, contextually relevant behavioral practices and interventions at each tier (Horner et al., 2010). Furthermore, the practices of PBS are unique to other programs in that they actively entail the use of psychometrically sound assessments of fidelity, thus ensuring that the practices are being implemented as directed in their evidence base (Cohen, Kincaid, & Childs, 2007; Horner et al., 2004; Vincent, et al., 2010). This provides a safeguard that the behavior of students is supported and measured against the reliable implementation of these practices (VanDerHeyden, et al, 2007). CICO is one practice utilized at the targeted tier of intervention and support (Hawken & Horner, 2003).

### **Check-In, Check-Out (CICO)**

As previously stated, CICO is a targeted group intervention that is evidence-based and therefore can be measured by the CICO-FIM. Due to their efficiency and effectiveness, targeted group interventions are an important component of a comprehensive prevention framework (Hawken, 2002; Walker et al., 1996). Efficiency is denoted both by the speed with which students identified as needing targeted support receives said support as well as the ease of use by school professionals (Hawken, 2002), while effectiveness is considered to be the efficacy in reducing future problem behaviors as well as the cost benefit of the intervention (Hawken, 2002). The most effective targeted intervention is cost effective to implement in that students can be supported quickly without large amounts of time and effort from staff (Hawken, 2002; Fairbanks et al., 2007). The intervention can be implemented with few resources outside of staff

training, time, and effort. Lane, Capizzi, Fisher, and Ennis (2012) measured the social validity of CICO and found that the time, effort, and other costs of implementation were valued and useful for staff and students. The defining features of CICO are built upon years of research and theory in education, applied behavior analysis, psychology, and mental health (Crone et al., 2010; Horner et al, 2010).

### **Theoretical Foundations of CICO**

Check-In, Check-Out uses behavioral principles, social validity theory, and ecological systems as its theoretical foundations to reduce risk and increase behavioral outcomes for students and families. First and foremost, CICO is built on the theoretical foundation of behavior theory (Hawken & Horner, 2003), which centers on understanding the antecedents and consequences that predict and reinforce problem behaviors, and utilizes these foundations to support and prevent problem behavior (March & Horner, 2002). CICO is designed for students whose behavior is reinforced by adult attention (Lane, et al., 2012; Hawken et al., 2007). Embedded, positive adult attention is utilized within this function based support model to prevent problem behavior in students whose behavior is maintained by receiving adult attention (March & Horner, 2002). Early investigation into positive adult attention as a component of a targeted group intervention was initiated by Sinclair et al. (1998), in an examination of the effectiveness of the Check and Connect program. Sinclair et al. (1998) found that utilizing positive adult feedback was effective in preventing problem behavior, and also demonstrated that positive adult feedback is most effective when it comes from a member of the school staff, rather than from a volunteer.

CICO also works to eliminate antecedents for students participating in the intervention, which is a validated behavioral practice present in CICO (Crone et al., 2010). Antecedents, or triggers for problem behavior, are eliminated through structures and practices of CICO. According to Crone et al. (2010), the first elimination of antecedents in CICO is the structures present at check-in, those specifically, related to school supplies. Making school supplies available to students prevents them from getting into trouble for coming to class without materials. The second set of elimination of antecedents is the use of positive prompts at check-in and throughout the day, as prompts delivered by adults throughout the day focus on preventing problem behavior. Prompts pre-correct students to have a good day and to focus on achieving expectations; thus, prompts eliminate antecedents and prevent problem behavior (Crone et al., 2010).

The third validated behavioral practice of CICO is the use of daily progress reports (DPR) to increase desired behavior through structured, consistent, and efficient feedback (Long & Edwards, 1994; Riley-Tillman, Chafouleas, Briesch, & Eckert, 2008). A standard DPR form is used to provide immediate feedback for each student participating in CICO in order to increase the efficiency of the intervention (Crone et al., 2010; Hawken, 2002; Hawken & Horner, 2003). Within behavioral interventions, the use of DPRs has an extensive evidence base as an effective means for providing accurate and timely feedback to students (Chafouleas, Riley-Tillman, Sassu, LaFrance, & Patwa, 2007; Long & Edwards, 1994).

While the foundations of CICO are firmly ensconced within behavioral practices, the intervention also includes elements found in the work of Vygotsky (1978) and Bronfenbrenner (1979). Vygotsky's (1978) theory of scaffolding is the process of school

professionals supporting the learning of students based on where they are in the development process. Students that need more support to be successful in developing receive more direct support from school professionals. CICO, as well as other targeted group interventions, are forms of developmental scaffolding, where students demonstrate a need for intervention with problem behaviors, and thus cannot be left to continue independently. However, such students do not demonstrate severe behavior that requires intensive adult attention. Instead, CICO scaffolds the support students receive between independence and intensive adult support. The practices of CICO serve as prompts to support further learning and development by students who need more to be successful.

The work of Bronfenbrenner (1979) and the ecological systems theory has influenced CICO interventions. Bronfenbrenner's ecological theory focuses on the role of various settings and organizations that influence thinking and growth in students, which is emblematic in the systems and the practices of CICO implementation. The systems approach to CICO is inclusive of an entire school, and it is clear that all school personnel influence students—within their microsystem, and not just their immediate teacher. The practices of using an adult mentor and increasing adult attention and feedback demonstrate the interaction across the microsystem between student and school personnel. The practice of the daily, two-way communication between the school and the participating families in the intervention is another example of institutions within the microsystem that influence each other to increase the development of the student in the intervention. The theoretical foundations of CICO serve as the backbone of the role the intervention plays in increasing support and decreasing the risk of students receiving suspension or expulsion.

## **Description of the CICO Intervention**

The key elements of implementing CICO intervention include staff agreement, personnel, student identification, the daily CICO cycle, and ongoing implementation and follow-up (Crone et al., 2010; Hawken & Horner, 2003; Hawken & Johnston, 2008). The daily cycle describes the use of the foundational components of CICO, including the use and practices of an adult mentor, the means toward increased positive feedback, and the use of a DPR.

### **Staff Agreement**

The adoption of CICO requires school staff to agree to implement CICO and other behavioral supports to meet the needs of students (Crone et al., 2010; Hawken, 2002, Hawken & Horner, 2003). Even though only 10% to 15% of the student population might be involved in the intervention at any one time, the practices of CICO are embedded throughout the school, and thus require all staff to buy-in and participate, as schools that have implemented CICO adopted the intervention based on whole-staff agreement (Crone et al., 2010; Hawken, 2002, Hawken & Horner, 2003).

### **Personnel**

After staff agreements have been acquired, issues related to personnel must be addressed, including include identifying a team to lead the implementation, identifying a CICO coordinator, and conducting training for the whole staff (Crone et al., 2010). CICO utilizes the PBS model of shared leadership in the form of team-based implementation (Sugai & Horner, 2002). Thus, a new or existing team is identified to lead the ongoing operation of the intervention. This team receives initial and follow-up training in the daily and ongoing implementation of CICO. During this training, the team learns how to

implement CICO, and also uses the time to create contextually relevant practices of intervention for their school. In advance of training, one member of the team is identified as the CICO coordinator. The CICO coordinator is responsible for day-to-day implementation, data entry, and leading ongoing team meetings for CICO. The CICO coordinator typically requires 10-15 hours of time each week dedicated to these tasks (Crone et al., 2010; Hawken & Johnston, 2008).

After completion of the CICO team training, team members return to the building and provide training in the basics of how to implement CICO to all school staff. This training includes both the systems required to execute the intervention school-wide as well as the skills required to implement the specific practices of the intervention. This training typically focuses on utilizing positive adult attention, instructional feedback, and accurate completion of the DPR, means of referring and exiting students, and the day-to-day logistics of implementing the intervention. Training is also done with participating students and families. Students participating in the intervention receive clear, developmentally appropriate instruction and training in the daily CICO cycle prior to implementation in order to support the successful application of the intervention. Families are trained in the implementation of CICO as well (Hawken, 2002). They are provided information on the presence of the intervention at the school, how to make a referral, how the intervention works, and the means toward achieving two-way communication that is specific to their student (Crone et al., 2010; Hawken, 2002; Hawken & Horner 2003).

## **Student Identification**

Student identification is an integral element of CICO, as well as any targeted group intervention (Anderson & Borgmeier, 2010; Crone et al., 2010). There are three possible ways in which a student is identified for entry into CICO. First, school behavior support teams utilize existing data to determine need. The sources of this data include attendance and office discipline referrals (ODRs). Attendance data that demonstrates a recent spike in absences or tardiness can be cause for further investigation, which could result in identifying the student for CICO intervention. Additionally, schools that utilize universal behavior supports with fidelity employ the use of ODRs for ongoing, data-based decision making. The rate of ODRs is unique to each school, and should be taken into account when devising a cut score for entry into CICO; however, this range is often between two to five ODRs in one school year (Crone et al., 2010). A sudden spike in ODRs can also be investigated to determine a student's need for CICO.

The second means for identifying students at-risk for problem behavior, and thus in need of CICO, is the use of universal behavior screening (Crone et al., 2010). Academic, behavioral, and emotional screening at the universal level is a critical prerequisite to provide appropriate instruction, support, and prevention within RTI logic (Glover & Albers, 2007). Universal behavior screening tools are utilized as a means to create a baseline of performance and early identification of at-risk students (Severson, Walker, Hope-Doolittle, Kratochwill, & Gresham, 2007). This screening is a systematic tool used to detect a subset of students from the whole population who are struggling behaviorally and/or are at-risk of short-term or severe problems, including students who would benefit from CICO (Sprague, Cook, Browning-Wright, & Sadler, 2008). Finally,

students can be identified to be involved in CICO by teacher or family referral (Crone et al., 2010). Once the family or teacher referral is initiated, the CICO team must use data to confirm that the student is an appropriate candidate for CICO (Crone et al., 2010). Thus, working with the family and the teacher to ensure student, teacher, and family needs are met using the appropriate mechanisms is essential.

### **Daily CICO Cycle**

Once staff agreements are in place, personnel are trained. Once clear mechanisms for student identification are in place, the daily implementation of the CICO cycle begins, beginning each day at check-in (Crone et al., 2010; Hawken & Johnston, 2008). Students participating in CICO arrive each day before school starts at a pre-determined spot for check-in with a school professional, often the CICO coordinator (Crone et al., 2010; Hawken & Johnston, 2008). During check-in, the CICO coordinator greets each student, thanks them for checking-in, does a quick scan of each student to ensure that they are ready for the day with appropriate materials, and provides them with the school's DPR. The DPR lists school-wide expectations and also provides places for teachers to rate student performance on each expectation for a specified time period. The time periods utilized in the DPR can represent very specific class periods or natural transition periods in the day. The DPR is made contextually relevant and developmentally appropriate by the CICO team, but it must be standardized in its expectations and the number of periods throughout the day in order to be a CICO DPR.

The student carries the DPR throughout the day and, after check-in, the student gives the DPR to his or her teachers. Teachers are trained to give a positive greeting along with the prompt to have a good day (Crone et al., 2010; Hawken & Johnston,

2008). Teachers provide students with feedback regarding their social behavior at the end of each period on the DPR. The teacher also rates the student on the DPR across each expectation for the associated period and briefly and positively instructs the student on their ratings.

At the end of the day, the student takes the DPR and checks-out with the CICO coordinator at a predetermined location. At check-out, points for the day are totaled. Point goals are typically 80%-85% of the total points available (Crone et al., 2010; Hawken & Johnston, 2008). However, this goal is adjusted based on need that is determined by the data. Students that meet their daily point goal receive a verbal and a small but tangible reinforcement. The DPR is signed by the CICO coordinator and sent home to initiate a two-way communication, allowing families to have ongoing information regarding their child's participation in CICO and to provide a platform for two-way communication with school personnel as students return the DPR during check-in on the next day of school (Crone et al., 2010; Hawken & Johnston, 2008).

### **Ongoing Implementation and Follow-up**

The procedures and practices of ongoing implementation and follow-up are essential to the viable and sustainable implementation of the CICO intervention. These procedures and practices include data-based decision making and ongoing communication to staff and families.

One aspect of the ongoing implementation of CICO focuses on data-based decision making, where the team responsible for the ongoing implementation of CICO meets bi-weekly (Crone et al., 2010; Hawken & Johnston, 2008). First, DPR data is examined for individual students to assess their performance in the intervention and to

determine if less support, more support, or no change in support is necessary (Crone et al., 2010; Hawken & Johnston, 2008). Specifically, the team uses DPR data in conjunction with attendance and ODR data to determine if students are making progress towards achieving their goals (Crone et al., 2010; Hawken & Johnston, 2008). Additionally, the team must assess if the intervention is being implemented as it was designed to ensure that the student has a fair shot at success. DPR data is also summarized across all students to investigate the efficacy of the intervention as a whole, which is determined based on the percent of students that check in and check out each day, the percent of students making progress or meeting goals, and the reduction of ODRs received by students in the intervention (Crone et al., 2010; Hawken & Johnston, 2008).

The second aspect of ongoing implementation and follow-up of CICO is ongoing communication with staff and families (Crone et al., 2010; Hawken, 2002; Hawken & Horner 2003). The team reviews the data of referred students to determine their appropriateness for the intervention (Crone et al., 2010; Hawken & Johnston, 2008). The status of the CICO intervention is shared with staff frequently. This communication includes updating personnel on the data reviewed by the CICO implementation team. Individual student performance is shared with the school personnel working directly with each student (Crone et al., 2010; Hawken, 2002; Hawken & Horner 2003). Communication with families regarding CICO intervention is provided on a daily, weekly, and monthly basis (Crone et al., 2010; Hawken, 2002). Each day, the school sends the DPR home to the students' families to inform them of daily progress and to initiate two-way communication (Crone et al., 2010; Hawken, 2002). At the end of each

week, families receive a summary of their students' performance in meeting daily and weekly CICO goals (Crone et al., 2010). Finally, the CICO implementation team shares a monthly assessment of student performance and plans for future implementation.

### **Evidence of the Effectiveness of CICO**

The CICO intervention has a clear theoretical foundation, and uses practices that have been validated from these theories. Initial validation and adoption of CICO is focused on the theoretical practices and elements of the intervention, since this is the time the intervention has been evaluated in its entirety.

In an early investigation of CICO, March and Horner (2003) found that students whose problem behavior was maintained by adult attention demonstrated decreased problem behavior and variability in behavior after implementation of the intervention. Additionally, the students were observed to have increased rates of academic engagement during their participation in CICO. Hawken and Horner (2003) found similar results with respect to decreased problem behavior and increased academic engagement in students whose problem behaviors were maintained by adult and peer attention. Hawken (2006) found that CICO was successful in reducing problem behavior for 70% of students participating in the intervention.

Moreover, further investigation into the data for the 30% of non-responders found that, in at least 10% of the students, their behaviors demonstrated that more intensive support was needed, thus reinforcing the premise that targeted group interventions are not intended to support severe problem behavior. These early investigations into CICO occurred at the middle-school level (Hawken, 2002; Hawken & Horner, 2003; March & Horner, 2003). More recently, Lane et al. (2012) and Simonsen, Myers, and Briere,

(2010) studied CICO in middle schools. Lane et al. (2012) utilized a single case design in four middle school students, and found that three of the students responded to CICO intervention by decreasing problem behaviors and attaining their DPR goals. Simonsen et al. (2010) used experimental group design in an urban middle school setting. The results of this study were mixed; it was found that students participating in CICO demonstrated an observed decrease in off-task behaviors. Also, the decrease in off-task behaviors actualized by students participating in CICO was statistically significant compared to the gains evidenced by the control group. Teacher ratings on student behaviors did not reflect this increase in on-task behaviors, however. The students participating in CICO also received fewer ODRs compared to peers in baseline and control groups. Simonsen et al. (2010) put forth that this discrepancy between observed change and social validity could be a result of the inconveniences experienced by the participating teachers, or because of the possibility that took longer for the perceptions of teachers to change. To date, there are no studies that have investigated CICO at a traditional high-school that have been published in a peer-reviewed journals (Maggin, Zurheide, Pickett, & Baillie, 2015).

An extensive number of research studies have been executed on the application of CICO in elementary schools. For example, Fairbanks et al. (2007) saw a decrease in ODRs; additionally, teachers involved in the intervention reported in a social validity questionnaire that the intervention was easy to use and that the intervention increased the general climate in the classroom. Filter et al. (2007) examined CICO at three elementary schools within one district. The results indicated a significant reduction in ODRs for the 12 students that participated the intervention. Teachers within this study rated positively all questions of perceived effectiveness and efficiency of the intervention. Todd et al.

(2008) also investigated the efficacy of CICO using single case design. Students that participated in the intervention had a reduction in observed problem behaviors and ODRs. Additionally, teachers and staff involved in the intervention reported satisfaction. In fact, during social validity testing at the end of the intervention, all staff reported that they would recommend use of CICO to other schools. Most recently, Miller, Dufrene, Sterling, and Olmi (2015) utilized an A-B-A-B withdrawal design to determine if the CICO intervention was successful in decreasing problem behaviors and increasing academic engagement for the three students that participated in the intervention.

Finally, CICO has also been investigated in alternative educational settings. Swoszowski, Jolivette, Fredrick, and Heflin (2012) implemented CICO with six students identified as having emotional and behavioral disorders in a residential facility that implemented universal PBS to criterion according to the evidence-based fidelity assessment, the School-Wide Evaluation Tool (SET; Sugai, Lewis Palmer, Todd, & Horner, 2001). Students were identified utilizing the criterion of receiving two to five ODRs. Swoszowski et al. (2012) found that four of the six students decreased problem behaviors, with the data from two students showing too much variability to draw any conclusions. This work clearly demonstrates that CICO is applicable and efficacious in alternative educational settings for students with disabilities. This work also adds to the growing evidence behind the use of CICO as a support for attention and escape-maintained behaviors. Swoszowski et al. (2012) put forth the theory that CICO supports escape-maintained behaviors due to the timing and distribution of adult attention, since required adult interactions can be very brief and are predominantly positive.

## **Summary of CICO**

Utilizing a multi-tiered framework that includes evidence-based practices, strategies, and programs is an essential step in increasing prosocial student success while decreasing the use of suspension and expulsion (Carr et al., 2002; Walker et al., 1996; Walker & Shinn, 2002). The CICO intervention is one such approach, and is the focus of this study. However, targeted practices and programs such as CICO require instruments to measure their implementation fidelity. Crone et al. (2010) published the CICO-FIM to support the implementation of CICO with integrity. This study is the first investigation into the psychometric properties of this instrument; thus, the following comprehensive review of the critical role of implementation fidelity is essential to provide context and relevance for the study.

### **Implementation of Fidelity**

Simply put, fidelity is defined as “the delivery of instruction in the way in which it was designed to be delivered” (Office of Special Education Programs, 2014). The investigation of measuring the fidelity of implementation in education has continued to grow during the past twenty years (Hagermoser Sanetti & Kratochwill, 2009). This growth is a product of the confluence of changes in federal legislation, guidance from professional organizations, proposal priorities from education granting agencies, and the increased adoption and utilization of RTI and PBS (Horner et al., 2004; *Individuals with Disabilities Education Improvement Act*, 2004; Institute of Education Sciences, 2011; National Association of School Psychologists [NASP], 2009; *No Child Left Behind*, 2002; Kratochwill, 2007; Office of Special Education Programs, 2014; VanDerHeyden et al., 2007). The following review presents the evolution of defining fidelity in the field of

education and the critical importance of fidelity assessment in research and practice. It concludes by outlining a comprehensive approach to fidelity assessment in education and a rationale for a CICO fidelity assessment.

### **Fidelity Assessment Defined**

In education, the modern-day definition of fidelity assessment is predominantly derived from the fields of medicine and psychology, as well as research in the social sciences. Early fidelity assessment was focused within the fields of medicine and clinical settings, and focused on the compliance to and dosage of a particular treatment (Harn et al., 2013). Fidelity assessment was holistically defined in education by Moncher and Prinz (1991), providing an early investigation and framework of fidelity assessment in schools through the practices of school psychologists. In this work, Moncher and Prinz (1991) utilized the phrase “treatment fidelity” to define and provide a historical perspective of fidelity assessment in the field of psychology. The collection and use of treatment integrity data came in response to a demand for greater accountability with respect to the effectiveness of psychotherapy in the 1950s, and also stemmed from the growth of community mental health centers in the 1960s (Moncher & Prinz, 1991). As the field of psychology grew and evolved, so did the measurement of treatment fidelity.

Today, the publication, dissemination, and use of treatment protocols and therapeutic manuals often include measures of treatment fidelity (Moncher & Prinz, 1991; Gresham, Gasnle, & Noelle, 1993). Furthermore, researchers across the social sciences use the concept of fidelity assessment. Specifically, researchers measure fidelity to ensure that the treatments or interventions being tested are implemented as defined in their study (Harn et al., 2013; Gresham, MacMillan, Beebe-Frankenberger, & Bocian,

2000). Ultimately, the viewpoints of researchers and practitioners merged to necessitate the measurement of fidelity and define it as the degree to which the change in the dependent variable is due to an independent variable (Moncher & Prinz, 1991; VanDerHeyden et al., 2007). Over the last fifteen years, the term “fidelity assessment” grew in use in education to describe the measurement of implementation of practices, curriculum, programs, and interventions (Durlack & DuPre, 2008; Harn et al., 2013; O’Donnell, 2008). The measurement of fidelity assures that outcomes obtained in the dependent variable are in fact a result of an independent variable and not the result of a variable not measured or controlled (Gresham et al., 2000).

### **The Importance of Fidelity Assessment**

The purpose of measuring fidelity in schools is to determine if the practitioners are implementing a strategy, program, or practice as described by the authors and the evidence base of the strategy, program, or practice. The importance of fidelity has implications for research and practice in the field of education.

The assessment of fidelity is important to research because it impacts internal validity, construct validity, external validity, and generalizability. Simply stated, internal validity describes changes in the dependent variable that are functionally related to controlled changes in the independent variable (Furr & Bacharach, 2014). This is easily illustrated in school-based interventions that focus on behavior, as school personnel must ensure that changes in the demonstrated behavior are a result of changes in the environment (Baer, Wolf, & Risley, 1987; Gresham, 1989). Unless the fidelity of implementation is measured, it cannot be determined whether the changes in the dependent variable were a result of the independent variable or were because of

extraneous circumstances (Baer et al., 1987; Gresham, 1989). Second, the measurement of fidelity also impacts external validity. Scientifically, it is important to measure and report fidelity of implementation in order to support further research, specifically as it relates to meta-analysis or systematic reviews that utilize published research that relies on homogeneity between studies to appropriately aggregate and analyze data (Flay et al., 2005; Carroll et al., 2007). Without these measures, there is no clear measure for determining homogeneity within the research, which presents a threat to the internal validity of the meta-analysis (Flay et al., 2005; Carroll et al., 2007). Finally, the measurement of fidelity is integral in generalizing results. As previously stated, fidelity is defined as the degree to which the change in the dependent variable is due to an independent variable. Thus, by its very definition, the presence of fidelity is required to generalize results purported by a study (Moncher & Prinz, 1991).

The assessment of fidelity is also important with respect to the practical application of interventions in schools. Fidelity assessment is important in the implementation of strategies, practices, programs, and interventions because it ensures that school professionals are implementing as they are designed in order to be effective (Gresham et al., 1993; Noell et al., 2005). Data from the assessment measures change in adult practice and supports the implementation of RTI and PBS in schools (Sugai & Horner, 2002; VanDerHeyden et al., 2007). Fidelity assessment also serves as a feedback loop to adults implementing intervention on how to better implement as designed and to support the evidence base, thus allowing for early detection and correction of errors in practice through ongoing action planning (Noell et al., 2005; Moncher & Prinz, 1991). Moreover, fidelity assessment is important in practice because it reduces the cost of

implementation (Moncher & Prinz, 1991; Noell et al., 2005). Utilizing fidelity promotes efficiency, prevents program failure, and corrects errors, thus reducing the costs associated with failed implementation (Moncher & Prinz, 1991; Noell et al., 2005). These are costs to the implementers as well as those receiving treatment.

In the case of this study, failed implementation without ongoing measurement could be a high cost for students in that it could be the difference between vital time in classroom receiving instruction or receiving a suspension or expulsion. Another important implication for the practical use of fidelity assessment is that a higher level of implementation fidelity results in increased outcomes for students (Durlack & DuPre, 2008; O'Donnell, 2008). In fact, studies that utilize fidelity assessments with acceptable psychometric properties demonstrate a consistent relationship between fidelity and outcomes (Biggs, Vernberg, Twemlow, Fonagy, & Dill, 2008; Forgatch, Patterson, & DeGarmo, 2005). Studies that utilized fidelity assessments without psychometric investigation (such as observation protocols or a checklist) lacked consistent evidence that increased fidelity, leading to better outcomes for students (Gansle & McMahon, 1997; Rhymer, Evans-Hampton, McCurdy, & Watson, 2002).

Finally, important practical applications of fidelity assessment also have ethical implications. The evaluation and documentation of fidelity is essential to the successful execution of a multi-tiered model of prevention (Noell & Gansle, 2009). A student receiving immediate and effective intervention and support at the time of need is a primary principle of RTI logic, and the assessment of fidelity of these interventions and support is central in understanding their efficacy in addressing such needs (Fairbanks et al., 2007; Cheney et al., 2008). VanDerHeyden et al. (2007) indicated that intervention

delivery is essential to RTI, but asks, without ongoing measurement of fidelity, how can one demonstrate if the failure was system or child-specific? A child-specific failure is when the intervention does not work for that particular student, and another option needs to be investigated (VanDerHeyden et al., 2007). VanDerHeyden et al. (2007) continue by indicating that within this model, a student's response to an evidence-based intervention implemented with fidelity is the primary factor in determining the level of intervention required. A systems failure is when the system has failed at implementing the intervention; thus, the system has failed the child. Ultimately, the delineation between system or child failure is the greatest primary factor in determining the level of intervention received and potentially a student's eligibility for special education services. Therefore, a system failure must be measured in the form of fidelity assessment in order to ensure that a student's protections to due process are being met. Moreover, a lack of assessment of fidelity leads to higher rates of defection from treatment and thus higher rates of student failure and implementation fatigue and/or school professionals opting out of implementation (Noell et al., 2005).

### **A Comprehensive Approach to Fidelity Assessment**

The assessment of fidelity is not a singular concept; rather, it is a construct made up of several components that are categorized into the following major areas: The first major area towards a comprehensive approach to fidelity assessment is termed "prerequisite considerations" (Moncher & Prinz, 1991). Prerequisite considerations are the tasks that are necessary to define the independent variable and to effectively train implementers in implementing the independent variable. In this study, the independent variable is the CICO intervention, as defined by Hawken and Horner (2003). In order to

train implementers effectively in the intervention, the Statewide Positive Behavioral Interventions and Supports (PBIS) initiative utilized high quality materials in conjunction with adult learning principles (Dunst & Trivette, 2009; Moncher & Prinz, 1991). The materials were created utilizing sources put forth in original research by Hawken and Horner (2003) in conjunction with training materials created and utilized by Leanne Hawken and the Office of Special Education (OSEP), which was funded by the Technical Assistance Center on Positive Behavioral Interventions and Supports (PBIS). Training materials also included strategies identified as adult learning principles. These approaches to training increase the likelihood that adult practitioners will learn and be able to apply skills and strategies presented during the training (Dunst & Trivette, 2009). Using adult learning principles and high quality materials increases the likelihood that the training will result in the implementation of CICO as it was designed (Bellg et al., 2004; Borrelli et al., 2005; Dunst & Trivette, 2009; Moncher & Prinz, 1991).

The second major area that must be considered when measuring fidelity is the use of treatment manuals and supervision (Bellg et al., 2004; Carroll et al., 2007; Moncher & Prinz, 1991). The Statewide PBIS initiative used three sources to serve as treatment manuals. First, school teams trained in CICO by the initiative receive one copy of the text written on CICO by Crone et al. (2010). This publication works as a manual, and is used throughout training to support learning and to serve as a guideline and resource for implementation once the training ends and teams return to their school to implement. Additionally, practitioners receive an example manual that a school created of the local systems, practices, and strategies created to implement CICO. This manual serves as another resource to guide implementation after training is over. Finally, the practitioners

receive a copy of the training materials to serve as a reference, resource, and a guide towards evidence-based implementation.

Ongoing supervision is another aspect of ensuring the precision of implementation. Due to the nature of the Statewide PBIS initiative with trained schools, the work in this area is more accurately described as coaching than as supervision. Once a school team is trained in CICO, a local implementation consultant hired and trained by the initiative coaches schools through the implementation of CICO. Implementation consultants were hired due to their expertise in PBIS and CICO, and are trained in the Statewide PBIS standardized approach to CICO as well as systems and instructional coaching practices that support increased implementation of trained materials at the district and school levels. Implementation consultants contact trained personnel on a monthly basis to check in regarding the implementation to ensure that the implementation is ongoing. Implementation consultants are also available when requested by school personnel.

Finally, in order to address fidelity assessment as a comprehensive approach to implementation measurement, a systemic approach to collecting and utilizing the fidelity data is needed (Moncher & Prinz, 1991). The Statewide PBIS initiative began utilizing CICO-FIM with schools that were trained during the 2012-2013 school year; this continues today. Archival data used in this study represents data from three cohorts of schools. These cohorts are defined as follows: Cohort 1 consists of schools trained during the 2012-2013 school year; Cohort 2 comprises schools trained during the 2013-2014 school year; and Cohort 3 includes schools trained during the 2014-2015 school year. The CICO-FIM instrument as previously described was developed by the authors of the

intervention, and was intended to measure the major activities necessary to implement CICO to criterion. The instrument requires the CICO team or CICO coordinator to self-report their implementation of CICO across 12 items on a scale of 0, 1, or 2. Schools are instructed to complete the initial evaluation of fidelity one to three months after training and then on an annual basis for each year implementation sustains. Cohort Three or schools trained during the 2014-2015 school year completed a CICO-FIM prior to training to serve as a pre-training baseline score. The instrument can be completed online in a form created on Adobe® FormsCentral or in paper and pencil format. The schools choosing to complete a hard copy of the CICO-FIM email or fax their responses to their implementation consultant. The implementation consultant then shares the data with The PBIS state-level personnel.

CICO-FIM data is then utilized in a number of ways to support increased, evidence-based implementation of CICO. Schools use the data to reflect on their implementation and plan areas to improve. Implementation consultants use the data to drive their coaching in the area of CICO and also to identify areas of need across their region for further training. Trainers in Cohort Three used the CICO-FIM instrument as a pre-assessment for training to assess previous knowledge of participants and to measure skills learned and used after training. Statewide PBIS initiative personnel use CICO-FIM data to assess the implementation of CICO across the state, identify regions and topical areas of need for future training, assess the efficacy of the training curriculum, and demonstrate the need for the project to funders.

## **Fidelity Assessment and Check-In, Check-Out**

Fidelity assessment is integral in the implementation of evidence-based interventions in schools. CICO is an evidence-based, targeted group intervention that does not have a fidelity assessment with demonstrated psychometric properties. This study purports to meet this void in the literature.

CICO needs a reliable and valid fidelity assessment to measure implementation for a number of reasons. First, CICO is widely utilized and has good results when implemented to the model (Campbell & Anderson, 2011; Fairbanks et al., 2007; Filter et al., 2007; Hawken & Horner, 2003; Hawken et al., 2007; March & Horner, 2002; Swoszowski et al., 2012; Todd et al., 2008). Second, numerous studies that have demonstrated the effectiveness of CICO also demonstrated a lack of consistency in the means to measuring fidelity. Maggin et al. (2015) investigated nine studies of the effectiveness of CICO, and while each measured fidelity of the intervention, no two studies used the same measurement, and thus cannot be compared. Furthermore, in the previous section on the efficacy of CICO, ten studies were presented. Each reported fidelity assessment, and each measured fidelity differently. Third, VanDerHeyden et al. (2007) presented a compelling argument that fidelity evaluation and documentation is essential in a tiered approach to ensure effective and efficient support for students.

A fourth reason that CICO needs a psychometrically sound fidelity assessment is due to purposeful intervention adaptation. Qualitative conversations with school-based professionals report that CICO is frequently implemented in a manner that is adjusted and adapted beyond the scope of its evidence-base, and thus were not implemented to fidelity. Because CICO is an intervention that is based on including contextually relevant

practices within the model of the intervention, any adaptation beyond this context is therefore an adaptation away from the evidence rather than towards a more contextually appropriate version of the intervention. Thus, any discussion of CICO requiring fidelity with flexibility is contrary to the model of the intervention itself (Hagermoser et al., 2009; APA, 2006). This is most often exemplified in the individualization of the intervention, rather than in keeping consistent across the entire targeted group as it was designed and measured.

Interventionist drift is another reason that CICO needs a valid fidelity assessment. Interventionist drift is the phenomena described as unintended changes made in the implementation of an intervention (Hagermoser et al., 2009). Thus, utilizing one universally available and valid fidelity measure of CICO is essential in the scientific and practical implementation of this school-based targeted group intervention.

Crone et al. (2010) created the CICO-FIM to measure implementation of the CICO intervention. The instrument contains 12 evaluation questions, each with a clearly defined, three-point scale. The CICO-FIM is a self-reporting tool that can be completed by the CICO coordinator or by the CICO team. The self-reporting nature of the tool addresses the social validity needs of efficiency and feasibility (Hagermoser et al., 2009; Sheridan, Swanger-Gagné, Welch, Kwon, & Garbacz, 2009; McKenna, Rosenfield, & Gravois, 2009). Additionally, self-report, provides on-site performance feedback, which is consistent within the PBIS model and thus provides a greater opportunity to build on the behavior momentum of this collection. Self-assessment data by school personnel is reliable (Hagermoser et al., 2009). Furthermore, fidelity assessments with acceptable psychometric properties have been found to increase the accuracy of self-reported data

from school personnel (Biggs et al., 2008). This study examines the psychometric properties of the CICO-FIM to determine if this tool can be the universally available and valid measure essential to support broad implementation of CICO in schools.

### **Summary**

A review of the research and literature across the domains of a multi-tiered behavioral framework of prevention, Check-In, Check-Out (CICO), and the critical role of implementation fidelity were covered in this chapter. Additionally, a rationale was presented on why a study is needed to examine the psychometric properties of a new fidelity measure associated with CICO, the CICO-FIM. Experts in the field of behavioral support and CICO created this fidelity instrument to advance research and the practical application of CICO in schools. A psychometrically sound fidelity instrument for CICO intervention may also help lead to increased positive outcomes for students receiving this intervention. A reliable and valid CICO fidelity instrument may help maximize student outcome and minimize student harm.

## **CHAPTER 3: METHOD**

This study examined the psychometric properties of the Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM). Check-In, Check-Out (CICO) is a targeted group intervention implemented at schools to prevent problem behaviors in students at-risk for school failure. The CICO-FIM is an instrument used to annually track the implementation of CICO by educators within a school. The CICO-FIM was designed to measure the essential practices of the school professionals in order to implement the CICO intervention as it was defined. This chapter presents the methodology that was utilized to investigate the properties of the CICO-FIM, including the study design, participants, instrumentation, procedures, and data analysis.

### **Study Design**

This study was the first investigation into the validity, factor structure, and reliability of the CICO-FIM. This study utilized archival data collected from schools participating in the statewide Positive Behavioral Interventions & Supports (PBIS) initiative out of a department of education in a western state. Participating schools were trained by the initiative in CICO and the use of the CICO-FIM from September 1, 2012 to May 31, 2015, yielding three cohorts of data. In Cohorts 1 and 2, the instrument was completed and collected at baseline (one to three months after training and initial implementation) and then annually as implementation was sustained. As a pre-assessment, Cohort 3 completed the instrument prior to training and then it was also

administered at baseline, one to three months after training, during initial implementation. This data were then used to answer the four psychometric research questions posed for this study to investigate the validity (using content validity, construct validity, and factor analysis) and reliability (using internal consistency and interrater reliability) discussed in more detail below.

### **Participants**

The overall sample that constituted the archival data sets used in the study and the specific sampling procedures used to address each research question are discussed below.

#### **Population**

As previously stated, CICO is one targeted group intervention utilized within the PBIS framework. According to the Office of Special Education funded Technical Assistance Center on PBIS, over 26,000 schools across the United States have been trained in the practices and interventions of PBIS (Educational and Community Supports, 2012). Due to the size and geographic area of this population, it is difficult to sample. Thus, a western state that provides professional development, technical assistance, and coaching for PBIS and CICO was used as the population. This state has trained 1,000 schools in the universal implementation of PBIS since 2001 and 119 schools in CICO since 2012.

Fidelity assessment is a standard practice across PBIS. The collection of fidelity data at the universal tier is a requirement to receive CICO training from the PBIS initiative of this state. During CICO trainings, participants are introduced to the CICO-FIM instrument and given directions on how to complete the instrument and the schedule for its completion. Of the 119 schools trained in CICO since September 2012, only 79

schools completed the CICO-FIM, which is a 66% response rate. For these 79 schools, the CICO-FIM was voluntarily completed and turned in as part of their involvement in the PBIS initiative at the state department. Table 1 describes the race and ethnicity of the students that attend the 79 schools utilized in this study. Table 2 describes the same sample of schools according to their grade level served and geographic region.

Table 1

*Race/Ethnicity of the Students in the Sample as Compared to the Students of the State*

Race/Ethnicity	Sample of 79 Schools		State	
	N	%	N	%
American Indian or Alaskan Native	242	0.65	6,537	0.74
Asian	649	1.73	27,297	3.07
Black or African American	2,234	5.96	41,660	4.69
Hispanic or Latino	12,552	33.52	294,435	33.11
White	20,391	54.45	484,305	54.48
Native Hawaiian or Other Pacific Islander	86	0.23	2,065	0.23
2 or more Races	1,293	3.45	32,707	3.68
Total	37,447		889,006	

Table 2

*Number of Schools within the Sample by Grade Level and Region*

Grade Level	Regions of the State							Total	
	Metro Area	North Central	NE	NW	South Central	SE	SW		West Central
Elementary Schools (ES)	18	23			8		1	9	59
ES/MS	1								1
ES/MS/HS		1						1	2
Middle Schools (MS)	2	4				1	1	6	14

MS/HS High School (HS)											0
									3		3
Total	21	28	0	0	8	1	2	19			79

The CICO-FIM was administered to schools prior to training, at baseline and annually thereafter, depending on the participating cohort. Schools in Cohort 1 (N=55) were trained in CICO during the 2012-2013 school year, and data were collected one to two months after initial implementation, in the first year of implementation (baseline), and each year after initial implementation, when implementation continued (annually). Likewise, schools in Cohort 2 (N=41) were trained in CICO during the 2013-2014 school year, and collected baseline and annual data when applicable. Finally, schools in Cohort 3 (N=23) were trained in CICO during the 2014-2015 school year, and data were collected prior to training, at baseline, and annually thereafter. Of the total 79 schools across the three cohorts, 37 schools had two or three submissions of the CICI-FIM. This is a result of school professionals from Cohorts 1 and 2 completing annual submissions of the instrument for their schools after the initial collection. This resulted in a total of 120 complete instruments collected from 79 schools across the three years of the archival data set.

This sample of 120 completed cases from 79 schools is a non-probability, convenience sample because the CICO-FIM data included in the study were collected and turned in by school professionals that voluntarily agreed to participate in the use and collection of this instrument. Convenience sampling is appropriate for this study for three reasons (Fowler, 2009). First, the size of the sample frame for the study (119 schools

trained in CICO) was too small for probability sampling. Second, the CICO-FIM instrument has not yet been psychometrically validated, and therefore has not been utilized in large-scale implementation nationally or in this western state. Thus, the convenience sample represents the only data available to carry out this psychometric study. Third, the state PBIS initiative followed clear procedures for training, collecting, and using the data that were commensurate with research in the collection and utilization of fidelity assessment data (Carroll et al., 2007; Moncher & Prinz, 1991).

The convenience sample of 79 schools with 120 completed instruments described above and in Tables 1 and 2 is the primary sample used within this study. However, other sample sizes and samples were used to investigate the four research questions posed to examine the psychometric properties of the CICO-FIM. The sample employed to address the first research question consisted of six content experts in conjunction with the sample of 120 completed cases to examine content and construct validity. The second and third research questions also employed the sample of 120 completed cases to investigate the factor structure and internal consistency of the CICO-FIM. The sample used to address the fourth research question was obtained from only 27 schools drawn from the 79 schools in the previously described convenience sample, resulting in 27 CICO-FIM instruments completed within the previous year. In addition, a second outside rater also completed forms for the same schools in order to investigate interrater reliability (IRR). The research questions for this study, along with the respective sample size and purposes of data analyses are presented in Table 2. Additional, information about the sampling procedures employed for each research question are described in more detail below.

Table 3

*Research Questions, Sample Size, and Data Analysis Means*

<b>Research Questions</b>	<b>Sample Size</b>	<b>Data Analysis</b>
Does the CICO-FIM demonstrate adequate content validity?	6	Mean and standard deviation per item across reviewers.
	120	Pearson correlation coefficient
What is the CICO-FIM factor structure?	120	Principal component analysis (PCA)
Does the CICO-FIM demonstrate adequate internal consistency?	120	Cronbach's alpha
Does the CICO-FIM demonstrate adequate interrater reliability?	27	Spearman's correlation & intraclass correlation coefficients

**Research Question 1 Sampling: Content and Construct Validity**

The first research question within this study includes two sub-questions that each utilized six content experts. The first sub-question examined content validity. These experts were recruited to provide data regarding the relevance, clarity, and difficulty of the items in measuring the implementation of CICO (Polit & Beck, 2006). Experts were chosen based on the criteria listed in Table 3. No one associated with the creation of the instrument was enlisted in the expert review. This sample includes researchers and practitioners. Together, six professionals from across the country satisfied the criteria set forth. The identified professionals were asked via email to participate in the study (see Appendix D for the email invitation). Each professional agreed to participate as a content expert. Three white females and three white males reviewed the items of CICO-FIM for relevance, clarity, and difficulty. Four of the experts hold a PhD. and work in academia;

one holds an EdD and works at an educational organization; and two have Master's degrees in school psychology and work in school districts in the western state studied here.

Table 4

*Criteria Utilized to Identify Content Experts*

#	Criteria
1.	Expertise and experience in psychometric measurement.
2.	Expertise and experience in school-based prevention.
3.	Expertise and experience in CICO.
4.	Expertise and experience in fidelity assessment in schools.
5.	Used the CICO-FIM in schools.

The second sub-question examined construct validity. This examination used focused association between two samples. The first sample was the six content review experts, while the second sample was the 120 completed instruments from 79 schools described previously and in Table 1.

**Research Questions 2 and 3 Sampling: Factor Structure and Internal Consistency**

The previously described non-probability convenience sample with 120 instruments completed from 79 schools was used to address the factor structure and internal consistency questions of this study. Of the 79 schools that completed the CICO-FIM, 37 schools completed the instrument two or three times for a total of 120 completed instruments. The repeated measures from some schools were included in this sample because it was assumed that each set of scores from each school was independent. This is

because the schools that administered the instrument more than once did so with different individuals or teams over time. Tinsley and Tinsley (1987) reported that the sample size necessary for the investigation of factor structure requires five to ten participants per item. Since the CICO-FIM is a 12-item instrument, the sample size of 120 completed instruments is within the appropriate range to measure factor structure. This sample of 120 completed instruments is also appropriate to estimate internal consistency. Cronbach's alpha was used to measure internal consistency in this study, which necessitates a minimum sample of 30 cases (Xu & Lorber, 2014). The use of the sample of 120 cases surpasses this guidance, thus increasing the likelihood of capturing an accurate alpha (Xu & Lorber, 2014).

#### **Research Question 4 Sampling: Interrater Reliability**

The convenience sample used in the first three questions of the study was also used to address interrater reliability. As described in the forthcoming procedures section the local ratings of school level professionals were compared with that of an outside rater for 27 schools. These 27 schools were a subset of the 79 schools utilized in the first three research questions described in Table 1. The 27 schools represented each cohort of implementation equally, and are located within a same region of the state. This region was identified by the statewide PBIS initiative as an optimal environment to test the interrater reliability of the CICO-FIM because it has characteristics that control for extraneous variables, including: equal number of schools from each cohort, the schools in the region have collected CICO-FIM data since initial training, the schools have district-level coaches, and the region has had a highly qualified technical assistance coordinator with extensive experience in CICO. Table 5 describes the race/ethnicity of the students

that attend the schools in the sub-sample utilized to address this research question as compared to the students enrolled in the state.

Table 5

*Race/Ethnicity of the Students in the 27 School Sample as Compared to the Students of the State*

Race/Ethnicity	Sample of 27 Schools		State	
	N	%	N	%
American Indian or Alaskan Native	46	0.41	6,537	0.74
Asian	116	1.04	27,297	3.07
Black or African American	106	0.95	41,660	4.69
Hispanic or Latino	2,396	21.4	294,435	33.11
White	8,236	73.58	484,305	54.48
Native Hawaiian or Other Pacific Islander	17	0.15	2,065	0.23
2 or more Races	277	2.47	32,707	3.68
Total	11,194		889,006	

Table 6

*Number of Schools within the 27 School Sample by Grade Level*

Grade Level	Number of Schools
Elementary Schools (ES)	23
ES/MS	
ES/MS/HS	1
Middle Schools (MS)	3
MS/HS	
High School (HS)	
Total	27

## **Consent from Participants**

Due to the archival nature of the data used in this study, participants did not provide informed consent prior to data collection and the data in this study were collected completely voluntarily as part of the statewide PBIS Initiative. Participants were informed that the purpose of data collection was to support improved school and district-level implementation and to support the statewide initiative in evaluating its training and technical assistance in CICO. The data were shared with the researcher after both the University of Denver's Institutional Review Board (IRB) and the state department of education's research review board granted exempt status. The data collected from the content experts did not require consent according to the University of Denver IRB. However, since participation in the study was voluntary, completion of the protocols denoted implied consent.

## **Instrument**

One instrument was employed in this study, the CICO-FIM (Crone et al., 2010), which was designed to measure the implementation of the CICO intervention. The historic development and description of the content of this instrument is presented followed by a description of how the instrument was completed by school professionals at the school-level.

## **Development of the CICO-FIM**

The CICO-FIM was constructed by Crone et al. (2010) and first published in *The Behavior Education Program: A Check-In, Check-Out Intervention for Students at Risk* as a practical self-assessment for school personnel to measure the implementation of CICO to criterion (see Appendix A). The CICO-FIM was developed to be conceptually

grounded and empirically derived to guide the creation, implementation, and continuous improvement of CICO. The authors first included the instrument in the second edition of their text on CICO. The instrument has not been used in previous studies to measure the implementation of CICO (Maggin et al., 2015).

### **Description of the CICO-FIM**

The CICO-FIM is a 12-item self-assessment that was designed to be a practical and efficient means of collecting fidelity assessment on CICO implementation at the school level. The instrument begins by collecting general data on the school from which the data are collected. These data include the name of the school, district, and state identifier as well as the date of completion. Additionally, the instrument requests the name of the data collector (a team or individual) and information describing the phase of implementation at the school level including pre-, baseline, or annual implementation. The instrument contains 12 evaluation questions; each is scored 0, 1, or 2, and each evaluation question has a specific definition for the 0, 1, or 2. Each question also provides the data collector with a suggested data source from which to glean the score. The data sources include permanent product, interview, and/or observation. Different sources of data are suggested for each question. The instrument can be completed utilizing a paper-pencil hard copy (Appendix A), a word processing document (Appendix A), or on a statewide initiative hosted online form (Appendix B).

The initial version of the instrument did not come with instructions. During the adoption of the instrument, the statewide PBIS initiative provided simple instructions at the beginning of the instrument in order to support the accurate collection of data using the instrument. The instructions describe the purpose of the instrument as a means of

measuring fidelity of implementation of CICO. The instructions also define who should complete the CICO-FIM and how often. First, the evaluation questions should be completed by a CICO coach or by the CICO school-based team. Second, the directions define the frequency of completion of the CICO-FIM by participating schools at pre-implementation, baseline or at one to three months after training, and then annually for each year of implementation.

### **Completion of the CICO-FIM**

The following discussion describes how the CICO-FIM is completed at the school as well as how this data was shared with the state PBIS initiative. As described in the directions, the CICO-FIM is intended to be completed either by a school-based team in charge of the operation and implementation of CICO, or by a school-based coach who is responsible for supporting the aforementioned team in the implementation of CICO. The team or coach responsible for CICO implementation completes the self-assessment by rating the school's implementation on a 0, 1, or 2 scale for each item. Each item has a corresponding data sources prompt for the team or coach to use to assess their implementation on the item. If a team completes the tool, the team must come to a consensus and provide one answer per question.

### **Procedures**

This study used archival data collected as part of a western state's PBIS initiative to examine the construct validity, factor structure, internal consistency, and interrater reliability (IRR) of the CICO-FIM. The procedures utilized to investigate each of the four research questions of this study are described below.

### **Research Question 1 Procedures: Content and Construct Validity**

The first research question for the study concerns investigating the content and construct validity of the CICO-FIM using content expert review. Once the criteria for identifying the experts were set, the principal investigator contacted professionals that met the criteria. Communication was initiated with each expert via email. The script of said communication can be found in Appendix C. After each expert agreed to participate, the content validity protocol was emailed to the expert with instructions for completion and a request to return the completed protocol within two weeks. A blank content validity protocol can be found in Appendix D. The protocols were sent to the experts on May 8, 2015; all were completed and returned by May 22, 2015. Experts emailed the completed protocols directly to the principal investigator for analysis. The investigation into construct validity in this research question also utilized the sample of 120 completed instruments from 79 schools. The description of the procedures for collecting this data is described below.

### **Research Questions 2 and 3 Procedures: Factor Structure and Internal Consistency**

The second and third research questions investigated the factor structure and internal consistency of the CICO-FIM using the previously identified archival data set that included a sample of 120 completed instruments from 79 schools. As previously described, in the Instrument section within this chapter, the CICO-FIM is completed at the school-level.

Once the school-level CICO team or coach has completed the instrument, the school professionals voluntarily shared the completed instrument with the technical assistance provider from the state PBIS initiative. The school professionals submitted

data either by an online form hosted by the statewide initiative or by emailing the completed instrument in Portable Document Format (PDF) or Microsoft Word. In order to procure these data at the state level, the technical assistance provider delivered ongoing reminders and prompts for schools to collect and use the data for action planning, ensuring implementation to the model. The technical assistance providers were rigorous in their pursuit of the data. The state PBIS initiative and technical assistance providers work in service of and as a support to districts and schools, and thus cannot require schools to collect data. However, the initiative and its technical assistance providers make a strong case for collecting and utilizing the CICO-FIM, and require data for follow-up training on CICO. This yielded a total archival data set of 79 schools, 37 of those schools completed the instrument two or three times, for a total of 120 unique completed instruments.

The data were collected by the state initiative in order to address the two organizational needs of program evaluation and performance assessment. As program evaluation, the completed instruments served as evidence of school professional behavior change as a result of training. These data are used to evaluate training and support the funding of future trainings. As performance assessment, the data were also used to provide information to support future training and technical assistance needs in the area of CICO. The state PBIS initiative did not previously use any other fidelity tool prior to this study, due to the absence of any psychometrically sound instrument to measure the implementation of CICO.

#### **Research Question 4 Procedures: Interrater Reliability**

An interrater estimation of reliability of the CICI-FIM measure was the focus of the fourth research question. This question was addressed using a convenience sample of 27 schools previously described in the Participants section of this Chapter. The following procedures describe the means utilized to collect data from schools and an outside rater.

In the spring of 2015, schools within this region were given a four-week window to complete their annual CICO-FIM either online or in paper-pencil format. Schools that completed the CICO-FIM using paper and pencil emailed their results directly to statewide PBIS initiative personnel. During the same four-week period, a technical assistance provider from the statewide visited each school and completed a CICO-FIM. The outside rater was trained in the collection and use of the instrument prior to collection, and completed the CICO-FIM at the school using the data sources suggested per item.

For items that required an interview, the outside rater interviewed the school-based CICO coach. Interviews were not performed with an administrator to minimize the bias caused by the social desirability of responses to items that implicated administrator responsibility. To further minimize social desirability bias, the outside rater provided clear instruction to the participating school personnel that the purpose of the rating was to find out more information regarding the instrument, not the school's performance on the instrument. The outside rater and school personnel were careful to ensure that their responses on the CICO-FIM were blind. The outside rater received additional training and guidelines on the administration of the CICO-FIM to limit experimenter bias. Finally, the outside rater utilized the paper and pencil format during the administration of the

instrument. After the outside rater completed each administration, the resulting CICO-FIM was scanned and emailed to statewide PBIS initiative personnel.

### **Data Analysis**

Data analysis included three phases: 1) data preparation and screening, 2) calculation of descriptive statistics, and 3) analysis of data pertinent to the four research questions. The general procedures followed during each of these phases are discussed below. The associated outcomes for each phase and the specific results obtained from each analysis are presented in the Results Chapter.

#### **Data Preparation, Screening, and Descriptive Statistics**

The first phase of data preparation and screening ensured that prior to analysis the data were reviewed for accuracy of entry, outliers, and missing data. The second phase focused on calculating descriptive statistics for the largest sample of 120 completed CICO-FIM instruments. Descriptive statistics were obtained for item means, medians, standard deviations, and ranges. The *IBM Statistical Package for the Social Sciences* (IBM, 2014) was utilized across each phase to examine the data and to conduct all final analyses.

#### **Research Question 1 Analysis: Content and Construct Validity**

The first research question examined the content and construct validity of the CICO-FIM. Content validity was addressed using descriptive statistics to depict item relevance, clarity, and difficulty for each item, as estimated by the content expert reviewers (Furr & Bacharach, 2014). Specifically, the item relevance and clarity means and standard deviations computed for each item across expert reviewers represented the perceived content validity for each item (Furr & Bacharach, 2014). Construct validity

was investigated in this research question through focused association. Focused association between the predictor and criterion item means, where the predictor item means were the expert responses per item on item difficulty and the criterion item means were the item means from the 120 completed case sample. The focused association is the correlation of the predictor item means with the criterion item means (Furr & Bacharach, 2014). Since these criteria are highly relevant to one another, this association provides more evidence regarding validity (Furr & Bacharach, 2014). Focused association was estimated using a Pearson correlation coefficient. For this study, a .8 coefficient was deemed adequate for a strong association between the predictor item difficulty and the criterion or empirically derived item difficulty from the sample.

### **Research Question 2 Analysis: Factor Structure**

The second research question examined the factor structure of the CICO-FIM using exploratory factor analysis (EFA). Prior to EFA, the sample was examined for assumptions necessary to investigate factor structure. The assumptions of normality and linearity were investigated using skewness, kurtosis, and scatterplots. Once the assumptions were established, EFA was investigated using a principal component analysis (PCA; Furr & Bacharach, 2014). PCA was the most appropriate analysis for the CICO-FIM because the instrument had not previously been examined for latent factors (Fabriger, Wegener, MacCallum, & Strahan, 1999). Moreover, PCA allows components to account for variances, rather than latent factors accounting for correlation, which is the case in principal axis factoring (Russell, 2002). PCA evaluates factor structure as well as item contribution and loading. Due to the unidimensional structure of the CICO-FIM, no

rotation procedure was necessary to simplify the data structure (Russell, 2002). Item loadings used the .4 or greater criterion for retention (Tinsley & Tinsley, 1987).

Before interpretation of EFA, the appropriateness of the use of factor analysis was established (Tabachnick & Fidell, 2007). Factorability was assessed using the Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy test and the Bartlett's Test of Sphericity (Tabachnick & Fidell, 2007). To interpret EFA, the KMO should be equal to or greater than .6, and the Bartlett's test should be statistically significant (Tabachnick & Fidell, 2007). The matrix also needed several substantial correlations to ensure factorability (Fabriger et al., 1999). The number of factors to interpret was determined by using a combination of inspection of eigenvalues and parallel analysis (Fabriger et al., 1999). Finally, communalities, or the portion of variance that is accounted for by common factors, were examined after extraction (MacCallum, Widaman, Zhang, & Hong, 1999).

### **Research Question 3 Analysis: Internal Consistency**

Reliability was first estimated using internal consistency. Internal consistency was estimated for the total CICO-FIM, using Cronbach's alpha. Cronbach's alpha was chosen due to the presence of multiple response options for each item in the instrument (Furr & Bacharach, 2014). The item-total statistics table was used for analysis at the item level. For this study, an internal consistency .7 or greater was considered acceptable.

### **Research Question 4 Analysis: Interrater Reliability**

Interrater reliability was obtained between the school-based rating and an outside rater. A Spearman Rho or rank order correlation was used to summarize the relationship between ratings of school-based raters and the outside rater. This correlation was used

because it measures the strength of association for ordinal scale data between the two raters, and reduces the impact of outliers (Bobko, 2001). For this study, a Spearman's Rho of .7 or greater was considered acceptable (DeVellis, 2003). Additionally, intraclass correlation coefficients (ICC) were used to assess the consistency of the ratings between the two different raters. An ICC found to be greater than .7 was considered adequate IRR for this study.

## **CHAPTER 4: RESULTS**

The major purpose of this study was to examine the psychometric properties of the Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM) instrument used by CICO teams or coordinators to measure the extent to which school professionals at a local level are implementing the components of CICO. Presented below are the outcomes associated with the preparation and screening of the original archival data set; the descriptive statistics for the 79 school sample with 120 completed instruments used to analyze the CICO-FIM; and finally, the results of the analyses associated with the four research questions posed in this study.

### **Data Preparation and Screening**

The first phase of data analysis was preparing and screening the data using SPSS (IBM, 2014). This exploratory data analysis ensured the accuracy of data entry, outliers, and missing data across each of the samples of data utilized for each research question. The data were checked by hand for errors in addition to using SPSS. After the raw data were entered into SPSS, the SPSS file was checked against the raw data to ensure that data entry errors were not made. Descriptive statistics were run, and the output was examined for outliers and missing data; none were observed. The exploratory data analysis phase did not include a test of the assumptions of the overall research questions because each statistic used within the questions has its own specific assumptions. Thus, in this chapter, each test of assumptions is presented with its statistical test.

## **Descriptive Statistics**

The second phase of data analysis focused on descriptive statistics. SPSS was used to calculate item means, medians, standard deviations, and ranges for each item, and utilized the sample for questions that investigated EFA and internal consistency, as well as a portion of the content expert review in construct validity. Table 4 presents the descriptive statistics for each item for this sample.

As described in Chapter 3, each CICO item was scored on a 0, 1, or 2 scale, where the scale is defined specifically for each item. Generally, a score of 2 represents meeting the highest standards of the item describing the implementation, and 0 represents the lowest level of implementation. In the sample used to address focused association, factor analysis, and internal consistency, the means were centered on a score of 1 and range from 0.60 to 1.51. Item 6 had the highest mean (1.51) and a standard deviation of .71. The mean of 1.51, the median of 2, and mode of 2 on this item revealed that a majority of participants reported that 90% of students in the CICO intervention check-in on a daily basis. Item 2 had the lowest mean (.60) and a standard deviation of .68. The mean of .60, median of 0, and mode of 0 on this item revealed that a majority of participants reported that their school budgets contained little to no money to maintain the CICO intervention. This item also had the least variability. Examining frequency distributions revealed that 67.50% of participants responded to item 2 with a 0; this was the lowest rated response of any item.

Table 7

*Item Description, Means, Medians, Mode, and Standard Deviations of the Largest Sample*

<b>Items</b>	<b>N</b>	<b>Mean</b>	<b>Median</b>	<b>Mode</b>	<b>Standard Dev</b>	<b>Range</b>
1. CICO coordinator employed	120	.66	1	0	.68	2
2. Budget	120	.60	0	0	.89	2
3. Support in week	120	1.35	2	2	.74	2
4. Administrator on CICO team	120	.94	1	1	.79	2
5. CICO taught	120	.95	1	0	.86	2
6. Check-in daily	120	1.51	2	2	.71	2
7. Check-out daily	120	1.37	2	2	.77	2
8. Students receive reinforcement	120	1.43	2	2	.80	2
9. Students receive feedback	120	1.32	1	2	.75	2
10. Receive feedback from parents	120	.68	2	0	.70	2
11. Enter DPR data	120	.83	2	0	.83	2
12. DPR data used	120	.81	2	0	.77	2

### **Research Question 1 Results: Content and Construct Validity**

The first research question investigated the content and construct validity by utilizing expert review. A review panel of six experts was enlisted to address two sub-questions. The first sub-question examined content validity. Six experts in Positive Behavioral Interventions and Supports (PBIS), Check-In, Check-Out (CICO), and fidelity from universities, professional organizations, and school districts completed a protocol that measured expert perception of item relevance, clarity, and difficulty. The means and standard deviations for item relevance and clarity served to appraise content validity. For the purpose of this study, high mean values were required to demonstrate content validity. The second sub-question examined construct validity through focused

association, which was examined by correlating the predicted item means with actual item means. The mean per item of the content experts' ratings of item difficulty served as the predicted item means, and was associated with actual item means from the sample of 120 cases from 79 schools. The association was estimated using a Pearson correlation coefficient.

### **Content Validity**

As previously described, the data were screened for accuracy, missing data, and outliers prior to analysis. Per the previously described procedures, no outliers or missing data were observed. Data analysis for the first sub-question on content validity examined item means and standard deviations per item for relevance and clarity, and were computed using SPSS. The examination demonstrated that item relevance and clarity had high means for each item. Item relevance means per item were between 1 and 2, and had a range of 1.50 to 1.83. Seven of the items had a mean of 1.83, with a standard deviation of .41, a median of 2, and a mode of 2. Examination of frequency tables for item relevance demonstrated that nine of the items had a median of 2 and a mode of 2; no item was given a score of 0. Item clarity means per item were between 1 and 2, with a range of 1.50 and 2.0. Item 2 was clear to all of the reviewers, as each rated the question with high clarity. Further examination of the frequency tables for item clarity revealed that ten of the items had a median of 2 and a mode of 2. Table 5 shows item means and standard deviations for item relevance and clarity based on expert review.

Table 8

*Item Relevance and Clarity Means and Standard Deviations*

Items	N		Relevance		Clarity	
	Valid	Missing	Mean	Standard Dev	Mean	Standard Dev
1. CICO coordinator employed	6	0	1.50	.55	1.50	.55
2. Budget	6	0	1.67	.52	2.00	.00
3. Support in week	6	0	1.67	.52	1.67	.52
4. Administrator on CICO team	6	0	1.50	.55	1.83	.41
5. CICO taught	6	0	1.83	.41	1.50	.55
6. Check-in daily	6	0	1.83	.41	1.83	.41
7. Check-out daily	6	0	1.83	.41	1.83	.41
8. Students receive reinforcement	6	0	1.83	.41	1.83	.41
9. Students receive feedback	6	0	1.83	.41	1.83	.41
10. Receive feedback form parents	6	0	1.50	.55	1.83	.41
11. Enter DPR data	6	0	1.83	.41	1.83	.41
12. DPR data used	6	0	1.83	.41	1.67	.52

**Construct Validity**

The second sub-question investigated the construct validity of the CICO-FIM instrument using focused association. The focused association compared the predictor and criterion item means. The mean values of the content experts on item difficulty served as the predictor item means, and the criterion item means were measured using the sample of the 120 cases that were utilized to investigate EFA and internal consistency. The association was estimated using a Pearson correlation coefficient. Table 6 shows the means and standard deviations for the predictor and criterion samples per item used in the focused association.

Table 9

*Item Means and Standard Deviations for the Predictor and Criterion for the Focused Association*

Items	Predictor		Criterion	
	Mean	Standard Deviation	Mean	Standard Deviation
1. CICO coordinator employed	1.50	.55	.66	.68
2. Budget	1.67	.52	.60	.89
3. Support in week	1.67	.52	1.35	.74
4. Administrator on CICO team	1.50	.55	.94	.79
5. CICO taught	1.83	.41	.95	.86
6. Check-in daily	1.83	.41	1.51	.71
7. Check-out daily	1.83	.41	1.37	.77
8. Students receive reinforcement	1.83	.41	1.43	.80
9. Students receive feedback	1.83	.41	1.32	.75
10. Receive feedback form parents	1.50	.55	.68	.70
11. Enter DPR data	1.83	.41	.83	.83
12. DPR data used	1.83	.41	.81	.77

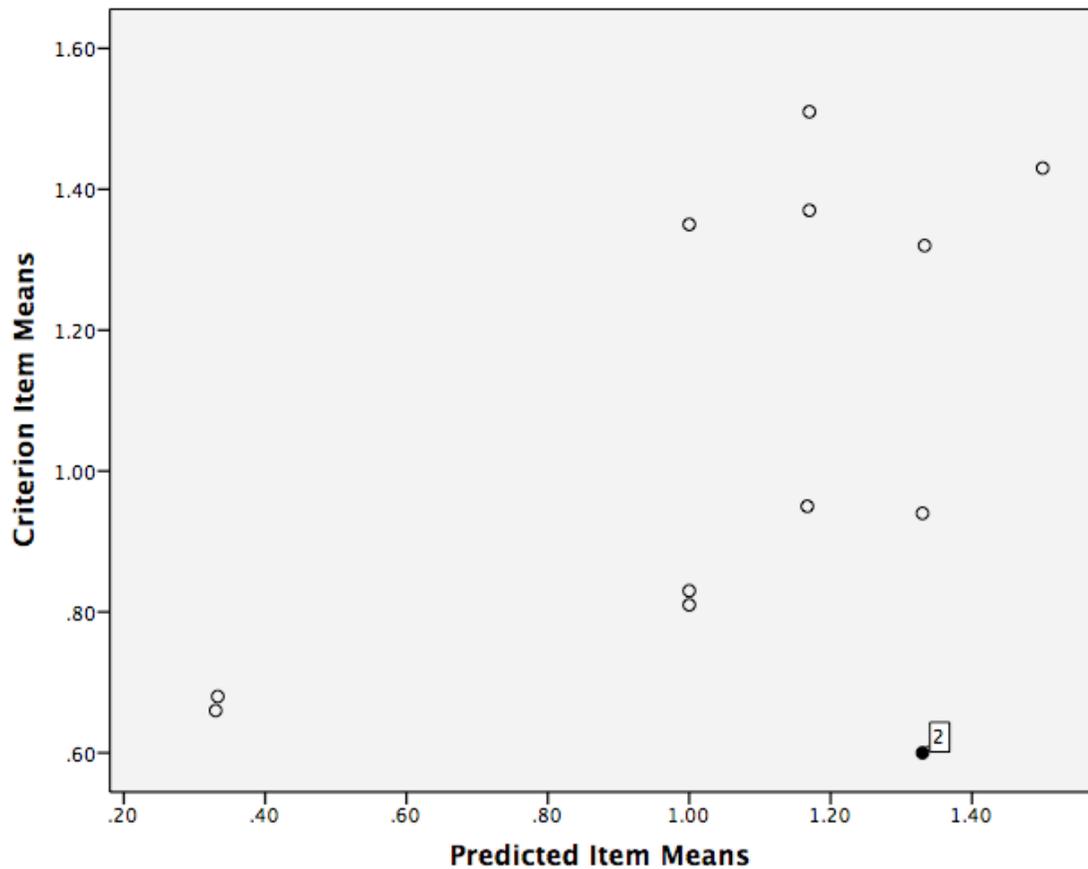
The Pearson correlation coefficient between the predictor and criterion means for the two samples was  $.53, p = .08$ . This represents a moderate, non-significant relationship between predicted and criterion item means. The scree plot provided in Figure 1 provides a visual demonstration of the relationship between predicted and criterion means. On the plot, Item 2 demonstrates the greatest mismatch between the actual and predicted means, with a predicted mean of 1.33 and criterion mean of .60. Item 2 asked the degree to which the school budget contained an allocated amount of money to maintain CICO.

The lack of association between predictor and criterion means was unexpected, and thus more investigation was necessary. First, open-ended response protocols were

examined to see if any comments specifically addressed Item 2 or other issues with item difficulty. No responses provided specifics regarding Item 2, with the exception of one expert. This expert suggested that there might be more specificity needed regarding the amount of money budgeted to maintain CICO, which would provide greater clarity to the item. The amount of money needed to maintain and sustain CICO might influence the difficulty of this item, but there is no current statistical evidence to support this appraisal. Second, expert responses were reexamined to include only those experts that were currently working at a district or school level. The assumption of this examination was that such experts understood the implementation of CICO at the school level better; therefore, they would have a greater association with the school-based responses. However, this examination yielded  $.43, p = .16$ —a smaller correlation. A third examination focused on the results of the scree plot. As previously stated, Item 2 was the most extreme outlier in this data set; therefore, the Pearson correlation coefficient was computed without Item 2 to determine the extent to which this outlier impacted the association. The Pearson correlation coefficient without Item 2 was  $.71, p = .02$ , representing a higher magnitude of association.

In summary, the results of the first research question that investigated content and construct validity utilizing content expert review were mixed. The examination of content validity through expert ratings of each item on a scale of 0, 1, or 2 for relevance and clarity asserted that the items of the CICO-FIM were relevant and clear. The examination of construct validity through the focused association of predicted and criterion item means showed that the predicted item means were derived from the six content experts' ratings on item difficulty, and that the criterion item means were derived from the actual

item means garnered from the sample of 120 cases utilized in the research questions, and were focused on factor analysis and internal consistency. The Pearson correlation coefficient ( $.53, p = .08$ ) was calculated with Item 2 removed; the correlation was found to be  $.71$ . Thus, the focused association between predictor and criterion for all 12 items had a moderate, non-significant correlation.



*Figure 1.* Scatter Plot of Predicted versus Criterion Item Means. This figure illustrates the coordinates of each item on its predicted and criterion means. Item 2 is shaded and denoted as it is the item with the largest difference between predicted and criterion means.

## **Research Question 2 Results: Factor Structure**

The second research question involved examining the factor structure of the CICO-FIM. The CICO-FIM has not yet been investigated for psychometric properties; thus, an EFA was the appropriate means to measure factor analysis. The data were screened as described in the beginning of this chapter; no outliers or missing data were found. The tests of assumptions for EFA are normality, linearity, independence, and outliers. Since the assumptions of independence and outliers have already been presented, the following section describes the results of tests of normality and linearity.

### **Test of Assumptions: Normality**

The assumption of normality is essential to utilize statistical inference to determine the number of factors of the CICO-FIM instrument (Tabachnick & Fidell, 2007). Skewness and kurtosis values of a normal distribution are 0.0 (Tabachnick & Fidell, 2007). A positive value for skewness indicates that the distribution is skewed to the right, and a negative value for skewness indicates that the distribution is skewed to the left. Values between -1 and 1 are deemed approximately normal and fit within the normal distribution (Leech, Barrett, & Morgan, 2015). A positive value for kurtosis indicates that the distribution of the data is too peaked, and a negative value for kurtosis indicates that the distribution of the data is too flat.

In this study, the assumption of normality was investigated by examining the skewness, kurtosis, and histogram for each item. The examination of skewness revealed that all of the items were within the range of -1 and 1 with the exception of Item 6, which showed at -1.10. Item 6 was not transformed to achieve normality because it was lightly skewed, and the instrument is published and in use in its current structure (Tabachnick &

Fidell, 2007). The examination of kurtosis revealed that items 2, 4, 5, 11, and 12 had negative kurtosis. This could be a result of the sample size, as negative kurtosis commonly disappears with samples of 200 or more; this sample had 120 cases (Tabachnick & Fidell, 2007). Since kurtosis does not typically affect the results of most statistics, kurtosis in this data set was ignored (Leech et al., 2015). Table 7 shows the values for mean, standard deviation, skewness, and kurtosis for each item.

Table 10

*Item Means, Standard Deviations, Skewness and Kurtosis*

Items	N		Mean	Standard Dev	Skewness		Kurtosis	
	Valid				Statistic	Std. Error	Statistic	Std. Error
1. CICO coordinator employed	120		.66	.68	.55	.22	-.74	.44
2. Budget	120		.60	.89	.88	.22	-1.16	.44
3. Support in week	120		1.35	.74	-.67	.22	-.89	.44
4. Administrator on CICO team	120		.94	.79	.11	.22	-1.39	.44
5. CICO taught	120		.95	.86	.10	.22	-1.64	.44
6. Check-in daily	120		1.51	.71	-1.10	.22	-.16	.44
7. Check-out daily	120		1.37	.77	-.74	.22	-.92	.44
8. Students receive reinforcement	120		1.43	.80	-.95	.22	-.76	.44
9. Students receive feedback	120		1.32	.75	-.59	.22	-.97	.44
10. Receive feedback from parents	120		.68	.70	.53	.22	-.83	.44
11. Enter DPR data	120		.83	.83	.32	.22	-1.49	.44
12. DPR data used	120		.81	.77	.35	.22	-1.23	.44

### **Test of Assumptions: Linearity**

The assumption of linearity is an essential assumption to exploratory factor analysis and parametric tests. A linearity test examines the assumption that variables are related in a linear way (Tab). To investigate linearity in this study, a matrix scatterplot

was created within SPSS to examine each item against each item to check for linearity of the sample specific to the investigation of exploratory factor analysis and internal consistency. The assumption of linearity with each item demonstrating a linear relationship with each corresponding item was not violated. Finding linearity in Item 6 supported the decision not to transform Item 6.

In summary, each research question sample was examined for accuracy of data entry, missing data, and outliers. This examination found accurate data entry, no missing data, and no outliers. The sample specific to the research questions examining exploratory factor analysis and internal consistency was tested for the assumptions of normality and linearity. Each item was normally distributed, with the exception of Item 6, which had slight negative skewness. The test of linearity found that each item had a linear relationship to other items, and no outliers were found. The assumption of independence was established with this sample, as each of the 120 cases provided the individual or team that completed the instrument. Of the 120 cases, 14 schools provided more than one case. Independence was established, because none of the repeated measures were completed by the same person or team. In summary, all statistical assumptions necessary for EFA were met with this sample.

### **Exploratory Factor Analysis**

To increase the ease of interpretation and to assess construct validity, principal components analysis (PCA) was used to investigate the grouping of the 12 items that describe CICO (Furr & Bacharach, 2014). PCA was used because it is more appropriate to examine an instrument that has not yet been assessed because it allows variables or items to account for the variance instead of latent factors (Tabachnick & Fidell, 2007).

Prior to this study, school-level CICO implementers used the CICO-FIM as a unidimensional instrument.

PCA was performed using SPSS on the 12 items for the sample of 120 cases. In order to determine the appropriateness of the use of factor analysis, the correlation of the items and of factorability were assessed. The correlation matrix demonstrated that the determinant is larger than 0.00 at .002, thus providing evidence that correlated factor analysis can be conducted. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy test was .86, which exceeds the minimum standard of .6 to demonstrate factorability. The Bartlett test was found to be significant ( $p < .001$ ), indicating that correlations are not near zero. All of these tests ensure that the sample demonstrated factorability.

Determining the number of components or factors is the first step in EFA. For the purpose of this study, the terms “factor” and “component” are used interchangeably. Though not completely accurate, this is consistent with the SPSS tests and output. This study used a combination of inspecting the initial eigenvalues in PCA, parallel analysis, and factor loadings to identify the number of factors underlying the responses to the instrument. PCA found two components with eigenvalues above 1.0. The first component had an eigenvalue of 5.15 with a 42.89% explained variance. The second component had an eigenvalue of 1.85 with 15.43% explained variance. Parallel analysis is assessed by comparing eigenvalues from randomly generated correlation matrices, with the eigenvalues extracted from the data set derived from this sample (O’Connor, 2000; Patil, Singh, Mishra, & Donovan, 2007). The parallel analysis for this study was developed using syntax to create randomly generated correlation matrices with eigenvalues as well

as the actual syntax utilized in SPSS (Patil, Singh, Mishra, & Donovan, 2007). Parallel analysis uses a Monte Carlo simulation that creates eigenvalues from 1,000 data sets at the 99th percentile for random and raw sample data (O'Connor, 2000). Raw data eigenvalues that are greater than the random data eigenvalues are considered appropriate possible components (O'Connor, 2000). For this study, this comparison demonstrated that two components had eigenvalues in the actual, raw data that were greater than the eigenvalue in the random data (O'Connor, 2000). Results of the initial eigenvalue and the parallel analysis were the same, which led to the conclusion that there could be one or, at most two, components extracted in the PCA process. The results of both statistics are found in Table 8.

Table 11

*Principal Component Analysis and Parallel Analysis Results of the CICO-FIM*

Component	Initial Eigenvalue		Parallel Analysis		
	Eigenvalue	% of Variance Explained	Raw Data Eigenvalues	Means	Random Data Eigenvalues
1	5.15	42.89	1.77	1.55	1.75
2	1.85	15.43	1.55	1.39	1.54
3	.95	7.89	1.39	1.28	1.40

Based on the results from a review of the factor loadings, one component was extracted. All loadings exceeded a typically used minimum value of .40. Factor loadings for this unidimensional model are given in Table 9. The one-factor model was extracted to achieve a parsimonious model.

Table 12

*Principal Component Loadings of the CICO-FIM*

<b>Items</b>	<b>Factor Loadings</b>
1. CICO coordinator employed	.46
2. Budget	.41
3. Support in week	.60
4. Administrator on CICO team	.50
5. CICO taught	.68
6. Check-in daily	.83
7. Check-out daily	.77
8. Students receive reinforcement	.79
9. Students receive feedback	.73
10. Receive feedback from parents	.60
11. Enter DPR data	.60
12. DPR data used	.76

A parsimonious model is a model that achieves a desired level of explanation with the fewest components possible. A one-component model for this extraction is a parsimonious model for the following reasons: First, one component was extracted because of the percent of explained variance. Component 1 accounted for 42.89% of the variance, while component 2 only accounted for 15.43% of the variance. A majority of the percent of variance was explained by the first component, and not enough was explained by component 2 to make it worth adding another factor to the model. Second, each of the factor loadings on the one-component model met the .4 or greater factor loading standard set for this study. This .4 standard was established in research to identify appropriate items loaded to a factor, and thus provides evidence that a parsimonious, one-factor model that includes all items is appropriate for this extraction (Tinsley & Tinsley, 1987). Third, the one-component model is appropriate for this extraction because the internal consistency of one factor is better than that of a two-factor model. As discussed in greater length in the third research question, the internal consistency of the one-factor

model is .87, while the internal consistency of the two-component model is .87 for the first factor and .80 for the second. Both models demonstrate adequate internal consistency, but the two-factor model has decreased consistency, which provides further evidence for a parsimonious one-factor model.

Finally, communalities—or the portion of variance in common with other variables—were examined after extracting the one-factor model. After extraction, communalities ranged from .17 to .68. This extraction found that all items met the .3 standard, except Items 1, 2, and 4, which had low ( $<.3$ ) communalities. Considering the robust results in the other areas of the EFA, these low communalities could be the result a small sample size.

### **Research Question 3 Results: Internal Consistency of the CICO-FIM**

The third research question investigated the internal consistency of the CICO-FIM. Internal consistency is the estimation of reliability that examines the extent to which items focus on the factor or factors measured by an instrument like the CICO-FIM (Furr & Bacharach, 2014). Items that are strongly related to each other are said to measure the same thing, and thus demonstrate internal consistency (Furr & Bacharach, 2014). The sample utilized to investigate internal consistency was the same sample used in the EFA. The sample had 120 cases across 79 schools, and was collected from September 2012 to June 2015.

As determined by the results of the EFA in the second research question, the CICO-FIM has one factor with 12 items; therefore, one measure of internal consistency was examined with all of the items. Cronbach's alpha was calculated to estimate the internal consistency of the 12 items of the one-factor CICO-FIM. The resulting

Cronbach's alpha was .87, which exceeds the standard set by this study that the CICO-FIM demonstrated adequate internal consistency with this sample. The item-total statistics table was examined for more information regarding the internal consistency of each item. Items with a corrected item-total correlation at or above .40 were considered acceptable. All item-total correlations were greater than .4, with the exception of Item 2. The correlation for Item 2 was .37. This item was included because, if the item were deleted, Cronbach's alpha would not increase. In summary, the one factor model of the CICO-FIM estimates reliable internal consistency across items in this archival sample was met with a Cronbach's alpha of .87. Additionally, each item demonstrated adequate internal consistency, as evidenced by the corrected item-total correlation for each item of .4 and above.

#### **Research Question 4 Results: Interrater Reliability for the CICO-FIM**

The fourth research question investigated the interrater reliability (IRR) of the CICO-FIM. The IRR investigation compared two observations of the implementation of CICO by using the CICO-FIM instrument to estimate the degree to which two observers estimated the same degree of implementation. The two observers were scores from the internal school professionals, which were compared to the scores of an outside expert in CICO-FIM. The sample for the IRR was 27 schools in the north-central region of this western state, with nine schools in Cohort 1, nine schools in Cohort 2, and nine schools in Cohort 3. The two observations were collected within two to three weeks of each other.

IRR was investigated using Spearman's Rho and intraclass correlation coefficients (ICC). Spearman's Rho summarized the relationship between the ratings of school-based raters and those of an outside, expert rater. For the sample, Spearman's Rho

was .89,  $p = .0001$ . Examination into the ICC first demonstrated that the within people-between items analysis of variance (ANOVA) was not significant, with an  $F = 3.28$  and a  $p = .08$ . This indicates that there was no significant difference between the means of the two observers. The index of IRR, determined by the ICC, was .84, with  $F(26) = 11.76$ ,  $p = .0001$ . This surpassed the standard (.7) set by the study. Thus, the sample demonstrated adequate IRR between the school raters and an outside rater.

### **Overall Summary of Results**

The initial phase of data analysis screened the archival data set for accuracy, missing values, and outliers across all three samples utilized to address the four research questions posed in this study. The results of this phase indicated that data were accurately entered and contained no missing values or outliers. Next, descriptive statistics were calculated for each sample utilized. The statistics for the largest sample had item means that centered around 1 with a range of 2, which also demonstrated adequate variance, linearity, and normal distribution.

Mixed results were obtained for the first research question that centered on an examination of CICO-FIM content and construct validity using expert review. First, the expert reviewers gave high item ratings for relevance and clarity. Thus, the hypothesis for content validity was confirmed, as the content expert review deemed each item relevant and clear. However, when construct validity was determined using focused association by comparing the predicted item means with criterion item means, using the sample of 120 cases, the predictor means only minimally and nonsignificantly correlated with the criterion item means (i.e., .53,  $p = .08$ ).

Results for the second research question were based on an examination of the factor structure of the instrument. PCA and parallel analysis established a one-factor model in support of the original hypothesis. This one-factor model met the assumption of item correlation, as evidenced by the determinant of .002, KMO of .86, and a significant Bartlett's test, at  $p < .001$ . Each of the 12 items loaded on the one-factor model and had loadings that met the standard of .4 or above and the one-factor model only had three items with low communalities, which were posited to be the result of the small sample size.

The results obtained for the third research question focused on internal consistency indicated there was high internal consistency in the one-factor model of the CICO-FIM. Internal consistency was analyzed for the 12 items of the CICO-FIM, and was found to be .87. The item-total statistics table was also assessed for the strength of the correlation between the item and the sum total. All items achieved the .4 or above standard, with the exception of Item 2. Item 2 was a slightly low, at .37. Deleting this item also did not improve Cronbach's alpha.

Finally, the results of the analysis associated with the fourth research question assessing interrater reliability indicated there was high consistency of CICO-FIM ratings by local level school professionals and an outside rater using a sample of 27 schools. The Spearman's Rho and ICC obtained suggested that the school and outside observers reached a high level of agreement on the items on the CICO-FIM.

## **CHAPTER 5: DISCUSSION**

This study was a psychometric investigation of the Check-In, Check-Out Fidelity of Implementation Measure (Crone et al, 2010). Overall, the results of this dissertation suggest that the CICO-FIM is a psychometrically promising instrument. In this chapter, the purpose and design of the study are reviewed and the major findings of the investigation are discussed. Notable implications of the study, limitations, and suggestions for future research follow. The chapter concludes with a discussion of fidelity assessment and the role of fidelity instruments like the CICO-FIM within large-scale school reform efforts and future education policy, especially in addressing the disparities and inequalities observed in many U.S schools.

### **Purpose and Design of the Study**

The purpose of this study was to ensure that the Check-In, Check-Out (CICO) intervention is being implemented with fidelity in order to prevent students at-risk of school failure from engaging in severe behavior problems, thus preventing the use of suspensions and expulsions. Utilizing suspensions and expulsions as a means of punishment is essentially flawed (Skiba & Sprague, 2008). Suspensions and expulsions do not extinguish problem behavior, and students that receive suspensions and expulsions are less likely to read at grade level and are more likely to dropout of school and be involved in the juvenile justice system (AAP, 2013; APA, 2008, 2006; Skiba & Sprague, 2008). Furthermore, students of color are more likely than their white peers to receive

suspensions and expulsions during their academic careers (Skiba et al., 2014; Skiba et al., 2011; Skiba et al., 2002; Tobin & Vincent, 2011). Focusing on preventing problem behaviors is a key component in decreasing the use of suspensions and expulsions in schools (Skiba & Sprague, 2008). Check-In, Check-Out (CICO) utilizes a systematic approach to provide increased rates of positive adult attention and feedback in order to prevent incidences of severe problem behavior in students that are at-risk (Crone, et al., 2011; Hawken & Horner, 2003). More to the point, the implementation of CICO at the school level is effective in reducing the number of office discipline referrals and suspensions in students receiving the intervention (Hawken et al., 2007; Lane et al., 2012; Miller et al., 2015; Todd et al., 2008). Measuring the fidelity of the implementation of school-based interventions is a burgeoning practice utilized to ensure interventions are implemented as designed and thus achieving intended outcomes (Moncher & Prinz, 1991; VanDerHeyden et al., 2001). To continue to bring this intervention to scale to prevent suspensions and expulsions, a psychometrically sound measure of fidelity is critical in supporting the research and practical application of CICO.

This study was designed to investigate the psychometric properties of the CICO-FIM utilizing an archival data set collected from a western state. This dissertation utilized three samples and six content experts to investigate four research questions. These research questions assessed the content validity, construct validity, factor structure, internal consistency, and interrater reliability (IRR) of the instrument.

### **Major Findings**

Four specific research questions were addressed in this study in order to assess the psychometric properties of the CICO-FIM. The first research question addressed content

and construct validity by utilizing content expert review from a sample of six experts from the fields of PBIS, CICO, and implementation fidelity. The second research question addressed construct validity by assessing the factor structure of the CICO-FIM across the total sample of 120 completed instruments. The third research question queried the internal consistency of the CICO-FIM with the sample of 120 completed instruments. The fourth research question investigated the IRR of the CICO-FIM utilizing two observers across 27 schools.

The assessment of content validity utilized responses on item relevance, clarity, and difficulty from content experts in the fields of PBIS, CICO, and fidelity assessment. The construct validity, factor structure, and internal consistency of the instrument were examined using an archival data set of 120 completed instruments from 79 schools. The IRR was assessed using a sample of 27 schools taken from the 79 schools, each with two observer ratings per school. One observer for the IRR was the school-based team or coordinator responsible for implementing CICO in each school, and the second observer was an outside expert in CICO and the CICO-FIM instrument.

### **Research Question 1 Major Findings: Content and Construct Validity**

The first research question addressed content and construct validity of the CICO-FIM, as evidenced by the assessment of content expert review and focused association. The first sub-question assessed content validity by examining the content expert review protocol item means for relevance and clarity. As expected, the item means for relevance and clarity were high demonstrating adequate content validity. These ratings suggest that the items are relevant to the implementation of CICO and easily understood by CICO implementers.

The second-sub-question examined construct validity through an investigation of focused association between the predicted and criterion item means. The predicted item means (item difficulty means from expert review) were correlated with the criterion item means (item means attained in the sample of 120 completed instruments). This unexpectedly found a moderate non-significant association. After further investigation, item 2 served as an outlier to this association. This was interesting to the researcher as item 2 centers on a budget for the implementation of CICO at the school-level. Considering the financial strain that many schools are experiencing, it is not surprising that schools find allocating resources towards a behavior intervention more difficult than experts.

### **Research Question 2 Major Findings: Factor Structure**

The second research question investigated the factor structure of the CICO-FIM. The analysis of the factor structure focused on exploratory factor analysis (EFA), parallel analysis, factorability, item loadings, and communalities. The results of the EFA and parallel analysis found that the CICO-FIM had no more than two components. Due to parsimony, the CICO-FIM was assessed to be unidimensional. Item loadings for the unidimensional model met the standard of .4 and above for each item. This parsimonious unidimensional model is consistent with how the instrument is currently used in the field.

### **Research Question 3 Major Findings: Internal Consistency**

The third research question addressed the internal consistency among the items that comprised the unidimensional structure found in the EFA. Cronbach's alpha was .87, indicating good to excellent internal consistency. An investigation of the item-total statistics table demonstrated that each of the items had a corrected item-total correlation

of .4 or above, with the exception of Item 2. Item 2 had a slightly low, at .37; however, the Cronbach's alpha would not increase if the item were deleted. Thus, the overall instrument and each item indicated good internal consistency. This also supports the parsimonious, unidimensional factor structure adopted in the second research question.

#### **Research Question 4 Major Findings: Interrater Reliability**

The fourth question assessed the IRR between the school-based ratings and the ratings of an outside expert of the 27 schools. IRR was measured by both the Spearman's Rho and intraclass correlation coefficient (ICC). Spearman's Rho for this sample was .89,  $p = .0001$ . The index of IRR assessed by the ICC was .84, with  $F(26) = 11.26$ ,  $p = .0001$ . Both of these test statistics demonstrated good, consistent IRR between the school-based and outside raters. Since the sample size for this question is small, the interpretations are made with caution. This limitation notwithstanding, these results support the use of self-assessment as consistent with outside experts in the measurement of fidelity assessment for this instrument with this sample.

#### **Implications of Results**

This investigation of the psychometric properties of the CICO-FIM has demonstrated promising results. Since this is an initial investigation, more research is needed to support these findings. The limitations of this study notwithstanding, the results denote an opportunity for increased use of the instrument in research and schools. The utilization of fidelity assessment is essential in demonstrating the effectiveness of an intervention in research and practice (Moncher & Prinz, 1991; VanDerHeyden et al., 2007). Thus, the major implications of the results of this study span research and practice and are described below.

## **Implications of Results in Research**

The implications of the results of this study within the realm of research are three-fold. The first implication of these results within research is the need to isolate the effects of the CICO intervention. According to Moncher and Prinz (1991), a psychometrically sound fidelity assessment enables a researcher to identify the specific effects of an independent variable on the dependent variable. Thus, a psychometrically sound fidelity assessment of CICO provides an instrument that enables a researcher to measure within the study if the intervention was implemented as it was designed. This measurement provides evidence that the assumption of the presence of the independent variable is true and thus its impact on the dependent variable can be measured. This in turn increases the use of the CICO-FIM in research.

The second implication of these results within research is consistency. Previous research studies into the effectiveness of CICO utilized fidelity assessment to ensure that the intervention is implemented as it was designed (Campbell & Anderson, 2011; Hawken, 2006; Hawken & Horner, 2003; Filter et al., 2007; March & Horner, 2002; Todd, Campbell, Meyer, & Horner, 2008). However, this research has not utilized the same instrument of fidelity across studies (Maggin et al., 2015). Without a standard measure of fidelity, an inconsistency is added when comparing outcomes across studies. Thus, consistently using the CICO-FIM as means to assess fidelity, researchers and consumers of research have a standardized measurement and definition of implementation fidelity in CICO. The wide-scale utilization of the Schoolwide Evaluation Tool (SET; Horner et al., 2004) to measure universal PBIS in research provides precedence for the use of a psychometrically sound instrument to measure the

fidelity of implementation (Cohen et al., 2007; Eber et al., 2001; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008; Vincent et al., 2010). Just this year, two studies have been published on the efficacy of CICO, which signifies that research on the intervention is an area of continued study and that there is a place within this research for the CICO-FIM to provide a defined, standardized approach to CICO fidelity assessment (Maggin et al, 2015; Miller et al., 2015).

Third, a psychometrically promising CICO-FIM supports further investigation into CICO. Specifically, more investigation is needed to link the fidelity of CICO to increased behavioral outcomes for students. More about directions for future research is presented later in this chapter.

### **Implications of Results in Practice**

The implications of the results of this study within the realm of practice are three-fold as well. First, as a result of this study, the use of this instrument can be increased in schools across the country implementing CICO. The number of schools this affects is yet unknown; however, over 23,000 schools in the United States, Australia, Canada, and Norway utilize the web-based PBIS Assessment application to measure fidelity of school-wide PBIS, demonstrating the large-scale possibilities of a psychometrically sound instrument measuring CICO (Educational and Community Supports, 2012).

Second, the high item mean scores on item relevance and clarity from the content expert reviews suggest that in conjunction with the instrument total score, individual items can support school, district, state, and national performance assessment and program evaluation. Thus, current and future schools can use the results of this instrument to guide implementation and to identify areas in need of improvement (Fixsen

et al., 2013). Districts can utilize the results of this instrument to identify coaching needs to support CICO implementation and to tailor professional development according to results obtained. Furthermore, state and national technical assistance centers can use these results to provide professional development and technical assistance that is focused on data-driven support. This data could also be used to tailor specific professional development, technical assistance, and coaching to follow up on initial training to schools and districts to increase the level of implementation fidelity. Additionally, the same organizations that provide professional development, technical assistance, and coaching on CICO can utilize the CICO-FIM instrument total score to provide a means to measure behavior change in the school professionals that are trained in the intervention, and thus become a data source for program evaluation.

The third implication of results of this study in practice is that the CICO-FIM builds on the evidence-base of the use of self-assessment in measuring performance assessment. The self-assessment has limited evidence-base in the literature. The instrument uses self-assessment by the CICO coordinator or team to assess fidelity. Self-assessment is an appropriate means of data collection on the CICO-FIM for a number of reasons. First, self-assessment is used frequently as a means of measuring fidelity because it is less expensive and less time-consuming than outside observation (O'Donnell, 2008; Bretenstein et al., 2010). Self-assessment is also a widely used means of self-appraisal for improvements, and schools often use self-assessment with students as a tool to achieve realistic goals and to direct areas in need of improvement (Bullock, 2010). School professionals, teams, coaches, and coordinators use the practice of self-assessment across the tiers of PBIS implementation (Cohen et al., 2007; Vincent et al.,

2010). Brand, Felner, Seitsinger, Burns, and Bolton (2008) found that with respect to school climate, teachers' ratings are consistent with student reports of school climate, which provides evidence that teachers can use self-assessment to appropriately measure school climate. This study adds to the evidence-base of the use of self-assessment in performance assessment.

### **Limitations**

This study was the initial psychometric investigation into the CICO-FIM. As is expected of an initial assessment, there are limitations to the results of this study. The limitations here are associated with the instrument, sampling, and dimensions of fidelity assessment.

### **Instrument**

The study investigated the 12-item self-assessment of the CICO-FIM. The instrument has limitations in three primary areas: 1) development, 2) response bias, and 3) lack of comparison measure.

The first limitation of the instrument is a lack of information on its development. The CICO-FIM was disseminated in the text written by Crone, Hawken, and Horner (2011) without any information on how it was designed and written. Furthermore, there is no published study that describes the creation, initial validation, or dimensionality of the instrument. The instrument was assumed to be unidimensional because it uses a composite score, but does not provide psychometric evidence of this factor structure. This limitation was known prior to the study, and was considered acceptable because the instrument was created and written by the authors of the intervention itself. Additionally,

one of the authors has experience creating fidelity assessments that are validated and widely utilized (Vincent et al., 2010).

The second limitation is the response bias present in utilizing an instrument that relies upon self-assessment. Specifically, the use of self-assessment data introduces a social desirability response bias (Furr & Bacharach, 2014). Social desirability describes the response bias in which participants are more likely to report or assess themselves in a way that is more socially appealing than provide a true score (Furr & Bacharach, 2014). An instrument that utilizes self-assessment is open to this response bias. The sources of social desirability response bias are test context, personality of participants, and test content. The data utilized for this study was archival; therefore, participants completed the instrument with the expectation of anonymity, thus diminishing the threat of response bias from the testing context (Furr & Bacharach, 2014). The threat of social desirability response bias was also reduced from the personality of participants. The CICO coordinator or the CICO implementation team determined the item responses on the CICO-FIM; therefore, the use of team consensus diminishes this threat. The test content as a source of social desirability was reduced due to the nature of the CICO-FIM. This instrument was used by school professionals to measure their implementation of CICO and, based on those results, the data were used to identify areas to improve the implementation of CICO. This instrument was used voluntarily, and is not associated with any high-stakes initiatives, which also lessens the threat of social desirability response bias with respect to test content (Furr & Bacharach, 2014). Finally, the instrument also estimated good IRR when school-level observations were correlated with

those of an outside expert. This correlation provides supporting evidence that the schools in the sample did not exhibit response bias.

The third limitation with respect to the CICO-FIM is that it is an instrument in isolation. There are no other instrument or observation tools that have been created to measure the implementation of CICO. Due to the lack of existing psychometrically sound or consistent measurement of CICO to assess the CICO-FIM, validity assessment of the instrument was hindered. The initial results of this study have begun to establish the validity of the CICO-FIM without an existing comparison instrument.

### **Sampling**

Another limitation of this study is associated with sampling, specifically the type of sampling procedures and the size of the sample. This study utilized a convenience sample of archival data from schools across a western state. A convenience sample is not representative of the population. Instead, a probability sample is representative of the population, which is necessary for generalizing results. Creswell (2013) states that it is often difficult to get a probability sample in education. Convenience samples, however, are standard in education when using naturally formed groups. This study used a convenience sample of the naturally formed group of CICO implementers trained by the statewide PBIS initiative that collected the CICO-FIM. Therefore, the results from this study can only be applied to this convenience sample. Additionally, Table 1 demonstrates that the race/ethnicity of the students within the schools in this sample are similar to those across the entire western state studies.

Furthermore, future research needs to include larger samples to increase the power of the investigation and utilize more sensitive analysis. The largest sample of the

study utilized a sample of 120 complete instruments from 79 schools—a sample that was appropriate for the size of the instrument, research questions the sample addressed, and for this initial investigation. However, Tables 2 and 6 demonstrates that the current sample had a large number of elementary schools, which denotes a need for a sample with more middle and high school representation. Additionally, the anticipated sample size for IRR was 30, due to the fact that it meets the minimum sample size for a correlation (Xu & Lorber, 2014). However, due to issues with instrument completion in the field, the actual sample size was 27. Since this sample size is close to 30, the IRR was still examined. The interpretations of these findings were done with caution due to the small sample size. The size of this sample supports replication with a larger sample to ensure the interpretations were appropriate. In order to address these sampling limitations, a large-scale, random sample examination of the CICO-FIM is needed to further substantiate the results from this initial investigation.

### **Dimensions of Fidelity Assessment**

Finally, the measurement of the construct of fidelity is a limitation of this study. Fidelity assessment in education is a burgeoning science that lacks consensus regarding the dimensions integral to its measurement (Moncher & Prinz, 1991). Therefore, more experimental research is needed to examine the variables that could influence fidelity (Moncher & Prinz, 1991; O'Donnell, 2008; VanDerHeyden et al., 2008). In the absence of one clear model of the dimensions of fidelity assessment in schools, empirical research is needed to evaluate the relationship among the various conceptual models of fidelity and their associated variables (Gerstner & Finney, 2013; Harn et al., 2013; Moncher & Prinz, 1991; O'Donnell, 2008). While the debate regarding the measurement of fidelity

continues, it is still essential that the fidelity of evidence-based practices continue to be measured (Fixsen et al., 2013). The procedures utilized in this archival data set address the current approach to the definition and dimensions of fidelity assessment in an educational setting in order to address this limitation (Carroll et al., 2007; Moncher & Prinz, 1991; O'Donnell, 2008).

### **Future Research**

The initial investigation of the CICO-FIM in this study has provided a solid starting point for research that provides a standard protocol for measuring the fidelity of CICO intervention implementation. Future research is necessary to build on this momentum. The four major areas to focus future research on this instrument suggested by the researcher are replication, sampling, more sensitive analysis, and identifying a critical threshold, which are discussed below.

First, future research is needed to replicate this study. The promising results of this initial study into the psychometric properties of the CICO-FIM are a good start but they are reflective of one convenience sample. In order to add to the research base on this instrument, this study needs to be replicated with a different sample.

Moreover, a second area in need of future research is the investigation of this instrument using a sample that is large-scale and ascertained via probability sampling procedures. The present study relied on a convenience sample that met the minimum requirements for sample size for each research question. Since this study was the initial investigation of the CICO-FIM, a smaller and non-parametric sampling was appropriate. However, subsequent investigations into this instrument would need to utilize a larger sample collected with probability sampling techniques.

The third area to focus future research on the CICO-FIM is utilizing more sensitive psychometric analysis. More investigation into the CICO-FIM would include a confirmatory factor analysis that examines the unidimensional component structure of the instrument. It also needs to examine item 2 under these new specifications to ensure that it continues yielding appropriate test statistics, which also sheds light on the need for more research on an item level into this instrument. Additionally, more research is needed using Rasch analysis to ensure the fundamental measurement of the instrument (Bond & Fox, 2007).

The fourth area of future research on the CICO-FIM instrument is establishing a critical threshold of fidelity. The CICO intervention has been demonstrated to increase behavioral outcomes for students who participate in the intervention (Hawken & Horner, 2003; Hawken et al., 2007; Lane et al., 2012; Miller et al., 2015; Todd et al., 2008). It is essential that fidelity assessment is associated with increased outcomes for students (Gerstner & Finney, 2013; Harn et al., 2013; Moncher & Prinz, 1991). The next area of research for this instrument is in determining the degree of fidelity achieved that is associated with increased student outcomes. This research will add to evidence of the construct validity of the CICO-FIM (Furr & Bacharach, 2014), and will do so by demonstrating that the instrument measures the implementation of CICO, but also that a precise degree of fidelity of implementation achieves results in the desired consequence of CICO (Furr & Bacharach, 2014).

### **Conclusion**

Check-In, Check-Out (CICO) is a targeted group behavioral intervention that is utilized within a response to intervention (RTI) framework in schools to prevent severe

problem behavior in students that are beginning to exhibit externalizing and/or internalizing behavioral needs (Crone et al., 2011; Hawken & Horner, 2003). CICO is one intervention utilized in a continuum of support to decrease the use of suspensions and expulsions of students demonstrating problem behavior (Campbell & Anderson, 2011; Hawken, 2006; Hawken & Horner, 2003). As evidence for the effectiveness of CICO grows, so does the number of schools implementing the intervention (Filter et al., 2007; March & Horner, 2002; Todd et al., 2008). This boom in implementation puts further emphasis on the need for a reliable and valid instrument to measure CICO interventions (Moncher & Prinz, 1991; VanDerHeyden, 2007).

The intention of this dissertation was to provide initial insight into the psychometric properties of the CICO-FIM (Crone et al., 2010), an instrument developed to measure the implementation of CICO. Prior to this study, the reliability and validity of the CICO-FIM had not been investigated, nor was there another psychometrically sound instrument utilized to measure CICO. Used in concert, the analysis of content and construct validity, factor analysis, internal consistency, and IRR provided evidence that the CICO-FIM is a psychometrically promising instrument.

Measuring the fidelity of CICO is essential to supporting students that are at-risk for severe problem behaviors that could get them suspended or expelled. Evidenced-based interventions that provide efficient and effective tools for school professionals and students to use to prevent future behavior problems are integral in decreasing the use of exclusionary discipline practices. This study has important implications for education today: the promising psychometric results support the use and further examination of the CICO-FIM in measuring the fidelity of CICO. As such, this study provides an instrument

to measure an important intervention in decreasing problem behaviors and the use of exclusionary discipline.

Today, the assessment of fidelity is an essential practice in the research and practical implementation of interventions in education (O'Donnell, 2008). Fidelity assessment is integral in research on interventions, as it is imperative that investigators isolate and demonstrate the impact of an independent variable on dependent variables (Moncher & Prinz, 1991). Measuring the fidelity of the implementation of an independent variable allows investigators to verify that the independent variable is being implemented as it was designed and, therefore, that changes in the dependent variable can be seen as a result of the independent (Moncher & Prinz, 1991). Fidelity assessment is also fundamental to the practical implementation of strategies, practices, and interventions in schools (Gerstner & Finney, 2013; Harn et al., 2013). To assess the effectiveness of an intervention addressing the needs of a student, school professionals must ensure that the intervention is implemented as it was designed (Gerstner & Finney, 2013; VanDerHeyden et al., 2007). Without measuring the fidelity of an intervention, school professionals cannot be sure that students are receiving the evidence-based support necessary to meet their academic or behavioral needs (VanDerHeyden et al., 2007). Practices and interventions implemented throughout the PBIS framework measure fidelity to ensure implementation to the defined model—including universal practices, classroom management strategies, and wraparound planning (Cohen et al., 2007; Eber et al., 2001; Horner, 2000; Simonsen, Fairbanks, Briesch, Myers, & Sugai, 2008; Vincent et al., 2010). This study adds to the growing research base and practice of utilizing fidelity assessment across the framework and practices of PBIS.

Guaranteeing that practices and interventions are implemented as they are designed is the promise that school professionals make to students and families when assessing a student's response to a practice, or when assessing a program in a RTI framework (VanDerHeyden et al., 2007). Measuring fidelity must become an essential practice in schools to ensure the implementation, sustainability, and replication of evidence-based practices (Fixsen et al., 2007). The school reform efforts of multi-tiered system of supports (MTSS), PBIS, and RTI focus on increasing academic and behavioral outcomes for students at the intersection of the implementation of evidence-based systems, practices, interventions, and data based problem-solving procedures (Sugai & Horner, 2003; Tilly, 2008). Fidelity assessment is the cornerstone of these efforts because it provides data to demonstrate that school professionals are making good on the promise of implementing a practice or intervention as it was designed.

In conclusion, educators today find themselves at a crossroads, of balancing the pressure of high stakes legislation that centers on increasing academic test scores, while trying to increase the behavioral, social, and emotional skills of students so that they are able to perform academically. A growing body of data from the federal government, state governments, universities, and non-profits demonstrates that students of color are being left behind in this tight rope act. The disengagement of students with school through the use of suspension and expulsion is a problem that our society cannot continue to ignore.

It is imperative that policy makers act to support educators in these efforts by ensuring that each and every student receives a free, appropriate education that will provide them with the tools to meet the demands of the 21st century work environment. This work environment does not just center on academic knowledge; rather, it requires

this knowledge to work in concert with problem solving, social skills, flexibility, initiative, and cross-cultural skills (just to name a few). As the discussion to reauthorize both the Elementary and Secondary Education Act (ESEA, 2001) and the Individuals with Disabilities Education Act (IDEA, 2004) continues, these policy makers must explicitly address the implementation with fidelity of practices, interventions, and programs that support student behavior, teach social and emotional skills, and deter the use of suspension and expulsion in order to keep good on the promise of educating each and every student. Fidelity assessment is key to this policy as it provides both a road map and a system for accountability that demonstrates educators are doing all that they can to support each and every student, especially those students that have been thus far left behind, disengaged by the system, or just asked to leave.

## REFERENCES

- American Academy of Pediatrics. (2013). Policy statement: Out-of-school suspension and expulsion. *Pediatrics*, *131*(3), e1000-e1007. doi:10.1542/peds.2012-3932
- American Psychological Association. (2006). Evidence-based practice in psychology. *American Psychologist*, *61*(4), 271-285. doi: 10.1037/0003-066X.61.4.271
- American Psychological Association Zero Tolerance Task Force. (2008). Are zero tolerance policies effective in the schools? An evidentiary review and recommendations. *American Psychologist*, *63*(9), 852-862. doi: 10.1037/0003-066X.63.9.852
- Anderson, A. R., Christenson, S. L., Sinclair, M. F., & Lehr, C. A. (2004). Check & connect: The importance of relationships for promoting engagement with school. *Journal of School Psychology*, *42*(2), 95-113.
- Anderson, C. M., & Borgmeier, C. (2010). Tier II interventions within the framework of school-wide positive behavior support: Essential features for design, implementation, and maintenance. *Behavior Analysis in Practice*, *3*(1), 33-45.
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1987). Some still-current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis*, *20*(4), 313-327. doi: 10.1901/jaba.1987.20-313
- Batsche, G., Elliot, J., Garden, J., Grimes, J., Kovalesski, J. F., Prasse, D., . . . Tilly, W. D. (2005). *Response to intervention: Policy considerations and implementation*. Alexandria, VA: National Association of State Directors of Special Education.

- Bellg, A., Borrelli, B., Resnick, B., Hecht, J., Minicucci, D. S., Ory, M., . . . Czajkowski, S. (2004). Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the Behavior Change Consortium. *Health Psychology, 23*, 443-451. doi: 10.1037/0278-6133.23.5.44
- Biggs, B. K., Vernberg, E. M., Twemlow, S. W., Fonagy, P., & Dill, E. J. (2008). Teacher adherence and its relation to teacher attitudes and student outcomes in an elementary school-based violence prevention program. *School Psychology Review, 37*(4), 533–549.
- Bobko, P. (2001). *Correlation and regression: Applications for industrial organizational psychology and management* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). New York, NY: Routledge.
- Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A. J., Czajkowski, S., Breger, R., . . . Orwig, D. (2005). A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of Consulting and Clinical Psychology, 73*(5), 852-860. doi: 10.1037/0022-006X.73.5.852
- Brand, S., Felner, R. D., Seitsinger, A., Burns, A., & Bolton, N. (2008). A large scale study of the assessment of the social environment of middle and secondary schools: The validity and utility of teachers' ratings of school climate, cultural pluralism, and safety problems for understanding school effects and school improvement. *Journal of School Psychology, 46*, 507-535. doi: 10.1016/j.jsp.2007.12.001

- Bretenstein, S. M., Gross, D., Garvey, C., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*(2), 164-173. doi: 10.1002/nur.20373
- Bronfenbrenner, U. (1979). *The ecology of human development*. Cambridge, MA: Harvard University Press.
- Bullock, D. (2010). Learner self-assessment: An investigation into teachers' beliefs. *ELT Journal, 65*(2), 114-125. doi: 10.1093/elt/ccq041
- Campbell, A., & Anderson, C. M. (2011). Check-in/check-out: A systematic evaluation and component analysis. *Journal of Applied Behavior Analysis, 44*, 315–326. doi: 10.1901/jaba.2011.44-315
- Carr, E. G., Dunlap, G., Horner, R. H., Koegel, R. L., Turnbull, A. P., Sailor, W., ... Fox, L. (2002). Positive behavior support: Evolution of an applied science. *Journal of Positive Behavior Interventions, 4*(1), 4-16. doi 10.1177/109830070200400102
- Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*(40). doi: 10.1186/1748-5908-2-40
- Chafouleas, S. M., McDougal, J. L., Riley-Tillman, T. C., Panahon, C. J., & Hilt, A. M. (2005). What do daily behavior report cards (DBRCs) measure? An initial comparison of DBRCs with direct observation for off-task behavior. *Psychology in the Schools, 42*(6), 669-676. doi: 10.1002/pits.20102
- Chafouleas, S. M., Riley-Tillman, T. C., & Sassu, K. A. (2006). Acceptability and reported use of daily behavior report cards among teachers. *Journal of Positive Behavior Interventions, 8*, 174-182.

- Chafouleas, S. M., Riley-Tillman, T. C., Sassu, K. A., LaFrance, M. J., & Patwa, S. S. (2007). Daily behavior report cards: An investigation of the consistency of on task data across raters and methods. *Journal of Positive Behavioral Interventions, 9*(1), 30-37. doi: 10.1177/10983007070090010401
- Cheney, D., Flower, A., & Templeton, T. (2008). Applying response to intervention metrics in the social domain for students at risk of developing emotional or behavioral disorders. *Journal of Special Education, 42*(2), 108-126. doi: 10.1177/0022466907313349
- Cohen, R., Kincaid, D., & Childs, K. E. (2007). Measuring school-wide positive behavior support implementation: Development and validation of the benchmarks of quality. *Journal of Positive Behavior Interventions, 9*(4), 203–213. doi:10.1177/10983007070090040301
- Colorado Department of Education (2014). *Colorado definition of public school levels*. Retrieved from <http://www.cde.state.co.us/datapipeline/codeofdefinitionspublicschoollevels>
- Colorado Department of Education (2015). Statewide preschool through 12th grade student enrollment continues to grow [Press release]. Retrieved from <http://www.cde.state.co.us/communications/20150115pupilenrollment>
- Colvin, G., Kame 'enui, E. J., & Sugai, G. (1993). Reconceptualizing behavior management and school wide discipline and general education. *Education & Treatment of Children, 16*(4), 361-381.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed method approaches* (4th ed.). Thousand Oaks, CA: SAGE Publications.

- Crone, D. A., Hawken, L. S., & Horner, R. H. (2010). *The behavior education program: A check-in, check-out intervention for students at risk* (2nd ed.). New York, NY: The Guilford Press.
- DeVellis, R. F. (2003), *Scale development: Theory and application* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Duncan, A. [used.gov]. (2014, January 8). *U.S. Departments of Education and Justice release school discipline guidance package* [Video file]. Retrieved from <https://www.youtube.com/watch?v=zOloBfI9SGo#t=23>
- Dunst, C. J., & Trivette, C. M. (2009). Let's be PALS: An evidence-based approach to professional development. *Infants & Young Children, 22*(3), 164-176. Retrieved from <http://dx.doi.org/10.1155/2009/576840>
- Durlack, J. A., & DuPre, E. P. (2008). Implementation matters: A review on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*, 327-350. doi:10.1007/s10464-008-9165-0
- Eber, L., Hyde, K., & Suter, J. C. (2011). Integrating wraparound into a school-wide system of positive behavior supports. *Journal of Child and Family Studies, 20*(6), 782-790. doi: 10.1007/s10826-010-9424-1
- Educational and Community Supports (2012). Who uses our Apps? Retrieved from [https:// www.pbisapps.org/About-Us/Who-Uses-Our-Apps.aspx](https://www.pbisapps.org/About-Us/Who-Uses-Our-Apps.aspx)
- Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Miner, P. M., & Booth, E. A. (2011). *Breaking schools' rules: A statewide study of how school discipline*

*relates to students' success and juvenile justice involvement.* New York, NY:  
Council of State Governments Justice Center.

- Fabriger, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299. doi: 10.1037/1082-989X.4.3.272
- Fairbanks, S., Sugai, G., Guardino, D., & Lathrop, M. (2007). Response to intervention: Examining classroom behavior support in second grade. *Exceptional Children, 73*, 288-310. doi: <http://dx.doi.org/10.1177/001440290707300302>
- Filter, K. J., McKenna, M. K., Benedict, E. A., Horner, R. H., Todd, A., & Watson, J. (2007). Check in/check out: A post-hoc evaluation of an efficient, secondary level targeted intervention for reducing problem behaviors in schools. *Education and Treatment of Children, 30*, 69-84. doi: <http://dx.doi.org/10.1353/etc.2007.0000>
- Fixsen, D., Blase, K., Metz, A., & Van Dyke, M. (2013). Statewide implementation of evidence-based programs. *Exceptional Children, 79*(2) 213-230. doi: 10.1177/001440291307900206
- Flay, B. R., Biglan, A., Boruch, R. F., Gonzalez Castro, F., Gottfredson, D., Kellam, S., & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*(3), 151-175. doi: 10.1007/s11121-005-5553
- Forgatch, M. S., Patterson, G. R., & DeGarmo, D. S. (2005). Evaluating fidelity: Predictive validity for a measure of component adherence to the Oregon model of parent management training. *Behavior Therapy, 36*(1), 3-13. doi: 10.1016/S00057894(05)80049-8

- Fowler, F. J. (2009). *Survey research methods* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: An introduction* (2nd ed.). Thousand Oaks, CA: Sage Publications.
- Gansle, K. A., & McMahon, C. M. (1997). Component integrity of teacher intervention management behavior using a student self-monitoring treatment: An experimental analysis. *Journal of Behavioral Education, 7*(4), 405–419. doi: 10.1023/A:1022851117439
- Gerstner, J. J., & Finney, S. J. (2013). Measuring the implementation fidelity of student affairs programs: A critical component of outcomes assessment cycle. *Research & Practice in Assessment, 8*, 15-28.
- Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessments. *Journal of School Psychology, 45*, 117-135. doi: 10.1016/j.jsp.2006.05.005
- Golly, A., Sprague, J., Walker, H., Beard, K., & Gorham, G. (2000). The First Step to Success program: An analysis of outcomes with identical twins across multiple baselines. *Behavioral Disorders, 25*, 170–182.
- Gresham, F. M. (1989). Assessment of treatment integrity in school consultation and prereferral intervention. *School Psychology Review, 18*(1), 37-50.
- Gresham, F. M., Gansle, K. A., & Noell, G. H. (1993). Treatment integrity in applied behavior analysis with children. *Journal of Applied Behavior Analysis, 26*(2), 257-263. doi: 10.1901/jaba.1993.26-257

- Gresham, F. M., MacMillan, D. L., Beebe-Frankenberger, M., & Bocian, K. M. (2000). Treatment integrity in learning disabilities intervention research: Do we really know how treatments are implemented? *Learning Disabilities Research & Practice, 15*(4), 198-205. doi: 10.1207/SLDRP1504\_4
- Griffin, D., & Gonzalez, R. (1995, Nov). Correlational analysis of dyad-level data in the exchangeable case. *Psychological Bulletin, 118*(3), 430-439.  
<http://dx.doi.org/10.1037/00332909.118.3.430>
- Hagermoser Sanetti, L. M., & Kratochwill, T. R. (2009). Treatment integrity assessment in the schools: An evaluation of the Treatment Integrity Planning Protocol. *School Psychology Quarterly, 24*(1), Mar 2009, 24-35. doi:  
<http://dx.doi.org/10.1037/a0015431>
- Harn, B., Parisi, D., & Stoolmiller, M. (2013). Balancing fidelity with flexibility and fit: What do we really know about fidelity of implementation in schools? *Exceptional Children, 79*(2), 181-193. doi: 10.1177/001440291307900204
- Hawken, L. S. (2002). *Evaluation of a targeted group intervention within a school-wide system of behavior support* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses database. (UMI No. 3055691)
- Hawken, L. S. (2006). School psychologists as leaders in the implementation of a targeted intervention: The behavior education program. *School Psychology Quarterly, 21*, 91–111. doi: <http://dx.doi.org/10.1521/scpq.2006.21.1.91>
- Hawken, L. S., & Horner, R. H. (2003). Evaluation of a targeted intervention within a school-wide system of behavior support. *Journal of Behavioral Education, 12*, 225–240. doi: 10.1023/A:1025512411930

- Hawken, L. S., & Johnston, S. S. (2008). Preventing severe problem behavior in young children: The behavior education program. *Journal of Early and Intensive Behavior Intervention, 4*(3), 2007, 599-613. doi:  
<http://dx.doi.org/10.1037/h0100394>
- Hawken, L. S., MacLeod, S. K., & Rawlings, L. (2007). Effects of the behavior education program (BEP) on office discipline referrals of elementary school students. *Journal of Positive Behavior Interventions, 9*, 94-101. doi:  
10.1177/10983007070090020601
- Heaviside, S., Rowand, C., Williams, C., & Farris, E. (1998). *Violence and discipline problems in U.S. public schools: 1996-97 (NCES 98-030)*. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.
- Horner, R. H. (2000). Positive behavior supports. *Focus on Autism and Other Developmental Disabilities, 15*(2), 97-105. doi: 10.1177/108835760001500205
- Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2004). The school-wide evaluation tool (SET): A research instrument for assessing school-wide positive behavior supports. *Journal of Positive Behavior Interventions, 6*(1), 3–12. doi: 10.1177/10983007040060010201
- Horner, R. H., Sugai, G., & Anderson, C. M. (2010). Examining the evidence base for school-wide positive behavior support. *Focus on Exceptional Children, 42*(8), 1-14.
- IBM Corp. (2013). *IBM SPSS Statistics for Windows, Version 22.0*. Armonk, NY: IBM Corp.

- Individuals with Disabilities Education Improvement Act of 2004, 20 U.S.C.1400 et seq. (2004). Retrieved from <http://idea.ed.gov/download/statute.html>.
- Institute of Education Sciences. (2011, February 28). *Education research grants CFDA 84.305A* Retrieved from [http://ies.ed.gov/funding/pdf/2012\\_84305A.pdf](http://ies.ed.gov/funding/pdf/2012_84305A.pdf)
- Kratochwill, T. R. (2007). Preparing school psychologists for evidence-based school practice: Lessons learned and challenges ahead. *American Psychologist, 62*(8), 829-843. doi: <http://dx.doi.org/10.1037/0003-066X.62.8.824a>
- Lane, K. L., Capizzi, A. M., Fisher, M. H., & Ennis, R. P. (2012). Secondary prevention efforts at the middle school level: An application of the behavior education Program. *Education and Treatment of Children, 35*(1), 51-90. doi: 10.1353/etc.2012.0002
- Leech, N. L., Barrett, K. C., & Morgan, G. A. (2015). *IBM SPSS for intermediate statistics: Use and interpretation* (5th ed.). New York, NY: Routledge.
- Long, N. & Edwards, M. (1994). The use of a daily report card to address children's school behavior problems. *Contemporary Education, 65*(3), 152-155.
- Maggin, D. M., Zurheide, J., Pickett, K. C., & Baillie, S. J. (2015). A systematic evidence review of the check-in/check-out program for reducing student challenging behaviors. *Journal of Positive Behavior Interventions*. (Advance online publication.) doi: 10.1177/1098300715573630
- March, R. E., & Horner, R. H. (2002). Feasibility and contributions of functional behavioral assessment in schools. *Journal of Emotional and Behavioral Disorders, 10*(3), 158–170. doi: 10.1177/10634266020100030401

- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99.
- McKenna, S. A., Rosenfield, S., & Gravois, T. A. (2009). Measuring the behavioral indicators of instructional consultation: A preliminary validity study. *School Psychology Review, 38*(4), 496-509.
- Miller, L. M., Dufrene, B. A., Sterling, H. E., Olmi, D. J., & Bachmayer, E. (2015). The effects of check-in/check-out on problem behavior and academic engagement in elementary school students. *Journal of Positive Behavior Interventions, 17*(1), 28-38. doi: 10.1177/1098300713517141
- Moncher, F. J., & Prinz, F. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11*, 247-266.
- Mundfrom, D. J., Shaw, D. G., & Lu Ke, T. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing, 5*(2), 159-158. doi: 10.1207/s15327574ijt0502\_4
- National Association of School Psychologists (NASP). (2009). *Appropriate behavioral, social, and emotional supports to meet the needs of all students (Position Statement)*. Bethesda, MD: National Association of School Psychologists.
- National Institutes of Health. (2003). *Pathways linking education to health*. Bethesda, MD: National Institutes of Health. Retrieved from [www.federalgrants.com/PATHWAYS LINKING EDUCATION-TO-HEALTH 4663.html](http://www.federalgrants.com/PATHWAYS LINKING EDUCATION-TO-HEALTH 4663.html)
- Newton, J. S., Horner, R. H., Algozine, B., Todd, A. W., & Algozine, K. (2012). A randomized wait-list controlled analysis of the implementation integrity of team

- initiated problem solving processes. *Journal of School Psychology, 50*(4), 421-441. doi: 10.1016/j.jsp.2012.04.002
- No Child Left Behind (NCLB) Act of 2001, Pub. L. No. 107-110, §115, Stat. 1425 (2002). Retrieved from <http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf>.
- Noell, G. H., & Gansle, K. A. (2009). Moving from good ideas in educational systems change to sustainable program implementation: Coming to terms with some of the realities. *Psychology in the Schools, 46*(1), 78-88. doi: 10.1002/pits.20355
- Noell, G. H., Witt, J. C., Slider, N. J., Connell, J. E., Gatti, S. L., Williams, K. L., ... & Resetar, J. L. (2005). Treatment implementation following behavioral consultation in schools: A comparison of three follow-up strategies. *School Psychology Review, 34*(1), 87-106.
- Office of Special Education Programs. (2014, December 22). *Request for application for State Personnel Development Grant Program (SPDG) CFDA 84.323A*. Retrieved from <http://www.signetwork.org>
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers, 32*(3), 396-402.
- O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K-12 curriculum intervention research. *Review of Educational Research, 78*(1), 33-84. doi: 10.3102/0034654307313793

- Patil, V. H., Singh, S. N., Mishra, S., & Donovan, D. T. (2007). Parallel analysis engine to aid determining number of factors to retain. (Computer software). Retrieved from <http://smishra.faculty.ku.edu/parallelengine.htm>
- Polit, D. F. & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health, 29*, 489-497. doi: 10.1002/nur.20147
- Rhymer, R. N., Evans-Hampton, T. N., McCurdy, M., & Watson, T. S. (2002). Effects of varying levels of treatment integrity on toddler aggressive behavior. *Special Services in Schools, 18*(1-2), 75-81. doi: 10.1300/J008v18n01\_05
- Riley-Tillman, T. C., Chafouleas, S. M., Briesch, A. M., & Eckert, T. L. (2008). Daily behavior report cards and systematic direct observation: An investigation of the acceptability, reported training and use, and decision reliability among school psychologists. *Journal of Behavioral Education, 17*(4), 313-327. doi: 10.1007/s10864-008-9070-5
- Russell, D. W. (2002). In search of underlying dimensions: The use (and abuse) of factor analysis in *Personality and Social Psychology Bulletin. Personality and Social Psychology Bulletin, 28*(12), 1629-1646. doi: 10.1177/014616702237645
- Severson, H. H., Walker, H. M., Hope-Doolittle, J., Kratochwill, T. R., & Gresham, F. M. (2007). Proactive, early screening to detect behaviorally at-risk students: Issues, approaches, emerging innovations, and professional practices. *Journal of School Psychology, 45*, 193-223. doi:10.1016/j.jsp.2006.11.003

- Simonsen, B., Fairbanks, S., Briesch, A., Myers, D., & Sugai, G. (2008). Evidence-based practices in classroom management: Considerations for research to practice. *Education and Treatment of Children, 31*(3), 351-380. doi: 10.1353/etc.0.0007
- Sheridan, S. M., Swanger-Gagné, M., Welch, G. W., Kwon, K., & Garbacz, S. A. (2009). Fidelity measurement in consultation: Psychometric issues and preliminary examination. *School Psychology Review, 38*, 476-495.
- Shore, R., & Shore, B. (2009). *Kids count indicator brief: Reducing the high school dropout rate*. Baltimore, MD: The Anna E. Casey Foundation. Retrieved from [http://www.aecf.org/m/resourcedoc/AECF ReducingTheHighSchoolDropoutRate-2009Full.pdf](http://www.aecf.org/m/resourcedoc/AECF%20ReducingTheHighSchoolDropoutRate-2009Full.pdf)
- Simonsen, B., Myers, D., & Briere, D. E. (2010). Comparing a behavioral check in/check-out (CICO) intervention to standard practice in an urban middle school setting using an experimental group design. *Journal of Positive Behavior Interventions, 13*(1), 31-48. doi: 10.1177/1098300709359026
- Sinclair, M. F., Christenson, S. L., Evelo, D. L., & Hurley, C. M. (1998). Dropout prevention for youth with disabilities: Efficacy of a sustained school engagement procedure. *Exceptional Children, 65*(1), 7-22.
- Skiba, R., Chung, C. G., Trachok, M., Baker, T. L., Sheya A., & Hughs R. L. (2014). Parsing disciplinary disproportionality: Contributions of infraction, student, and school characteristics to out-of-school suspension and expulsion. *American Educational Research Journal, 51*(4), 640-670. doi: 10.3102/0002831214541670

- Skiba, R. J., Horner, R. H., Chung, C. G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review, 40*(1), 85-107.
- Skiba, R. J., Michael, R. S., Nardo, A. C., & Peterson, R. L. (2002). The color of discipline: Sources of racial and gender disproportionality in school punishment. *The Urban Review, 34*, 317-342. doi: 10.1023/A:1021320817372
- Skiba, R. J., & Sprague, J. (2008). Safety without suspensions. *Educational Leadership, 66*(1), 38-43.
- Sprague, J., Cook, C. R., Browning Wright, D., & Sadler, C. (2008). *RTI and behavior: A guide to integrating behavioral and academic supports*. Horsham, PA: LRP Publications.
- Sugai, G., & Horner, R. H. (2002). The evolution of discipline practices: School-wide positive behavior supports. *Child and Family Behavior Therapy, 24*, 23–50. doi: 10.1300/J019v24n01\_03
- Sugai, G., & Horner, R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School Psychology Review, 35*, 245-259.
- Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., ...Ruef, M. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions, 2*(3), 131–143. doi:10.1177/109830070000200302
- Sugai, G., Horner, R. H., & Todd, A. W. (2000). *Effective behavior support: Self-assessment survey*. Eugene, OR: University of Oregon Press.

- Sugai, G., Lewis-Palmer, T., Todd, A. W., & Horner, R. H. (2001). *School-wide evaluation tool*. Eugene, OR: University of Oregon Press.
- Swoszowski, N. C., Jolivette, K., Fredrick, L. D., & Heflin, L. J. (2012). Check in/check out: Effects on students with emotional and behavioral disorders with attention or escape-maintained behavior in a residential facility. *Exceptionality, 20*(3), 163-178. doi: 10.1080/09362835.2012.694613
- Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education.
- Tilly, W. D., III (2008). The evolution of school psychology to a science-based practice: Problem solving and the three-tiered model. In J. Grimes, & A. Thomas (Eds.), *Best practices in school psychology* (pp. 17–36). Bethesda, MD: National Association of School Psychologists.
- Tinsley, H. E. A., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of Counseling Psychology, 34*(4), 414-424.
- Tobin, T., Sugai, G., & Colvin, G. (1996). Patterns in middle school discipline records. *Journal of Emotional and Behavioral Disorders, 4*, 82–94. doi: 10.1177/106342669600400203
- Tobin, T. J., & Vincent, C. G. (2011). Strategies for preventing disproportionate exclusions of African American students. *Preventing School Failure, 55*, 192-201. doi: 10.1080/1045988X.2010.532520
- Todd, A. W., Campbell, A. L., Meyer, G. G., & Horner, R. H. (2008). The effects of a targeted intervention to reduce problem behaviors: Elementary school

- implementation of check-in check-out. *Journal of Positive Behavior Interventions*, 10, 46-55. doi: 10.1177/1098300707311369
- U.S. Census Bureau (2015, March). *U.S. Census Bureau: State and county QuickFacts*. Retrieved from <http://quickfacts.census.gov/qfd/states/08000.html>
- U.S. Department of Education. (2014). *Guiding principles: A resource guide for improving school climate and discipline*. Retrieved from <http://www2.ed.gov/policy/gen/guid/school-discipline/guiding-principles.pdf>
- U. S. Department of Education, Office of Civil Rights. (2014). Civil rights data collection data snapshot: school discipline. Retrieved from <http://www2.ed.gov/about/offices/list/ocr/docs/crdc-discipline-snapshot.pdf>
- VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, 45, 225-256. doi: 10.1016/j.jsp.2006.11.004
- Vincent, C., Spaulding, S., & Tobin, T. J. (2010). A reexamination of the psychometric properties of the School-Wide Evaluation Tool (SET). *Journal of Positive Behavior Interventions*, 12(3), 161-179. doi: 10.1177/1098300709332345
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Walker, H. M., Horner, R. H., Sugai, G., Bullis, M., Sprague, J. R., Bricker, D., & Kaufman, M. J. (1996). Integrated approaches to preventing antisocial behavior patterns among school-age children and youth. *Journal of Emotional and Behavioral Disorders*, 4(4), 194-209. doi:10.1177/106342669600400401

- Walker, H. M., & Shinn, M. R. (2002). Structuring school-based interventions to achieve integrated primary, secondary, and tertiary prevention goals for safe and effective schools. In M. R. Shinn, G. Stoner, & H. M. Walker (Eds.), *Interventions for academic and behavior problems: Preventive and remedial approaches* (pp. 1-26). Silver Springs, MD: National Association of School Psychologists.
- Xu, S. & Lorber, M. F. (2014). Interrater agreement statistics with skewed data: Evaluation of alternatives to Cohen's Kappa. *Journal of Consulting and Clinical Psychology, 82*(6), 1219-1227. <http://dx.doi.org/10.1037/a0037489>

## APPENDIX A

### Check-In/Check-Out Fidelity Of Implementation Measure (CICO-FIM)

Version - Paper-Pencil/Microsoft Word

#### **Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM)**

**Directions:** The purpose of this tool is to measure the fidelity of implementation of Check-In, Check-Out (CICO) in a school or facility. The items in this measure will also help teams' action plan to increase the efficacy of their CICO implementation.

Please complete the CICO-FIM one to three months after initial CICO training (Baseline) and then annually each year after training. The evaluation questions should be completed by the CICO coordinator or by the CICO school-based team. If a team completes the CICO-FIM, the team must come to consensus and provide one answer per question. It is also important to note that each question provides a prompt for a data source to be utilized to score each item. The collection of this data is not required but suggested.

Thank you for taking the time to complete this tool.

**Positive Behavioral Interventions and Supports (PBIS)  
Check-In, Check-Out Fidelity of Implementation Measure**

**School:** \_\_\_\_\_ **Date:** \_\_\_\_\_ **Pre:** \_\_\_\_\_ **Baseline:** \_\_\_\_\_ **Annual:** \_\_\_\_\_

**District:** \_\_\_\_\_ **State:** \_\_\_\_\_ **Data collector:** \_\_\_\_\_

<u>Evaluation Question</u>	<u>Data Source</u> <b>P = Permanent product; I = Interview; O= Observation</b>	<u>Score</u> <b>0-2</b>
1. Does the school employ a CICO coordinator whose job is to manage the CICO (10-15 hours per week allocated)? (0 = No CICO Coordinator, 1 = CICO coordinator but less than 10 hours per week allocated, 2= CICO Coordinator, 10-15 hours per week allocated)	Interviews with Administrator & CICO Coordinator	<b>I</b>
2. Does the school budget contain an allocated amount of money to maintain the CICO ?(e.g. money for reinforcers, DPR forms, etc.)? (0 = No, 2 = Yes)	CICO Budget Interviews	<b>P / I</b>
3. Do students who are referred to the CICO receive support within a week? (0 = more than 2 weeks between referral and CICO support, 1 = within 2 weeks, 2 = within a week)	Interview CICO Referrals & CICO Start dates	<b>P / I</b>
4. Does the administrator serve on the CICO team or review CICO data on a regular basis? (0 = no, 1 = yes, but not consistently, 2 = yes)	Interview	<b>I</b>
5. Do 90% of CICO team members state that the CICO system has been taught/ reviewed on an annual basis? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)	Interview	<b>I</b>
6. Do 90% of the students on the CICO <u>check-in daily</u> ? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)	CICO recording form	<b>P</b>
7. Do 90% of students on the CICO <u>check-out daily</u> ? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)	CICO recording form	<b>P</b>
8. Do 90% of students on the CICO report that they receive reinforcement (e.g. verbal, tangible) for meeting daily goals? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)	Interview students on CICO	<b>I</b>
9. Do 90% of students on the CICO receive regular feedback from teachers? (randomly sample 50% of student DPR's across 3 days) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)	CICO Daily Progress Reports	<b>P</b>
10. Do 90% of students on the CICO receive feedback from their parents? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)	CICO Daily Progress Reports	<b>P</b>
11. Does the CICO coordinator enter DPR data at least once a week? (0 = no, 1 = 1= every other week, 2 =once a week)	Interview	<b>I</b>
12. Do 90% of CICO team members indicate that the daily CICO data is used for decision-making? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)	Interview	<b>I</b>

Adapted from Crone, Hawken, &, Horner, 2010

**APPENDIX B**

Check-In, Check-Out Fidelity Of Implementation Measure (CICO-FIM)

Version - Online Form

## Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM)

**Directions:**

The purpose of this tool is to measure the fidelity of implementation of Check-In, Check-Out (CICO) in a school or facility. The items in this tool will also help teams' action plan for increasing the efficacy of their CICO implementation.

Please complete the form below one to three months after initial CICO training (Baseline) and then annually each year after training. The evaluation questions should be completed by the CICO coordinator or by the CICO school based team. If a team completes the CICO-FIM, the team must come to consensus and provide one answer per question. All questions with red stars require an answer for completion of the tool. It is also important to note that each question provides a prompt for a data source to be utilized to score each item. The collection of this data is not required but suggested.

To ensure your measure has been collected in its' entirety, please choose submit at the end of the tool. The PBIS Implementation Consultant (IC) that supports your district in PBIS will follow-up with your sites CICO-FIM data and action plan.

Thank you for taking the time to complete this tool. Annual results will be shared with you by the CO PBIS initiative.

Date

School Name

District Name

State

Name of Person Completing Form  
(If team completed, please just write "Team")

Please provide email address for data  
report and follow-up:

Trained in CICO by

- Colorado Department of Education Personnel
- District Level Personnel
- Self (e.g. Video/ text)
- Other

Submission Type:

- Pre Implementation
- Baseline (post training initial implementation)
- Annual

---

EVALUATION QUESTIONS:

1. Does the school employ a Check-In, Check-Out (CICO) coordinator whose job is to manage the CICO with 10-15 hours per week allocated?

(0 = No CICO Coordinator, 1 = CICO Coordinator but less than 10 hours per week allocated, 2 = CICO coordinator with 10-15 hours per week allocated)

Data Source

Score

- Interview(s) with administrator & CICO coordinator
- Permanent Product
- Observation
- Other

- 0
- 1
- 2

---

2. Does the school budget contain an allocated amount of money to maintain the CICO? (e.g. money for reinforcers, Daily Progress Reports (DPR) forms, etc)

(0 = No, 2 = Yes)

Data Source

Score

- Interview(s)
- Permanent Product - CICO Budget
- Observation
- Other

- 0
- 1
- 2

---

3. Do students who are referred to CICO receive support within a week of the referral?

(0 = more than 2 weeks between referral and CICO support, 1 = within 2 weeks, 2 = within a week)

Data Source

Score

- Interview(s)
- Permanent Product - CICO referrals & start dates
- Observation
- Other

- 0
- 1
- 2

---

4. Does the administrator serve on the CICO team or review CICO data on a regular basis?  
(0 = no, 1 = yes but not consistently, 2 = yes)

Data Source	Score
<input type="checkbox"/> Interview(s) <input type="checkbox"/> Permanent Product	<input type="radio"/> 0
<input type="checkbox"/> Observation <input type="checkbox"/> Other	<input type="radio"/> 1
<input type="text"/>	<input type="radio"/> 2

---

5. Do 90% of CICO team members state that the CICO system has been taught/reviewed on an annual basis?  
(0 = 0-50%, 1 = 51%-89%, 2 = 90%-100%)

Data Source	Score
<input type="checkbox"/> Interview(s) <input type="checkbox"/> Permanent Product	<input type="radio"/> 0
<input type="checkbox"/> Observation <input type="checkbox"/> Other	<input type="radio"/> 1
<input type="text"/>	<input type="radio"/> 2

---

6. Do 90% of the students on CICO check in daily (randomly sample 3 days of recording)?  
(0 = 0-50%, 1 = 51%-89%, 2 = 90%-100%)

Data Source	Score
<input type="checkbox"/> Interview	<input type="radio"/> 0
<input type="checkbox"/> Permanent Product - CICO Forms/DPRs	<input type="radio"/> 1
<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

---

7. Do 90% of the students on CICO check out daily (randomly sample 3 days of recording)?  
(0 = 0-50%, 1 = 51%-89%, 2 = 90%-100%)

Data Source	Score
<input type="checkbox"/> Interview	<input type="radio"/> 0
<input type="checkbox"/> Permanent Product - CICO Forms/DPRs	<input type="radio"/> 1
<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

---

<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

8. Do 90% of the students on CICO report that they receive reinforcement (e.g. verbal, tangible) for meeting daily goals?  
(0 = 0-50%, 1 = 51%-89%, 2 = 90%-100%)

Data Source	Score
<input type="checkbox"/> Interviews with students on CICO	<input type="radio"/> 0
<input type="checkbox"/> Permanent Product	<input type="radio"/> 1
<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

---

9. Do 90% of the students on CICO report that they receive regular feedback from teachers (randomly sample 50% of student DPRs across 3 days)?  
(0 = 0-50%, 1 = 51%-89%, 2 = 90%-100%)

Data Source	Score
<input type="checkbox"/> Interview	<input type="radio"/> 0
<input type="checkbox"/> Permanent Product - CICO DPR	<input type="radio"/> 1
<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

---

10. Do 90% of the students on CICO report that they receive regular feedback from their parents (randomly sample 50% of student DPRs across 3 days)?  
(0 = 0-50%, 1 = 51%-89%, 2 = 90%-100%)

Data Source	Score
<input type="checkbox"/> Interview	<input type="radio"/> 0
<input type="checkbox"/> Permanent Product - CICO DPR	<input type="radio"/> 1
<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

---

11. Does the CICO coordinator enter DPR data at least once a week?  
(0 = no, 1 = every other week, 2 = once a week)

Data Source	Score
<input type="checkbox"/> Interview(s)	<input type="radio"/> 0
<input type="checkbox"/> Permanent Product	<input type="radio"/> 1
<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

---

12. Do 90% of the CICO team members indicate that the daily CICO data is used for decision making?

(0 = 0-50%, 1 = 51%-89%, 2 = 90%-100%)

Data Source	Score
<input type="checkbox"/> Interview(s)	<input type="radio"/> 0
<input type="checkbox"/> Permanent Product	<input type="radio"/> 1
<input type="checkbox"/> Observation	<input type="radio"/> 2
<input type="checkbox"/> Other <input type="text"/>	

---

Adapted from Crone, Hawken, & Horner, 2010

Submit

## APPENDIX C

### Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM)

#### Content Review Expert Invitation

Erin A. Sullivan, Ph.D. Candidate  
Morgridge College of Education  
University of Denver

Dear [Fill in blank].

Hello. My name is Erin A. Sullivan and I am a Ph.D. candidate at the University of Denver. I am writing to invite you to review the Check-In/Check-Out Fidelity Of Implementation Measure (CICO-FIM) for content validation purposes. This instrument was designed to assess the implementation fidelity of the Check-In/Check-Out (CICO) intervention. As a recognized expert, your participation this project will involve examining the CICO-FIM for content validity. Specifically, it will entail assessing the appropriateness of the content of this instrument in measuring the implementation of CICO. Participation in this project is strictly voluntary and confidential and should take about 45 minutes of your time. A copy of the study results will be made available to you, if interested.

This study is being conducted to fulfill the requirements of a Ph.D. dissertation and has been approved by the University of Denver Institutional Review Board (pending). This project is supervised by Dr. Gloria Miller, Morgridge College of Education, University of Denver, Denver, CO 80208, 303-871- 3340. [Gloria.Miller@du.edu](mailto:Gloria.Miller@du.edu). Results of this study will be used for presentation, publication, and in fulfillment of the requirements of a Ph.D. If you have questions, I can be reached via email at [Erin.Sullivan@du.edu](mailto:Erin.Sullivan@du.edu) or by phone at 773-817-4088.

If you are interested in participating, please read the attached informed consent letter and sign the signature page. The signature page can be returned as a pdf file via email. If you would prefer a hard copy, please let me know, and I would be happy to provide the forms with a self-addressed, stamped return envelope. I truly appreciate your time, expertise, and assistance.

Sincerely,

Erin A. Sullivan, M.S.Ed., M.A.  
Ph.D. Candidate, Child, Family, and School Psychology  
Morgridge College of Education, University of Denver  
[Erin.Sullivan@du.edu](mailto:Erin.Sullivan@du.edu)  
773-817-4088

## **APPENDIX D**

### **Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM)**

#### **Content Review Protocol**

Erin A. Sullivan, Ph.D. Candidate  
Morgridge College of Education  
University of Denver

As an expert in the field, you have been invited to review the Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM) for content validation purposes. There are two components of the review of the content of the CICO-FIM. The first component is the appraisal of item relevance, clarity, and difficulty to agree with. Please use the tables provided to rate each item according to its degree of relevance, clarity, and difficulty in implementing CICO. The second component is an open-ended reflection of the instrument as a whole. Please refer to the CICO-FIM (attached to original email) to respond to these items. Please also feel free to make any suggestions, comments, or recommendations on the open-ended reflection protocol.

Please return the completed form electronically to Erin A. Sullivan at [erinasullivan@me.com](mailto:erinasullivan@me.com).

Thank you for taking the time to share your expertise.

Check-In, Check-Out Fidelity of Implementation Measure (CICO-FIM)

**I. Content Review Protocol: Item Relevance**

**Directions:** Please mark the degree of relevance of each item to the implementation of Check-In, Check-Out (CICO) in a school. Mark each item with an “X” in the appropriate box denoting the degree of relevance of each item.

	Evaluation Question	<i>How relevant is this item to the school-based implementation of Check-In, Check-Out?</i>		
		Highly Relevant	Moderately Relevant	Low Level of Relevance
1	Does the school employ a CICO coordinator whose job is to manage the CICO (10-15 hours per week allocated)? (0 = No CICO Coordinator, 1 = CICO coordinator but less than 10 hours per week allocated, 2= CICO Coordinator, 10-15 hours per week allocated)			
2	Does the school budget contain an allocated amount of money to maintain the CICO ?(e.g. money for reinforcers, DPR forms, etc.)? (0 = No, 2 = Yes)			
3	Do students who are referred to the CICO receive support within a week? (0 = more than 2 weeks between referral and CICO support, 1 = within 2 weeks, 2 = within a week)			
4	Does the administrator serve on the CICO team or review CICO data on a regular basis? (0 = no, 1 = yes, but not consistently, 2 = yes)			
5	Do 90% of CICO team members state that the CICO system has been taught/ reviewed on an annual basis? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			
6	Do 90% of the students on the CICO <u>check-in daily</u> ? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			

7	Do 90% of students on the CICO <u>check-out daily</u> ? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
8	Do 90% of students on the CICO report that they receive reinforcement (e.g. verbal, tangible) for meeting daily goals? (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
9	Do 90% of students on the CICO receive regular feedback from teachers? (randomly sample 50% of student DPR's across 3 days) (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
10	Do 90% of students on the CICO receive feedback from their parents? (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
11	Does the CICO coordinator enter DPR data at least once a week? (0 = no, 1 = once every other week, 2 =once a week)			
12	Do 90% of CICO team members indicate that the daily CICO data is used for decision-making? (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			

Evaluation questions and scale are adapted from Crone, Hawken, and Horner, 2010

## II. Content Review Protocol: Item Clarity

**Directions:** Please mark the degree of clarity of each item. Mark each item with an “X” in the appropriate box denoting the degree of clarity of each item. For the purpose of this study, an item with high clarity is an item that is brief, explicit, uses common language, avoids jargon or slang, and asks one clear question.

	Evaluation Question	<i>Please rate the clarity of each item.</i>		
		High	Moderate	Low
1	Does the school employ a CICO coordinator whose job is to manage the CICO (10-15 hours per week allocated)? (0 = No CICO Coordinator, 1 = CICO coordinator but less than 10 hours per week allocated, 2= CICO Coordinator, 10-15 hours per week allocated)			
2	Does the school budget contain an allocated amount of money to maintain the CICO ?(e.g. money for reinforcers, DPR forms, etc.)? (0 = No, 2 = Yes)			
3	Do students who are referred to the CICO receive support within a week? (0 = more than 2 weeks between referral and CICO support, 1 = within 2 weeks, 2 = within a week)			
4	Does the administrator serve on the CICO team or review CICO data on a regular basis? (0 = no, 1 = yes, but not consistently, 2 = yes)			
5	Do 90% of CICO team members state that the CICO system has been taught/ reviewed on an annual basis? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			
6	Do 90% of the students on the CICO <u>check-in</u> daily? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			
7	Do 90% of students on the CICO <u>check-out</u> daily? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			
8	Do 90% of students on the CICO report that they receive reinforcement (e.g. verbal, tangible) for meeting daily goals? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			

9	Do 90% of students on the CICO receive regular feedback from teachers? (randomly sample 50% of student DPR's across 3 days) (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
10	Do 90% of students on the CICO receive feedback from their parents? (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
11	Does the CICO coordinator enter DPR data at least once a week? (0 = no, 1 = once every other week, 2 = once a week)			
12	Do 90% of CICO team members indicate that the daily CICO data is used for decision-making? (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			

Evaluation questions and scale are adapted from Crone, Hawken, and Horner, 2010

### III. Content Review Protocol: Item Difficulty

**Directions:** Please mark the degree of difficulty of each item. Mark each item with an “X” in the appropriate box denoting the degree of difficulty of each item. Please rate each item according to how *difficult* you perceive it would be for a school implementing CICO to obtain the highest score relative to this item.

Item	Evaluation Question	<i>Please rate the difficulty of each item.</i>		
		High	Moderate	Low
1	Does the school employ a CICO coordinator whose job is to manage the CICO (10-15 hours per week allocated)? (0 = No CICO Coordinator, 1 = CICO coordinator but less than 10 hours per week allocated, 2= CICO Coordinator, 10-15 hours per week allocated)			
2	Does the school budget contain an allocated amount of money to maintain the CICO ?(e.g. money for reinforcers, DPR forms, etc.)? (0 = No, 2 = Yes)			
3	Do students who are referred to the CICO receive support within a week? (0 = more than 2 weeks between referral and CICO support, 1 = within 2 weeks, 2 = within a week)			
4	Does the administrator serve on the CICO team or review CICO data on a regular basis? (0 = no, 1 = yes, but not consistently, 2 = yes)			
5	Do 90% of CICO team members state that the CICO system has been taught/ reviewed on an annual basis? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			
6	Do 90% of the students on the CICO <u>check-in</u> daily? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			
7	Do 90% of students on the CICO <u>check-out</u> daily? (Randomly sample 3 days for recording) (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			
8	Do 90% of students on the CICO report that they receive reinforcement (e.g. verbal, tangible) for meeting daily goals? (0 = 0-50%, 1 = 51-89%, 2 = 90–100%)			

9	Do 90% of students on the CICO receive regular feedback from teachers? (randomly sample 50% of student DPR's across 3 days) (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
10	Do 90% of students on the CICO receive feedback from their parents? (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			
11	Does the CICO coordinator enter DPR data at least once a week? (0 = no, 1 = once every other week, 2 = once a week)			
12	Do 90% of CICO team members indicate that the daily CICO data is used for decision-making? (0 = 0-50%, 1 = 51-89%, 2 = 90-100%)			

Evaluation questions and scale are adapted from Crone, Hawken, and Horner, 2010

#### IV. Content Review Protocol: Open-Ended Response

Directions: Please respond to the following questions using the attached copy of the CICO-FIM. Please mark yes or no for question 1, 2, and 3 and then respond accordingly.

1. Overall, does the CICO-FIM reflect the necessary components to implement Check-In, Check-Out (CICO)?	Yes	No
If no, please explain.		
2. Are the instructions clear and concise?	Yes	No
If no, what would make them better?		
3. Are there any items that are awkward or confusing (Please refer to the Item Clarity Protocol)?	Yes	No
If yes, please suggest alternative wordings?		
4. Is there a component necessary to implement CICO that the instrument is missing or anything you think should be included that was not measured?		
5. Are there any questions that are redundant or not essential to implement CICO? If so, which items would you delete or change?		
6. Any other suggestions, questions, or comments?		