



Published in final edited form as:

Drug Alcohol Depend. 2010 June 1; 109(1-3): 167–174. doi:10.1016/j.drugalcdep.2009.12.027.

Psychometric Properties of the Contingency Management Competence Scale

Nancy M. Petry^{a,*}, Sheila M. Alessi^a, David M. Ledgerwood^{a,b}, and Sean Sierra^a

^a Calhoun Cardiology Center, MC-3944, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030-3944

^b Department of Psychiatry and Behavioral Neurosciences, Wayne State University School of Medicine, 2761 E. Jefferson Ave., Detroit, MI 48207

Abstract

Contingency management (CM) is an evidence-based treatment, and clinicians are beginning to implement this intervention in practice. However, little research exists on methods for assuring appropriate implementation of CM. This study describes the development and psychometric properties of the 12-item CM Competence Scale (CMCS). Thirty-five therapists from 9 community-based clinics participated; following a training period, a randomized trial evaluated the efficacy of CM in cocaine abusing patients. Analyses of the CMCS are based on ratings from 1,613 audiotapes of therapist interactions with 78 patients enrolled in the training phase and 103 patients in the randomized phase. Inter-rater reliability from 11 raters and internal consistency of items on the CMCS was good to excellent. Items loaded onto two factors: one contained items specific to discussions of the outcomes of urine testing and reinforcement, and the other contained general items related to use of praise, communication of confidence, empathy, skillfulness, and maintaining session structure, as well as discussions of self-reports of drug use when they occurred. During the training phase in CM delivery, scores on the CMCS rose significantly between earlier and later training sessions, and during the randomized phase, CM sessions were rated more highly than non-CM sessions. Scores on the subscale assessing general items were significantly correlated with indices of the therapeutic alliance and predictive of durations of cocaine abstinence achieved. These data suggest that the CMCS is reliable and valid in assessing delivery of CM and that competence in CM delivery is associated with improved patient outcomes.

Keywords

contingency management; therapist; training; competence

1. Introduction

Contingency management (CM) is an empirically-based intervention for treating substance use disorders. In a recent meta-analysis and review of the literature, Dutra et al. (2008) concluded

* To whom all correspondence should be addressed. Calhoun Cardiology Center, MC-3944, University of Connecticut School of Medicine, 263 Farmington Avenue, Farmington, CT 06030-3944 Phone: 860-679-2593, Fax: 860-679-1312, Npetry@uchc.edu.

Clinical trials registry: NCT00249418

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

that CM was the psychosocial intervention with the greatest evidence of efficacy for substance abusing patients. Theoretically, CM is based on basic operant principles-- a behavior that is reinforced will increase in frequency. The central tenants of CM are to: frequently monitor the behavior targeted for change, reinforce the behavior tangibly each time it occurs, and withhold reinforcement each time the behavior does not occur. Typically, in CM treatment of substance use disorders, drug abstinence is the behavior targeted for change. Urine sample monitoring occurs 2-3 times weekly, and patients earn reinforcers consisting of vouchers worth escalating monetary values or chances to win prizes ranging from \$1-\$100 in value each time abstinence is detected (Higgins et al., 2000; Petry, 2000).

In the recent National Institute on Drug Abuse's Clinical Trials Network (CTN) studies of CM in the United States (Peirce et al., 2006; Petry et al., 2005a), over 800 stimulant abusing patients from 14 community clinics were randomized to standard care plus frequent urine testing or that same treatment plus the chance to win prizes contingent upon abstinence. The longest duration of abstinence achieved increased significantly in patients assigned to the CM condition relative to those assigned to the standard care condition. Meta-analyses likewise point to the efficacy of CM in treating substance abusers (Lussier et al., 2006; Prendergast et al., 2006).

With the completion of the CTN studies in CM, the National Institute on Drug Abuse and the Substance Abuse and Mental Health Services Administration in the United States created dissemination materials entitled Promoting Awareness of Motivational Incentives. Clinics in the United States and in other countries are beginning to implement CM in clinical settings (Cameron & Ritter, 2007; Garcia-Rodriguez et al., 2009; Lott and Jencius, 2009; Kellogg et al., 2005; Squires et al., 2008). As interest in CM grows, more effort needs to be directed toward training resources in CM delivery.

In the only known study systematically evaluating adherence to CM implementation, Andrzejewski et al. (2001) noted that without regular feedback from CM experts, clinicians' administration of CM was often suboptimal, with clinicians failing to monitor and reinforce behaviors according to pre-specified criteria. As with any therapy, and perhaps particularly behavioral therapies, effectiveness is likely to be diminished if the intervention is not administered appropriately.

To date, relatively little research has attended to evaluation of the adherence and competence of CM delivery. In most research trials, CM administrators are carefully selected, highly trained, and closely supervised research assistants or research therapists (Budney et al., 2006; Higgins et al., 2007; Petry et al., 2000, 2004, 2005b, 2007). Principal investigators or project directors with extensive experience directly train and oversee CM administration and quickly correct problems that may arise. Sometimes, checklists assist in appropriate protocol specific administration of CM (Petry et al., 2005ab, 2007), but no known studies have reported on psychometric properties of specific instruments for evaluating competence in CM delivery.

Chapman and colleagues (2008) describe a scale that assesses aspects of adherence to CM. The CM-Therapist Adherence Measure was designed to evaluate adherence to cognitive-behavioral therapy in conjunction with CM when administered to substance abusing adolescents participating in family-based interventions. This scale contains 3-4 items that assess two aspects of CM (conducting drug testing and providing consequences in a general manner), and the other 5 items evaluate use of cognitive-behavioral strategies. Henggeler et al. (2008) report that overall therapist-reported adherence on this instrument is significantly associated with reductions in biological and self-report measures of marijuana use among substance abusing youths. This instrument, however, was not designed to evaluate features of CM such as specifying the nature of the reinforcement and providing escalating reinforcers, both of which are associated with sustained behavioral change (Roll et al., 1996; Higgins et

al., 2000; Petry et al., 2005b, 2006b, 2007). Although useful for assessing adherence to aspects of family-based interventions for adolescents, this instrument was not designed to assess therapists' competence or quality of CM delivery, nor has it been evaluated with respect to CM delivery with adults.

The absence of tools for evaluating competence in CM treatment delivery may impede effective implementation of CM in community settings. Tools for monitoring competence of therapy delivery exist for other interventions including supportive-expressive therapy (Barber et al., 1996a), motivational interviewing (Martino et al., 2008; Miller et al., 2003; Moyers et al., 2003), cognitive-behavioral therapy (Carroll et al., 2000), and drug counseling (Barber et al., 1996a). Therapist competence is associated with improved outcomes of patients who are treated via these interventions (Barber et al., 1996a; Martino et al., 2008).

The purpose of this study is to describe the development and psychometric properties of the Contingency Management Competence Scale (CMCS). Because a fidelity instrument will ultimately be most useful for practice settings, data were drawn from community-based therapists. We evaluated inter-rater reliability, internal consistency, factor structure, and discriminant, concurrent and predictive validity of the CMCS.

2. Methods

2.1 Participants

Therapist subjects were 35 clinicians from one of nine substance abuse treatment programs (3 were methadone programs and the remainder outpatient psychosocial programs). All had at least one year of experience treating substance abusers, and at least 6 months employment at the participating clinic, reported sufficient time availability for training in CM delivery, and were willing to allow random assignment of their patients to CM and standard care conditions. Average (and standard deviation) age was 46.2 ± 12.4 years, 48.6% ($n = 17$) were male, 5.7% ($n = 2$) were African American, 14.3% ($n = 5$) were Hispanic, 71.4% ($n = 25$) were Caucasian, and 8.6% ($n = 3$) were of other ethnicities. In terms of education, 51.3% ($n = 18$) had a master's degree, and 2.9% ($n = 1$) a Psy.D., with the remainder having a bachelors degree (28.6%, $n = 30$) or less education (14.4%, $n = 5$). The median (interquartile range) length of experience in substance abuse treatment was 69 (144) months, and 42.9% ($n = 15$) of the therapists had a license to practice in the mental health field.

Patients ($n = 181$) were enrolled in a study comparing the efficacy of standard care plus frequent urine testing to that same treatment with CM ($n = 103$) or participated in a training phase in CM delivery associated with that same study ($n = 78$). Each patient met the following inclusion criteria: cocaine abusing or dependent, enrolled in substance abuse treatment, and age 18 years or older. Significant uncontrolled psychiatric conditions (e.g., suicidality, psychosis) and in recovery from pathological gambling (due to the potential similarity between gambling and the reinforcement system) were exclusionary criteria. On average, patients were 37.4 ± 8.7 years of age, and 52.4% ($n = 95$) were male. Eight percent ($n = 14$) were African American, 17.5% ($n = 32$) Hispanic, 61.2% ($n = 111$) Caucasian and 13.3% ($n = 24$) of other ethnicities. Twenty percent ($n = 36$) were married or cohabitating, 49.5% ($n = 90$) had never married, and the remainder were divorced, separated or widowed (30.5%, $n = 55$). On average, patients had 2.1 ± 1.7 prior substance abuse treatment attempts. In addition to meeting criteria for a cocaine use disorder, 12.6% ($n = 23$) of patients also met past-year criteria for alcohol dependence and 26.2% ($n = 47$) for marijuana dependence. Opioid-dependent patients ($n = 103$) maintained on methadone received a mean dose of 88.0 ± 25.6 mg/day.

The University of Connecticut Health Center Institutional Review Board approved this study. Therapist and patient participants signed informed consent for study procedures, including audiotaping of interactions.

2.2 Development of the CMCS

The CM specific items were generated from review of session audiotapes and treatment manuals on CM delivery (e.g., Petry and Stitzer, 2003). As recommended by Waltz et al. (1993), items were selected that reflected use of elements that are unique and essential to CM but would not be found in other treatment approaches. We initially included four items that assessed to what extent the therapist discussed draws earned at that session, draws available at the next session, the patient's desire for prize items, and enthusiasm for the patient's choice of prizes. In a review of 195 audiotapes of CM sessions, ratings on these last two items correlated at $r = 0.94$, and therefore, the item about therapist's enthusiasm for the client's preference of prizes was removed from subsequent analyses. The item about assessing the patient's desire for prizes was selected because it is applicable to all patients receiving CM, whereas the enthusiasm item is most relevant when patients win prizes. The three items included in the final scale are shown in Table 1 (items 2-4).

Waltz et al. (1993) also recommended rating interventions essential to the treatment being evaluated but not unique to it. The CMCS included four such items that are commonplace in many treatments including CM, but are not unique to particular interventions (numbered 1, 5, 6 and 7 in Table 1). These items related to discussions of the outcomes of testing and assessment of self-reports of substance use; if self-reported drug use occurred, two additional items scored including relating self-reports to objective indicators, and relating self-reports to consequences of positive samples.

We also included 5 non-specific items. These items are labeled 8-12 in Table 1, and they tap therapists' use of praise, confidence, general effectiveness, maintenance of session structure, and empathy. Each item was rated using a 7-point scale, with higher numbers reflecting better usage. Competence ratings anchored the 1 with the phrase "Very poor," and the 7 was anchored with the word "Excellent." The interim numbers included the following descriptions: 2 = "Poor," 3 = "Barely acceptable," 4 = "Acceptable," 5 = "Good," and 6 = "Very good."

Because the CMCS is intended to assess therapists' ability to deliver CM, items were written to focus on the therapist, not the patient. Thus, each item was phrased as, "To what extent did the therapist...". To increase reliability, items were worded so as to be as specific and concrete as possible and to focus on observable therapist behaviors.

As suggested by Carroll et al. (2000), most items were rated with respect to both adherence (the "quantity" or degree to which it was present in the session) and competence (the "quality" or "skillfulness" with which the therapist delivered the intervention). However, after the initial review of 195 audiotaped sessions, adherence and competence ratings correlated at 0.92 - 0.99 for each item. These high levels of correlation suggest that the adherence and competence ratings were not unique, and only competence ratings are presented henceforth. Competence ratings were chosen over adherence ratings for consistency with the three general items (skillfulness, maintaining structure and empathy), which were rated only with respect to competence as adherence ratings seemed inappropriate. Further, one can adhere to a therapy without doing so competently, so competence ratings appeared most germane.

2.3 Rater selection and training

Raters in this study were bachelor to master level research assistants with experience in CM delivery ($n = 9$) or doctoral level researchers with expertise in CM administration ($n = 2$).

Procedures for rater training paralleled those used in other studies (Carroll et al., 1994; Carroll and Nuro, 1997). Raters attended a didactic seminar that included detailed review of the rating sheet and several group practice ratings using taped examples. They then rated at least 10 audiotapes, which were previously rated with respect to consensus of the larger group. Raters were considered “certified” only after their ratings achieved acceptable reliability ($> .80$) with respect to the consensus ratings. Rater recalibration sessions were held regularly to monitor and prevent rater drift, and reliability was formally evaluated several times during the conduct of the study. In this process, each of the raters rated a small set of tapes, and they were unaware of which tapes were used to calculate interrater reliability. Because several raters left over the time frame of this study and several others joined the project, no sessions were rated by all raters, but 99 sessions were rated by three or more of the 11 individuals, and were included in our inter-rater reliability analysis.

2.4 Overview of the clinical trial

This report focuses on the development and psychometric properties of the CMCS, so we only briefly describe the clinical trial (Petry et al., unpublished data) from which data were drawn to validate this instrument. The trial involved training community-based therapists to deliver CM to cocaine abusing patients and evaluating the efficacy of CM when delivered by community-based therapists.

Following a training period in which clinicians attended a didactic seminar and read information about CM and the specific trial, they participated in at least three mock CM sessions in which they played the part of a therapist and a research assistant assumed the role of a patient. Role plays were rated by a CM expert, and when they achieved average ratings of “4” or higher on CMCS items (indicative of “Acceptable” quality), therapists proceeded to a training phase in CM delivery. In this phase, patients earned one draw from a bowl with a chance of winning small (\$1), large (\$20) or jumbo (\$100) prizes for each cocaine-negative urine sample submitted, and number of draws increased with successive submissions of negative samples. Delivery of CM was audiotaped twice weekly for up to 12 weeks for 2-3 patients, and a CM expert reviewed each tape and provided feedback to the therapists weekly regarding their implementation of CM. After the therapist achieved ratings equal to or exceeding 4 for at least three sessions in a row with at least two patients, the therapist then proceeded to a randomized phase.

In this final phase, cocaine abusing patients were randomized to standard care with twice weekly urine monitoring or that same treatment and the chance of winning prizes contingent upon submission of cocaine negative samples. The CM administered during the randomized phase was identical to that administered in the training phases; however, in the randomized phase, therapists were not provided feedback regarding their implementation of CM or standard care, and they administered both conditions to patients randomized in this phase. In both the CM training phase and the randomized phase, interactions related to urine collection and CM delivery were audiotaped, and these tapes serve as the basis of the analyses described in this report.

2.5 Data analysis

Inter-rater reliability was examined via intraclass correlation coefficients (ICC) for each of the CMCS items with respect to competence. Shrout and Fleiss (1979) random effects model estimated reliabilities among raters. A total of 99 sessions were independently rated by three or more of the 11 raters and were used in the analyses of inter-rater reliability.

Cronbach's alpha evaluate the internal consistency of the CMCS items. Internal consistency was assessed in a randomly selected sample of 177 CM sessions from the randomized phase

in which 22 fully trained therapists participated, after having successfully completed the CM training phase. Because three of the CMCS items were relevant only when self-reports of substance use occurred during sessions, the initial analyses excluded the self-report of drug use items. Internal consistency of the full 12-item scale was assessed with another 78 audiotapes obtained from CM sessions occurring in the randomized phase in which patients self-reported substance use.

The factor structure of the CMCS was also evaluated. The number of factors comprising the scale was not decided a priori, so data were subjected to a Principal Components Analysis (PCA), using varimax rotation and Kaiser Normalization. PCA was conducted twice-- from the larger group of randomly selected tapes ($n = 177$) during which self-reported substance use may or may not have occurred, and from tapes of sessions in which patients self-reported drug use ($n = 78$). In the former analyses, the three items related to self-reports of drug use were excluded, while in the secondary analyses, the PCA included these three items.

Discriminant validity of the CMCS was evaluated along two dimensions. First, we assessed whether therapists learning to administer CM showed improvement in CMCS scores during the training period. Thirty-five therapists participated in the training phase and collectively treated 78 patients using CM. In total, 809 CM sessions were audiotaped during the training phase. Data for the discriminant validity analyses were drawn from 342 of these sessions; the 467 sessions occurring in the middle of the training phase for each therapist were not included in the analyses. For each therapist, ratings from 5 early sessions in CM delivery were averaged, and these ratings were compared with average ratings from 5 later sessions derived from the same therapist. In two cases, therapists did not have 10 audiotapes of sessions available, and data from the initial 2-4 sessions were then compared with data from the last 2-4 sessions for that therapist. Paired t-tests compared average ratings from the early and later training sessions between each therapist.

The second evaluation of discriminant validity involved comparing CMCS ratings of CM sessions to ratings from non-CM sessions during the randomized phase. For each patient randomly assigned to a treatment condition ($n = 103$), ratings on the overall scale and for each item were averaged across all the sessions audiotaped for that patient (range: 1-21; the mean and standard deviation was 7.8 ± 5.7 tapes rated per patient, and 804 tapes were collected and rated in total during this phase). Independent t-tests compared average CMCS ratings for patients randomized to the CM condition versus average CMCS ratings for patients randomized to the standard care condition.

Concurrent validity of the CMCS was assessed by correlating scores on this instrument with measurements of the therapeutic alliance. For each audiotaped session, a CMCS total score was derived by averaging the mean competence ratings across all items (9 items were averaged if no self-reported drug use occurred in the session and 12 items if self-reported drug use occurred). An overall score was then created for each randomized patient, using average total ratings from the 1-21 sessions rated for each patient. The 19-item Helping Alliance Scale (HAQ; Luborsky et al., 1996) was administered to patients ($n = 96$) and therapists ($n = 22$), who rated their perception of alliances with each of their patients. In total, therapists rated alliances with 92 patients at the end of the 3-month treatment period, and there were 91 pairs of HAQ ratings available from both the patient and the therapist. The missing data relate to failure of some patients ($n = 7$) to complete the follow-up, or misplaced or forgotten administrations of the HAQ therapist ratings for 11 patients. The HAQ is scored by summing up responses to the 19 items, after reverse scoring the four items that assess negative relationships or alliance. In this sample, Cronbach's alpha of the patient version of the HAQ was 0.883, and for the therapist version it was 0.863. Pearson's correlations evaluated

concordance between HAQ scores from both the patient and the therapist with CMCS overall and subscale scores.

Finally, predictive validity of the CMCS was evaluated among the 103 patients participating in the randomized phase. The main outcome measure was longest duration of cocaine abstinence achieved, which was calculated from the twice-weekly urine samples during the 12-week intervention period (range 0-12 weeks). Missed samples were considered positive and broke a string of abstinence. A univariate analysis evaluated the association of subscale scores on the CMCS with longest durations of abstinence achieved, after controlling for treatment group assignment. For these analyses, competence ratings from all audiotaped sessions available (range 1-21 per patient and 804 in total) were averaged for each patient.

3. Results

3.1 Inter-rater reliability

Intraclass correlation coefficients (ICC) were calculated to provide an estimate of reliability for each item of the CMCS and the overall scale. The 99 sessions that were rated by three or more of the 11 raters were included in these analyses, and ICCs are presented in Table 1. All the items were highly reliable, and ICCs ranged from 0.67 to 0.94 for the 12 competence items. The ICC for the overall scale was 0.92.

3.2 Internal consistency

Internal consistency of the CMCS scale was determined via 177 tapes of CM sessions from fully trained therapists. Only 9 of the 12 competence items were included in the primary analysis because self-reports of drug use did not occur in most CM sessions. Cronbach's alpha for the 9-item CMCS scale was 0.834. When sessions in which self-reports of drug use occurred and all 12 items were rated ($n = 78$ tapes), the internal consistency of the scale was 0.903.

3.3 Factor structure

The factor structure of the CMCS was evaluated using these same 177 tapes of CM and the 9 items that were rated in all sessions (i.e., regardless of whether or not self-reports of substance use occurred). Two factors emerged with Eigenvalues greater than 1, and factor loadings from the rotated component matrix are shown in the left-hand columns of Table 2. Six items loaded on the first factor. They consisted of items related to the therapist's assessment of the patients' desire for prizes, use of praise, communication of confidence, general effectiveness, maintenance of structure, and empathy. This factor explained 46.5% of the variance and was termed the General subscale. The second factor was termed the Draw subscale, and it contained three items: discussions of the outcomes of the testing and the number of draws earned and the number of draws possible at the next session. This factor explained 17.0% of the variance, and the two factors combined explained 63.5% of the variance.

A second PCA was conducted with the 78 tapes of CM sessions in which self-reported drug use occurred during the session. Again, two factors emerged, with the first reflecting general items along with the self-report items, and the second containing the three items related to testing outcomes and draws. Factor loadings of the items are also shown in Table 2. Two items (relating self-reports of drug use to consequences of positive samples and maintaining session structure) loaded above 0.40 on both factors, but all other items loaded on just a single factor. The factors explained 41.5% and 21.3% of the variance, respectively, for a total of 62.8%.

3.4 Discriminant validity

Table 3 shows mean and standard deviations of CMCS ratings for each therapists' initial attempts at CM delivery during the training phase. During even the first five sessions, therapists

administered CM competently with average scores between “5” and “6” reflecting “good” to “very good” competence for the scale as a whole and each item. Nevertheless, administration of CM improved by the end of the training period. Ratings from the later CM sessions averaged over a score of “6” for the overall scale and were greater than 6 on 10 of the 12 individual items. Significant differences ($p < .05$) between early and later administration of CM were noted in terms of overall CMCS and General subscale scores, as well as on five items: assessing the patients' desire for prizes, complimenting or praising the patients' efforts toward abstinence, communicating confidence, general skillfulness, and maintaining structure.

Table 4 shows mean and standard deviation therapist competence ratings derived from sessions of patients randomized to the CM condition ($n = 63$) versus those randomized to the standard care plus frequent urine testing condition ($n = 40$). Ratings from CM sessions averaged higher than “4” (indicating “adequate” or better) competence for each item and near “6” (“very good”) for the scale as a whole. Ratings from standard care sessions were significantly lower for the scale overall, as well as for both subscales. As expected, the items assessing the number of draws earned, draws possible at the next session, and desire for prizes (none of which should occur in standard care sessions), were all rated significantly lower during standard care sessions, with means of “1”. In addition, therapists were less skilled at relating self-reports to consequences of positive samples, complimenting or praising patients' efforts toward abstinence, and communicating confidence that the patients' efforts will yield success in standard care relative to CM conditions. General skillfulness, maintenance of session structure, empathy, and discussions of outcomes of biological testing were implemented with similar degrees of competence in CM and standard care sessions.

3.5 Concurrent validity

For each patient, scores on the CMCS, averaged from all audiotaped sessions available for that patient, were correlated with measures of therapeutic alliance using both the therapist and the patient version of the HAQ. Scores on the patient and therapist version of the HAQ were significantly associated, $r^2 = 0.36$, $p < .001$, with 91 pairs of ratings available from both patients and therapists. Patients' ratings of alliance ($n = 96$) correlated at $r^2 = 0.17$, $p = .10$, with overall CMCS scores, and $r^2 = 0.20$, $p = .06$ with General subscale and $r^2 = -0.06$, $p = .59$ with Draw subscale scores. The therapists' ratings of their alliance with each patient ($n = 92$) were significantly correlated with mean overall CMCS scores from sessions with that patient, $r^2 = 0.25$, $p < .05$ and $r^2 = 0.34$, $p < .001$ and $r^2 = 0.08$, $p = .47$, with General and Draw subscale scores, respectively.

3.6 Predictive validity

The longest duration of continuous abstinence achieved by patients randomized to CM was 4.7 ± 4.8 weeks versus 2.3 ± 3.0 weeks for patients randomized to standard care. After controlling for treatment group assignment, General subscale scores ($F(1, 99) = 3.83$, $p < .05$), but not Draw subscale scores ($F(1, 99) = 0.07$, $p < .79$), were significantly associated with longest duration of abstinence achieved.

4. Discussion

These data suggest the CMCS is a reliable and valid measure of CM administration. Inter-rater reliability for each item and the overall scale fell in the range of good to excellent (Cicchetti, 1994), and internal consistency was good (Cronbach, 1951). Scores on the CMCS rose as therapists were learning to administer CM, and CM sessions were readily distinguishable from non-CM sessions with respect to the overall score and 6 of the 12 items. Concurrent validity was established via significant correlations between therapists' rating of alliance with each of their patients and CMCS scores from those same therapist-patient interactions as assessed by

independent raters. Further, adherence to items on the General subscale of the CMCS were significantly associated with patients' treatment outcomes.

Therapists exhibited competent administration of CM even during the training phase. In part, the high levels of competence noted during initial attempts at CM delivery may relate to the overall training experience. Prior to administering CM with patients, therapists successfully practiced CM in at least three role play scenarios (e.g., when a negative sample was submitted, when a positive sample was submitted, and when a difficult situation arose such as self-reports of use being discordant with sample results). Because therapists were trained to criterion levels of "Adequate" or better on CMCS items prior to delivering CM to any patients, their initial ratings in CM competence were high in the training phase. Nevertheless, with continued practice and feedback, competence ratings significantly increased, especially with respect to items that assessed general skills common across therapeutic interventions. These data suggest that general therapy skills may have been enhanced via trainings and experience delivering CM.

High levels of competence in CM delivery persisted during the randomized phase, in which therapists audiotaped their interactions with patients but no further feedback was provided with respect to their performance. Thus, following adequate training, community-based clinicians competently administer CM, and competence in CM delivery appears to endure even without formal feedback. Nevertheless, data derived from this study do not speak to whether such levels of competence would continue if therapist-patient interactions were not audiotaped. Taping and rating of competence in terms of other aspects of counseling (e.g., standard care therapy) did not occur, so these data also do not address whether other aspects of counseling delivery improved.

Ratings of competence were highly and significantly associated with ratings of adherence on CMCS items, suggesting that these constructs are not independent during CM sessions. Similar effects have been reported with scales assessing adherence and competence to other interventions, including motivational enhancement and cognitive-behavioral therapies and drug abuse counseling (Barber et al., 1996b, 1997; Carroll et al., 2000; Martino et al., 2008), calling into question the need to evaluate both dimensions when assessing therapist performance. We elected to focus on competence ratings for several reasons. Competence presumes adherence so competence is a higher order skill, and studies show that competence is more closely related to patient outcomes than adherence (Barber et al., 1996a). Nevertheless, in certain situations and settings, adherence ratings may be preferable as they can be quickly measured and may not require audiotaping and independent evaluations by trained raters. Subsequent research evaluating self-reported adherence ratings along the dimensions tapped by the CMCS may be useful. In particular, following achievement of good competence in CM delivery, therapists may be able to self-monitor adherence, easing the burden associated with continued audiotaping and rating of sessions by independent raters.

As in studies of other psychotherapies, competent therapy delivery was positively related to other areas. CMCS overall scores and General subscale scores were associated with therapeutic alliance as assessed by the therapist and longest duration of abstinence achieved. Competence on the General subscale items was more closely aligned with alliance than competence on the Draw subscale, which contained three items assessing discussions of outcomes of biological testing and number of draws available. Although contingent provision of reinforcement is a key principle in CM and ought to be assessed for fidelity purposes, perhaps the straightforward nature of these discussions make them less associated with the therapeutic alliance and outcomes than the CMCS items that capture more general therapeutic skills.

Limitations of this study include the relatively small sample size of therapists, which obviated the ability to assess associations between therapist characteristics and competence. Although therapists in this study had a range of educational backgrounds, they may have been more highly educated and experienced than substance abuse counselors in other areas or settings, which may limit generalization of the findings. In addition, this study did not evaluate construct validity and discriminability of treatments via multiple groups profile analysis. Some studies have used this approach (Carroll et al., 2000) which allows for simultaneous evaluation of the magnitude of different sources of variance, such as therapist, time, site, rater, etc. This approach was not adapted because it requires a full set of data from all participants and a larger number of therapists than participated in the present study.

Although the 11 raters included in this study were readily able to achieve high levels of inter-rater reliability, this study did not evaluate methods for training raters. The CMCS assumes that raters are aware of the reinforcement schedules in effect, and it was developed to evaluate competence of CM delivery with respect to reinforcing abstinence using the prize-based CM system (Petry et al., 2000, 2004, 2005ab, 2007). Although items could be re-worded to capture other reinforcement systems, its reliability and validity has not been evaluated in other contexts.

Strengths of the study are the inclusion of community-based therapists from a fairly large number of clinics, spanning outpatient psychosocial clinics to methadone clinics. Over 1,500 audiotapes of interactions, from nearly 200 patients, were rated and analyzed. These characteristics of the study design enhance generalization of findings.

CM is a brief intervention that can have pronounced effects on behavior. On average, CM sessions were 6.5 ± 3.0 minutes in duration. CM increased durations of abstinence achieved by two-fold relative to standard care treatment conditions that did not include CM. Although it is not possible to distinguish whether receipt of contingent reinforcement per se or receipt of contingent reinforcement in the context of the interaction with the therapist was responsible for the improved outcome, results from this study find that competence on the CMCS General subscale was significantly related to abstinence, even after controlling for the effect of treatment group assignment. These data suggest that general competence is important for improving patient outcomes, effects that are noted with other therapies as well (Carroll et al., 2000). Importantly, CM can be added onto any number of other psychotherapeutic approaches, including standard care (as it was in the present study), community-reinforcement, motivational enhancement, and cognitive-behavioral therapy (e.g., Budney et al., 2006; Higgins et al., 2000; Rawson et al., 2002, 2006). Thus, CM is widely applicable, and the CMCS could be utilized, perhaps in conjunction with other process measures, to ensure appropriate delivery of the intended combined intervention (e.g., CM+CBT). Such combined interventions, especially when delivered competently, may assist in further reducing drug use and relapse rates.

The CMCS is the first known instrument of its kind to evaluate competence with respect to CM delivery, and it may be useful for training and monitoring competence in CM delivery as this empirically-based intervention becomes integrated in community settings. This therapy rating system has other potential applications as well, including evaluating contributions of specific therapeutic elements to outcomes, examining sources of therapist effects, and exploring how treatment implementation varies with patient characteristics, such as severity of dependence, motivation, and antisocial personality disorder. Analyses of therapist-patient interactions using the CMCS may assist in identifying therapist behaviors that may be particularly effective, which ultimately may be garnered to improve outcomes of patients treated with CM.

References

- Andrzejewski ME, Kirby KC, Morral AR, Iguchi MY. Technology transfer through performance management: The effects of graphical feedback and positive reinforcement on drug treatment counselors' behavior. *Drug Alcohol Depend* 2001;63:179–186. [PubMed: 11376922]
- Barber JP, Crits-Christoph P, Luborsky L. Effects of therapist adherence and competence on patient outcome in brief dynamic therapy. *J Consult Clin Psychol* 1996a;64:619–622. [PubMed: 8698958]
- Barber JP, Krakauer I, Calvo N, Badgio PC, Faude J. Measuring adherence and competence of dynamic therapists in the treatment of cocaine dependence. *J Psychother Pract Res* 1997;6:12–24. [PubMed: 9058557]
- Barber JP, Mercer D, Krakauer I, Calvo N. Development of an adherence/competence rating scale for individual drug counseling. *Drug Alcohol Depend* 1996b;43:125–132. [PubMed: 9023068]
- Budney AJ, Moore BA, Rocha HL, Higgins ST. Clinical trial of abstinence-based vouchers and cognitive-behavioral therapy for cannabis dependence. *J Consult Clin Psychol* 2006;74:307–316. [PubMed: 16649875]
- Cameron J, Ritter A. Contingency management: perspectives of Australian service providers. *Drug Alcohol Rev* 2007;26:183–189. [PubMed: 17364854]
- Carroll KM, Kadden R, Donovan D, Zweben A, Rounsaville BJ. Implementing treatment and protecting the validity of the independent variable in treatment matching studies. *J Stud Alcohol Suppl* 1994;12:149–155. [PubMed: 7722991]
- Carroll KM, Nich C, Sifry R, Nuro K, Frankforter T, Ball SA, Fenton L, Rounsaville BJ. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug Alcohol Depend* 2000;57:225–238. [PubMed: 10661673]
- Carroll, KM.; Nuro, KF. Yale Psychotherapy Development Center Training Series Number 1. New Haven, CT: 1997. The Technology Model: An introduction to psychotherapy research in substance abuse. Unpublished manuscript and videotape
- Chapman JE, Sheidow AJ, Henggeler SW, Halliday-Boykins C, Cunningham PB. Developing a measure of therapist adherence to contingency management: An application of the many-facet Rasch model. *J Child Adoles Substance Abuse* 2008;17:47–68.
- Cicchetti DV. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychol Assess* 1994;6:284–290.
- Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297–334.
- Dutra L, Stathopoulou G, Basden SL, Leyro TM, Powers MB, Otto MW. A meta-analytic review of psychosocial interventions for substance use disorders. *Am J Psychiatry* 2008;165:179–187. [PubMed: 18198270]
- Garcia-Rodriguez O, Secades-Villa R, Higgins ST, Fernandez-Hermida JR, Carballo JL, Errasti Perez JM, Al-halabi Diaz S. Effects of voucher-based intervention on abstinence and retention in an outpatient treatment for cocaine addiction: a randomized controlled trial. *Exp Clin Psychopharmacol* 2009;17:131–138. [PubMed: 19586227]
- Henggeler SW, Chapman JE, Rowland M, Halliday-Boykins C, Randall J, Shackelford J, Schoenwald SK. Statewide adoption and initial implementation of contingency management for substance-abusing adolescents. *J Consult Clin Psychol* 2008;76:556–567. [PubMed: 18665685]
- Higgins ST, Heil SH, Dantona R, Donham R, Matthews M, Badger GJ. Effects of varying the monetary value of voucher-based incentives on abstinence achieved during and following treatment among cocaine-dependent outpatients. *Addiction* 2007;102:271–281. [PubMed: 17222282]
- Higgins ST, Wong CJ, Badger GJ, Ogden DEH, Dantona RL. Contingency reinforcement increases cocaine abstinence during outpatient treatment and at 1 year follow-up. *J Consult Clin Psychol* 2000;68:64–72. [PubMed: 10710841]
- Kellogg SH, Burns M, Coleman P, Stitzer M, Wale JB, Kreek MJ. Something of value: The introduction of contingency management interventions into the New York City Health and Hospital Addiction Treatment Service. *J Subst Abuse Treat* 2005;28:57–65. [PubMed: 15723733]
- Lott DC, Jencius S. Effectiveness of very low-cost contingency management in a community adolescent treatment program. *Drug Alcohol Depend* 2009;102:162–165. [PubMed: 19250774]

- Luborsky L, Barber JP, Siqueland L, Johnson S, Najavits LM, Frank A, Daley D. The revised Helping Alliance Questionnaire (HAQ-II): Psychometric properties. *J Psychother Pract Res* 1996;5:260–271.
- Lussier JP, Heil S, Mongeon J, Badger G, Higgins ST. A meta-analysis of voucher-based reinforcement therapy for substance use disorders. *Addiction* 2006;101:192–203. [PubMed: 16445548]
- Martino S, Ball S, Nich C, Frankforter TL, Carroll KM. Community program therapist adherence and competence in motivational enhancement therapy. *Drug Alcohol Depend* 2008;96:37–48. [PubMed: 18328638]
- Miller WR, Yahne CE, Tonigan JS. Motivational interviewing in drug abuse services: A randomized trial. *J Consult Clin Psychol* 2003;71:754–763. [PubMed: 12924680]
- Moyers T, Martin T, Catley D, Harris KJ, Ahluwalia JS. Assessing the integrity of motivational interviewing interventions: Reliability of the motivational interviewing skills code. *Behav Cog Psychother* 2003;31:177–184.
- Peirce JM, Petry NM, Stitzer ML, Blaine J, Kellogg S, Satterfield F, Schwartz M, Krasnansky J, Pencer E, Silva-Vazquez L, Kirby KC, Royer-Malvestuto C, Roll JM, Cohen A, Copersino ML, Kolodner K, Li R. Effects of lower-cost incentives on stimulant abstinence in methadone maintenance treatment: A National Drug Abuse Treatment Clinical Trials Network study. *Arch Gen Psychiatry* 2006;63:201–208. [PubMed: 16461864]
- Petry NM. A comprehensive guide to the application of contingency management procedures in clinical settings. *Drug Alcohol Depend* 2000;58:27–33. [PubMed: 10669052]
- Petry NM, Alessi SM, Carroll KM, Hanson T, MacKinnon S, Rounsaville B, Sierra S. Contingency management treatments: Reinforcing abstinence versus adherence with goal-related activities. *J Consult Clin Psychol* 2006;74:592–601. [PubMed: 16822115]
- Petry NM, Alessi SM, Hanson T, Sierra S. Randomized trial of contingent prizes vs. vouchers in cocaine-using methadone patients. *J Consult Clin Psychol* 2007;75:983–991. [PubMed: 18085914]
- Petry NM, Alessi SM, Marx J, Austin M, Tardiff M. Vouchers versus prizes: Contingency management treatment of substance abusers in community settings. *J Consult Clin Psychol* 2005b;73:1005–1014. [PubMed: 16392974]
- Petry NM, Martin B, Cooney JL, Kranzler HR. Give them prizes and they will come: Contingency management for treatment of alcohol dependence. *J Consult Clin Psychol* 2000;68:250–257. [PubMed: 10780125]
- Petry NM, Peirce JM, Stitzer ML, Blaine J, Roll JM, Cohen A, Obert J, Killeen T, Saladin ME, Cowell M, Kirby KC, Sterling R, Royer-Malvestuto C, Hamilton J, Booth RE, Macdonald M, Liebert M, Rader L, Burns R, DiMaria J, Copersino M, Stabile P, Kolodner K, Li R. Effects of prize-based incentives on outcomes in stimulant abusers in outpatient psychosocial treatment program: A National Drug Abuse Treatment Clinical Trials Network Study. *Arch Gen Psychiatry* 2005a; 62:1148–1156. [PubMed: 16203960]
- Petry, NM.; Stitzer, ML. Contingency management: Using motivational incentives to improve drug abuse treatment Treatment Manual. Yale University Psychotherapy Development Center; West Haven, CT: 2003.
- Petry NM, Tedford J, Austin M, Nich C, Carroll KM, Rounsaville BJ. Prize reinforcement contingency management for treating cocaine users: How low can we go, and with whom? *Addiction* 2004;99:349–360. [PubMed: 14982548]
- Prendergast M, Podus D, Finney J, Greenwell L, Roll J. Contingency management for treatment of substance use disorders: A meta-analysis. *Addiction* 2006;101:1546–1560. [PubMed: 17034434]
- Rawson RA, Huber A, McCann M, Shoptaw S, Farabee D, Reiber C, Ling W. A comparison of contingency management and cognitive-behavioral approaches during methadone maintenance treatment for cocaine dependence. *Arch Gen Psychiatry* 2002;59:817–824. [PubMed: 12215081]
- Rawson RA, McCann MJ, Flammino F, Shoptaw S, Miotto K, Reiber C, Ling W. A comparison of contingency management and cognitive-behavioral approaches for stimulant-dependent individuals. *Addiction* 2006;101:267–274. [PubMed: 16445555]
- Roll JM, Higgins ST, Badger GJ. An experimental comparison of three different schedules of reinforcement of drug abstinence using cigarette smoking as an exemplar. *J Appl Behav Anal* 1996;29:495–505. [PubMed: 8995832]

- Shrout PE, Fleiss JL. Intraclass correlations: Uses in assessing rater reliabilities. *Psychological Bulletin* 1979;86:420–428. [PubMed: 18839484]
- Squires DD, Gumbley SJ, Storti S. Training substance abuse treatment organizations to adopt evidence-based practices: The Addiction Technology Transfer Center of New England Science to Service Laboratory. *J Subst Abuse Treat* 2008;34:293–301. [PubMed: 17600652]
- Waltz J, Addis ME, Koerner K, Jacobson NS. Testing the integrity of a psychotherapy protocol: Assessment of adherence and competence. *J Consult Clin Psychol* 1993;61:620–630. [PubMed: 8370857]

Table 1

Items on the Contingency Management Competence Scale and Intraclass correlation coefficients (ICC; n = 99 sessions)

<i>Item</i>	ICC
1. To what extent did the therapist discuss outcomes of urine and breath sample monitoring?	0.78
2. To what extent did the therapist state how many draws were earned at this session?	0.90
3. To what extent did the therapist state how many draws would be earned at the next session if the client were abstinent?	0.90
4. To what extent did the therapist assess the client's desire for items in the prize cabinet?	0.94
5. To what extent did the therapist discuss the client's self-report of substance use?	0.76
6. If the client self-reported substance use, to what extent did the therapist relate self-reports of substance use to objective indicators of substance use?	0.75
7. If the client self-reported substance use, to what extent did the therapist relate self-report of substance use to consequences of positive samples?	0.77
8. To what extent did the therapist compliment/praise client's efforts toward abstinence?	0.85
9. To what extent did the therapist communicate confidence that the client's efforts will yield success in the future?	0.80
10. General skillfulness (expertise, competence and commitment, engages client in discussion, interventions made at appropriate times—not missed or made too early).	0.77
11. Maintaining session structure (maintains session focus, sets appropriate tone and structure, appropriate level of therapist activity/directiveness, duration).	0.67
12. Empathy (conveys warmth and sensitivity, demonstrates genuine concern and a non-judgmental stance, understands/expresses clients' feelings and concerns).	0.80

Table 2

Principal components analysis. Values represent factor loadings.

<u>Item</u>	Excluding self-report items (n = 177 sessions)		With self-report items (n = 78 sessions)	
	<i>General factor</i>	<i>Draw factor</i>	<i>General factor</i>	<i>Draw factor</i>
Outcome of testing	.143	.544	.037	.851
Draws earned	-.071	.741	.139	.823
Draws possible at next session	.188	.739	.217	.457
Desire for prizes	.665	.142	.514	.314
Discussion of self-reports	--	--	.792	.171
Relate self-report to objective indicators	--	--	.669	.340
Relate self-reports to consequences of positive samples	--	--	.570	.532
Compliment/praise patient's efforts	.860	.127	.867	.073
Communicate confidence	.867	.065	.858	-.031
General skillfulness	.895	-.012	.781	.275
Maintain structure	.781	.302	.679	.494
Empathy	.885	.097	.827	.307

Table 3

Ratings during early and later training periods in contingency management delivery.

Variable	Early sessions	Late sessions	t (d.f.), p value
Overall score	5.9 ± 0.6	6.3 ± 0.4	t (34) = 2.75, p < .01
General subscale score	5.8 ± 0.7	6.1 ± 0.4	t (34) = 2.74, p < .01
Draw subscale score	6.4 ± 0.6	6.5 ± 0.6	t (34) = 1.19, p = .24
<u>Item</u>			
Outcome of testing	6.6 ± 0.6	6.8 ± 0.5	t (34) = 1.64, p = .11
Draws earned	6.6 ± 0.4	6.6 ± 0.8	t (34) = 0.01, p = .99
Draws possible at next session	5.9 ± 1.1	6.2 ± 1.1	t (34) = 1.09, p = .28
Desire for prizes	5.0 ± 1.4	5.6 ± 1.1	t (34) = 2.41, p < .05
Discussion of self-reports	5.5 ± 0.8	6.0 ± 1.4	t (11) = 1.03, p = .32
Relate self-report to objective indicators	5.5 ± 0.8	5.9 ± 1.2	t (10) = 0.98, p = .35
Relate self-reports to consequences of positive samples	5.5 ± 1.0	6.0 ± 0.9	t (11) = 1.77, p = .10
Compliment/praise patient's efforts	5.7 ± 1.1	6.3 ± 0.5	t (34) = 3.47, p < .001
Communicate confidence	5.5 ± 1.2	6.0 ± 0.6	t (34) = 2.74, p < .01
General skillfulness	5.9 ± 0.7	6.2 ± 0.5	t (34) = 2.10, p < .05
Maintain structure	6.0 ± 0.6	6.3 ± 0.5	t (34) = 2.60, p < .05
Empathy	6.3 ± 0.6	6.3 ± 0.5	t (34) = 0.10, p = .92

Values represent means and standard deviations of ratings of sessions (up to 5 for earlier and later sessions, and 342 sessions in total) for each of the 35 therapists during the training phase in contingency management delivery.

Table 4

Ratings from contingency management and standard care sessions. Values represent mean and standard deviations of session ratings (n = 804) for 103 patients randomized to conditions.

Variable	Standard Care Sessions	Contingency Management Sessions	t (d.f.), p value
Overall score	3.9 ± 0.8	5.7 ± 0.7	t (101) = 11.33, p < .001
General subscale score	4.0 ± 0.9	5.3 ± 0.9	t (101) = 6.91, p < .001
Draw subscale score	2.8 ± 0.3	6.2 ± 0.9	t (101) = 23.23, p < .001
<u>Item</u>			
Outcome of testing	6.4 ± 0.8	6.4 ± 0.8	t (101) = 0.26, p = .79
Draws earned	1.0 ± 0.0	6.4 ± 1.3	t (101) = 26.04, p < .001
Draws possible at next session	1.0 ± 0.3	5.7 ± 1.7	t (101) = 17.98, p < .001
Desire for prizes	1.0 ± 0.2	4.8 ± 1.3	t (101) = 17.76, p < .001
Discussion of self-reports	3.8 ± 2.4	4.8 ± 2.0	t (60) = 1.90, p = .06
Relate self-report to objective indicators	3.8 ± 2.2	4.8 ± 2.0	t (60) = 1.79, p = .08
Relate self-reports to consequences of positive samples	3.2 ± 2.2	5.0 ± 1.8	t (60) = 3.60, p < .001
Compliment/praise patient's efforts	4.9 ± 1.7	5.4 ± 0.9	t (101) = 2.06, p < .05
Communicate confidence	4.5 ± 1.8	5.4 ± 0.9	t (101) = 3.00, p < .01
General skillfulness	5.6 ± 1.0	5.8 ± 0.8	t (101) = 0.96, p = .34
Maintain structure	5.7 ± 0.9	5.8 ± 0.8	t (101) = 0.60, p = .55
Empathy	5.8 ± 0.9	5.8 ± 0.8	t (101) = 0.34, p = .74