Development and Psychometric Evaluation of the Gambling Treatment Outcome Monitoring System (GAMTOMS)

Randy Stinchfield, Ken C. Winters, Andria Botzet, Sarah Jerstad, and Jessie Breyer
University of Minnesota Medical School

This article describes the development and initial reliability and validity estimates of the Gambling Treatment Outcome Monitoring System (GAMTOMS), a multi-instrument and multidimensional outcome assessment battery designed to measure gambling treatment outcomes. Reliability methods included both internal consistency of scales and a 1-week test–retest temporal stability procedure. Validity was examined with procedures to estimate content, convergent, discriminant, predictive, and construct validity. Data were collected from 2 separate studies, 1 on the questionnaire version and 1 on the interview version. The questionnaire study included 46 female and 41 male gambling treatment clients and 22 female and 2 male nonclinical participants. The interview study included 88 female and 62 male gambling treatment clients and 16 female and 9 male nonclinical participants. The GAMTOMS was found to yield satisfactory estimates of internal consistency, and 1-week test–retest reliability and was also found to demonstrate satisfactory content, convergent, discriminant, predictive, and construct validity.

Keywords: gambling, treatment outcome assessment

With the expansion of legalized gambling, there has also come an increasing awareness of those gamblers who experience problems associated with their gambling. With this growing public awareness of pathological gambling (PG), there has been a concomitant growth in epidemiological surveys and clinical investigations. PG impacts multiple domains of the gambler's life, in-

cluding his or her emotional/mental health, marital relationship, family relationships, and employment. PG causes severe pain and suffering to the gambler and his or her significant others, as well as significant costs to society in terms of health care expenses and lost productivity (National Research Council, 1999). Since the first instrument was reported in the literature in 1975 as part of the Survey of American Gambling Attitudes and Behavior Study (Kallick, Suits, Dielman, & Hybels, 1979), at least 16 different problem gambling assessment instruments have been developed (National Research Council, 1999). Many of these measures have not been psychometrically evaluated, and others have received minimal evaluation (Stinchfield, Govoni, & Frisch, 2004). The exception is the South Oaks Gambling Screen (SOGS), which has been widely used in numerous epidemiological studies and has received extensive psychometric evaluation in a variety of populations and settings (Lesieur & Blume, 1987; National Research Council, 1999; Shaffer, Hall, & Vander Bilt, 1997; Stinchfield, 2002).

Related to instrumentation is an interest in measuring gambling treatment outcomes. Treatment providers, payers, clients, and policy makers are interested in the effectiveness of treatment and how treatment may be improved, both in terms of therapeutic effectiveness and cost effectiveness. There is a growing number of gambling treatment outcome studies, and this literature has been reviewed by several other investigators (Knapp & Lech, 1987; Murray, 1993; National Research Council, 1999; Petry & Armentano, 1999; Stinchfield & Winters, 2001; Viets & Miller, 1997; Walker, 1993). This literature does not indicate a standard set of outcome variables or instruments, nor is there agreement on how treatment success should be defined. Thus, no standard clinical or treatment outcome measure has been developed that has achieved recognition as a benchmark tool in the field. Gambling treatment investigators are left to either develop their own clinical and treatment outcome instruments or use psychometrically untested

ones. In sum, gambling treatment outcome research is hampered by a lack of psychometrically sound measures of treatment effectiveness.

In its report to the President of the United States and the Congress, the National Gambling Impact Study Commission (1999) specifically recommended that a treatment outcome mechanism be developed that compiles data on the efficacy of various treatment methods and services. Treatment attempts to reduce both the human and economic costs of this disorder, but in order to evaluate its effectiveness, it is necessary to have a reliable and valid treatment outcome assessment system. The Gambling Treatment Outcome Monitoring System (GAMTOMS) was originally developed in 1992 to evaluate gambling treatment in Minnesota (Stinchfield & Winters, 1996, 2001). The original GAMTOMS consisted of a battery of self-administered paper-and-pencil questionnaires. Due to counselor recommendations, the GAMTOMS was subsequently revised into a counselor-administered interview version. The purpose of the current research was to describe the development and initial reliability and validity of both the questionnaire and interview versions of the GAMTOMS.

Method

Participants

The study of the questionnaire version included a treatment-seeking sample ($n = 87$) of 46 women and 41 men recruited from two treatment programs in Minneapolis, Minnesota, and one program in Duluth, Minnesota. Ninety-nine clients were asked to participate, and 12 refused. Participants were recruited between November 2000 and August 2001. The nonclinical sample ($n = 24$), recruited for a validity procedure, included 22 women and 2 men recruited from clinic staff and significant others of the clinical sample. The treatment-seeking clinical sample ranged in age from 20 to 63 years, with a mean age of 42.4 ($SD = 10.3$) years; 53% were women, and 89% were White. More than half (54%) were not married, 25% were college graduates, 70% were employed full time, and the modal annual household income category was $30,000 to $40,000. Comparison of clinical and nonclinical samples on demographic variables included a between-groups t test on age, $t(109) = 2.3$, $p = .02$, where the clinical sample was older (M age $= 42.4$) than the nonclinical sample (M age $= 36.9$). There was a greater proportion of women in the nonclinical sample (95.7%) than in the clinical sample (52.9%), and a chi-square test was statistically significant, $\chi^2(1, N = 110) = 14.1$, $p < .001$; however, there was no significant difference for race (White vs. non-White), $\chi^2(1, N = 108) = 0.5$, $p = .48$.

The study of the interview version included a clinical sample ($n = 150$) of 88 women and 62 men recruited from the same two treatment programs in Minneapolis, MN, that were used in the questionnaire study. One hundred and sixty three clients were asked to participate, and 13 refused. Participants were recruited between June 2002 and June 2004. The nonclinical sample ($n = 25$), recruited for a validity procedure, included 16 women and 9 men who were family members of client participants. The age of the clinical sample ranged from 18 to 67 years with a mean of 42.7 years ($SD = 10.3$); 59% were women and 41% were men; 88% were White and 12% were non-White. More than half (54%) were not married, 26% were college graduates, 69% were employed full

time, and the modal annual household income category was $40,000 to $50,000. Comparison of clinical and nonclinical samples on demographic variables found no differences in age, $t(172) = 1.2$, $p = .22$; gender, $\chi^2(1, N = 174) = 0.25$, $p = .62$; and race (White vs. non-White), $\chi^2(1, N = 174) = 1.4$, $p = .23$.

Comparison of the questionnaire and interview clinical samples on demographic variables found no differences in age, $t(235) = 0.2$, $p = .85$; gender, $\chi^2(1, N = 237) = 0.75$, $p = .39$; race (White vs. non-White), $\chi^2(1, N = 232) = 1.35$, $p = .24$; marital status (married/partnered vs. single), $\chi^2(1, N = 236) = 0.48$, $p = .49$; education level, $\chi^2(6, N = 233) = 3.78$, $p = .71$; employment status (full time vs. other), $\chi^2(1, N = 236) = 0.13$, $p = .72$; and income, $\chi^2(7, N = 235) = 11.0$, $p = .14$.

Instruments

Development of the GAMTOMS. The GAMTOMS was developed in 1992 to evaluate the outcome of PG treatment in Minnesota (Stinchfield & Winters, 1996, 2001). Given the lack of a standardized gambling treatment outcome assessment battery, the research team reviewed the gambling treatment literature, examined treatment outcome assessment instruments, and developed an outcome assessment battery that largely drew from substance abuse treatment outcome assessment instruments. The development of the GAMTOMS was also aided by reviews from treatment providers and a panel of experts in the gambling field.

The GAMTOMS is a multidimensional battery of assessment tools that consists of admission, discharge, and follow-up instruments to be used with adult clients in treatment for PG. In addition to client assessment, the GAMTOMS also measures types and amount of treatment services administered to the client. The GAMTOMS includes the following instruments: (a) Gambling Treatment Admission Questionnaire/Interview (GTAQ/GTAI), (b) Gambling Treatment Discharge Questionnaire/Interview (GTDQ/GTDI), (c) Gambling Treatment Services Questionnaire (GTSQ), and (d) Gambling Treatment Follow-up Questionnaire/Interview (GTFQ/GTFI). This assessment system was designed to provide tools to evaluate treatment effectiveness in a manner that is practical, scientifically valid, and economical, while minimizing the intrusion and burden on treatment providers and clients. The GAMTOMS is an integrated system that includes assessment questionnaires and interviews, a user's manual, and scoring information. A copy of the GAMTOMS may be obtained from the first author. See the Appendix for content of GAMTOMS instruments at each measurement point.

GTAQ/GTAI. The GTAQ/GTAI consist of 162 items that are administered to the client at the time of admission to treatment. Both the self-administered paper-and-pencil GTAQ and the counselor-administered GTAI take about 45 to 60 min to complete. They are multidimensional instruments that measure the following domains: client demographics (5 min); clinical and treatment history (5 min); recent gambling behaviors (including gambling frequency for each game and gambling debt; 10 min); gambling problem severity (including the SOGS; Lesieur & Blume, 1987) and diagnostic criteria of PG (10 min); alcohol/drug use frequency (2 min); stage of change (SOC; 1 min); gambling-related financial (5 min), legal (5 min), and occupational (1 min) problems; and psychiatric symptoms (both lifetime and past 30 days; 5 min). *Diagnostic and Statistical Manual of Mental Disorders* (4th edi-

tion; *DSM–IV*) diagnostic criteria for PG (American Psychiatric Association, 1994) were paraphrased into 10 items and have demonstrated satisfactory reliability and validity (Stinchfield, 2003; Stinchfield, Govoni, & Frisch, 2005). Although the SOGS and *DSM–IV* have some content overlap, they are not identical; both are included in the GAMTOMS because they each assess unique content and both are commonly used for assessment, diagnosis, and outcome assessment (Stinchfield, 2002). The psychiatric symptom items were borrowed from the public-domain Addiction Severity Index (ASI; McLellan, Luborsky, Woody, & O'Brien, 1980) and include scales for both lifetime and past-30-days timeframes. The reading level of the GTAQ is sixth grade, computed using Fry's (1977) Readability Graph.

The GTAQ/GTAI was administered at the time of admission to treatment along with validity measures, including the timeline follow back (TLFB; Sobell, Sobell, Maisto, & Cooper, 1985) and BASIS-32 (Eisen, Dill, & Grob, 1994; Eisen, Grob, & Klein, 1986). The TLFB was revised for gambling (TLFB-RG) and assessed gambling behavior over the past 4 weeks. The TLFB was adapted for gambling by other investigators who reported satisfactory psychometric characteristics (Weinstock, Whelan, & Meyers, 2004). The BASIS-32 is a brief, standardized, 32-item instrument that measures the following five mental health domains: (a) relation to self and others, (b) depression and anxiety, (c) daily living and role functioning, (d) impulsive and addictive behavior, and (e) psychosis. The BASIS-32 also provides one total scale score. The BASIS-32 has demonstrated satisfactory reliability, both in terms of internal consistency and temporal stability. Internal consistency has been demonstrated with coefficient alphas ranging from .63 to .80 for the subscales and .89 for the total scale (Eisen et al., 1994). Temporal stability has been demonstrated with 2- to 3-day test–retest reliability coefficients ranging from .65 to .81 for the subscales. The BASIS-32 is a copyrighted instrument and the copyright holder, McLean Hospital in Belmont, Massachusetts, permits mental health care providers or facilities to use the instrument for the purpose of outcome assessment of their own patients or clients provided that McLean Hospital is clearly identified as the copyright holder (Eisen et al., 1986).

GTDQ/GTDI. Both versions of this instrument include 88 items that are administered to the client at discharge from treatment and take about 30 min to complete. This multidimensional instrument measures the following domains: recent gambling, TLFB-RG, SOC, efforts at recovery, psychiatric symptoms, treatment component helpfulness, and client satisfaction with treatment. The Client Satisfaction scale (Larsen, Attkisson, Hargreaves, & Nguyen, 1979; Nguyen, Attkisson, & Stegner, 1983) has been adapted for addressing the services of addictions treatment and measures the client's level of satisfaction with the treatment received. The reading level of the GTDQ is sixth grade, computed using Fry's (1977) Readability Graph.

GTSQ. The GTSQ is a summary of the client's treatment experience as reported by the treatment provider. The GTSQ measures the types and amount of treatment services received by the client during the course of treatment and was adapted from the Treatment Services Review (McLellan, Alterman, Cacciola, Metzger, & O'Brien, 1992). The GTSQ measures the following domains: assessment/diagnostic workup; number of individual, group, family, and marital therapy sessions; number of aftercare group sessions; admission and discharge dates; referral source;

admission and discharge status; length of treatment; health insurance and source of payment; diagnoses; and discharge referrals. The GTSQ is completed by the treatment provider at the time of the client's discharge and takes about 15 min to complete. The reading level of the GTSQ is eighth grade, computed using Fry's (1977) Readability Graph.

GTFQ/GTFI. Both versions of this follow-up instrument include 95 items administered to the client after a period of time following discharge from treatment (typically at 6 and/or 12 months after discharge); they take 30 to 45 min to complete. It is a multidimensional instrument that measures the following domains: client demographics; gambling behaviors (including gambling frequency for each game and gambling debt); SOC; gambling problem severity (*DSM–IV* and SOGS); alcohol/drug use frequency; posttreatment services utilization; gambling-related financial problems, illegal activities/arrests, and occupational problems; psychiatric symptoms; client satisfaction; and treatment outcome. The reading level of the GTFQ is sixth grade, computed using Fry's (1977) Readability Graph.

Procedures

The outpatient treatment programs from which clients were recruited were similar in terms of therapeutic orientations, treatment methods, and delivery of therapeutic services. These programs were eclectic but emphasized a Gamblers Anonymous 12-step approach that encouraged participation at meetings, taught Gamblers Anonymous principles, and facilitated the client's completion of the steps. The programs viewed PG as a primary, progressive, and chronic condition and included structured group therapy, family therapy, individual therapy, educational activities such as lectures and homework assignments, and financial counseling. Treatment was multimodal as noted previously, with the predominant treatment modality being group therapy. The therapeutic goal was abstinence from gambling, and treatment provided the setting and skills for clients to learn new ways of living without gambling. The frequency and length of treatment sessions varied somewhat across programs but was typically 2 to 3 therapy sessions per week for approximately 2 months, with aftercare lasting up to 1 year. Family involvement in treatment was encouraged.

Research staff introduced themselves to clients and described the purpose of the study and what would be asked of them. If the client consented to participate, he or she signed the consent form and was administered the GTAQ/GTAI within the first few days of treatment. The GTAQ/GTAI retest was administered 1 week after initial GTAQ/GTAI administration. The GTDQ/GTDI was administered at the time of discharge from treatment. Treatment providers completed the GTSQ at the time of the client's discharge from treatment. In the interview study, the GTFI was administered by research staff to the client via telephone 6 months after discharge from treatment. The chart review was conducted by research staff shortly after discharge from treatment. Clients signed a medical release form at the time of consent that gave the treatment provider permission to allow research staff access to clinical records.

All clients completed their admission assessment (GTAQ $n = 87$, GTAI $n = 150$), but fewer clients completed the 1-week retest (GTAQ $n = 72$, GTAI $n = 104$) and discharge assessment (GTDQ $n = 56$, GTDI $n = 88$). Follow-up data was collected in the interview study (GTFI $n = 54$) but not in the questionnaire study

due to the fact that the length of the questionnaire study was limited to just 1 year. Data were collected over the course of longitudinal studies; not all clients had complete data sets because some had not aged to discharge assessment (GTDQ $n = 4$, GTDI $n = 13$), and some had not aged to follow-up assessment at the end of the interview study (GTFI $n = 41$). Some clients could not be contacted for 1-week retest (GTAQ $n = 15$, GTAI $n = 46$), discharge (GTDQ $n = 27$, GTDI $n = 49$), or follow-up for the interview study (GTFI $n = 55$). Nonclinical samples were also recruited in both studies to provide validity information for the GTAQ/GTAI. Nonclinical samples included clients' significant others and clinic staff (GTAQ $n = 24$, GTAI $n = 25$). Treatment providers completed the GTSQ on 51 clients in the questionnaire study and 139 clients in the interview study. Research staff completed chart reviews on 51 clients in the questionnaire study and 91 clients in the interview study. Treatment staff did not know that their answers on the GTSQ would be compared to the chart reviews.

Psychometric Evaluation

Reliability. Two types of reliability were measured: temporal stability and internal consistency. Temporal stability was examined with a 1-week test–retest procedure of the GTAQ/GTAI (administered at admission and again 1 week later). In order to demonstrate satisfactory temporal stability, a test–retest correlation of $r = .60$ or higher needs to be obtained (Cicchetti, 1994). Temporal stability was measured with intraclass correlation (ICC) coefficients because Pearson product–moment correlations are not sensitive to changes in scale means and variance between test and retest and can be misleading. For example, if all clients increased their SOGS score by 10 points at retest, the Pearson product–moment correlation would be a perfect 1.0, in spite of the large change in score from test to retest. We therefore computed ICCs, which are sensitive to test–retest mean changes (Bartko, 1976).

The second type of reliability, *internal consistency*, refers to the concept that all items in a scale measure the same construct. Internal consistency analyses included principal components analysis (PCA) and Cronbach's (1951) coefficient alpha. PCA was used to determine whether scales were uni- or multifactorial and to identify which items were associated with each factor or scale (Tabachnick & Fidell, 2001). Item analyses, including item-to-total score correlations, were computed to determine if each item was related to the total scale score on all GAMTOMS scales. Internal consistency estimates were computed using Cronbach's alpha (1951), which is interpreted on a scale from 0 to 1. The higher the alpha, the higher the internal consistency of the scale. Nunnally (1978) suggested that scales have an alpha of .70 or greater to be considered as having a minimal level of internal consistency for research purposes and .90 or to be used to make important decisions. Internal consistency was computed on all scales at admission, discharge, and follow-up assessment.

Validity. Three basic types of validity—content, criterion-related, and construct—were examined for the GAMTOMS. Content validity was examined by having a panel of expert scientists and clinicians in each study review the GAMTOMS content and provide feedback as to whether the GAMTOMS included all of the important content domains that should be assessed in gambling treatment outcome. In the questionnaire study, three scientists and

three clinicians served as expert panel members who reviewed the content of the GAMTOMS and provided feedback. In the interview study, three individuals who were both scientists and expert clinicians reviewed the GAMTOMS. The expert panel gave written feedback as to content that should be revised or added to the GAMTOMS to fully represent the content domains that should be present in gambling treatment outcome. Criterion-related validity was examined with measures of both convergent and discriminant validity (Campbell & Fiske, 1959). Convergent validity refers to how well a scale correlates with an independent criterion or an existing measure of the same construct. If the scale obtains a moderate to high correlation with the independent measure, this validity coefficient provides evidence for the validity of this scale. To demonstrate evidence of convergent validity, the GAMTOMS scale should obtain moderate to high correlations ($r > .30$) with other measures of gambling problem severity (Cicchetti, 1994).

The GAMTOMS includes a number of measures of gambling problem severity, including gambling frequency, *DSM–IV* diagnostic criteria for PG, gambling-related financial problems, gambling-related illegal activities/arrests, and psychiatric symptoms. Alternative or validity measures of gambling frequency, gambling problem severity, and mental health included the TLFB-RG, SOGS, and BASIS-32, respectively. Validity of the *DSM–IV* diagnostic criteria was measured with correlations with the SOGS, Gambling-Related Financial Problems scale, and the BASIS-32 Impulsive/Addictive scale. Validity of gambling frequency was measured with correlations with the TLFB-RG. It was hypothesized that the GAMTOMS measures of highest level of gambling frequency as well as average level would be moderately correlated ($r > .30$) with TLFB-RG total days and total hours. The validity of the Financial Problems scale would be evident if it was correlated with the *DSM–IV* and the SOGS (which is heavily weighted with financial items). Validity of the Psychiatric Symptoms scale would be demonstrated if it was correlated with the BASIS-32 Depression/Anxiety scale and the BASIS-32 total score.

Testing criterion-related validity for the GTSQ included comparing information from the GTSQ with data from an independent chart review regarding treatment services, such as number and types of treatment sessions. The GTSQ was completed by the counselor in both the questionnaire and interview studies, so the data from both studies were combined for this comparison. The greater the agreement between the GTSQ and chart review, the greater the validity of the GTSQ.

Discriminant validity is present when a scale obtains low correlations with variables that are purportedly unrelated to the construct of interest. PG is purported to be unrelated to demographic variables such as gender, age, race, marital status, employment status, education, and income. The GAMTOMS gambling problem severity measures of *DSM–IV*, gambling frequency, financial problems, and illegal activities would demonstrate discriminant validity if there were low correlations ($r < .20$) with the aforementioned demographic variables.

The SOC item (McConnaughy, Prochaska, & Velicer, 1983; Prochaska & DiClemente, 1992) asks clients to select one of five response options (i.e., precontemplation, contemplation, decision making, action, and maintenance) that reflects their stage of readiness to make changes in their gambling behavior. It is purported that clients should move from lower stages to higher stages over the course of treatment and follow-up. Therefore, evidence for the

predictive validity of the SOC item would be demonstrated if it showed a change from a lower stage to a higher stage over the course of treatment and into follow-up. Further evidence of the predictive validity of the SOC item would be demonstrated if the admission SOC item was correlated with treatment completion and treatment outcome (as measured by gambling frequency and SOGS scores).

Construct validity of the GAMTOMS was examined by computing between-group tests of the clinical and nonclinical samples. It was predicted that the GAMTOMS measures of PG would imply group differences in test scores between clinical and nonclinical samples. Therefore, the GAMTOMS would demonstrate evidence of construct validity if it yielded statistically significant differences on measures of problem gambling severity between the clinical and nonclinical samples. Construct validity of the GAMTOMS Gambling Problem Severity scales would be evident if the scores in the nonclinical sample were significantly lower than the scores in the clinical (gambling treatment) sample.

Results

Internal Consistency Reliability

The PCA for the GTAQ/GTAI scales of *DSM–IV*, SOGS, Financial Problems, and Psychiatric Symptoms (lifetime) all yielded unifactorial solutions. The PCA of the Psychiatric Symptoms scale (past 30 days) yielded one dominant factor (eigenvalue = 2.5, accounting for 28% of scale variance) made up of five items and one other factor; however, the ASI scoring instructions were to score all nine items in one scale, and therefore we followed the ASI protocol. The PCA for the GTDQ scales of Client Effort at Recovery, Psychiatric Symptoms (past 30 days), Treatment Helpfulness, and Client Satisfaction all yielded unifactorial solutions. The PCA of the GTFI scales of *DSM–IV*, SOGS, Financial Prob-

lems, General Outcome, and Psychiatric Symptoms (past 30 days) all yielded unifactorial solutions.

Internal consistency coefficient alphas for the GAMTOMS scales are shown in Table 1 for both the questionnaire and interview versions. GTAQ/GTAI internal consistency coefficient alphas ranged from .59 to .79. The scales with coefficient alphas below the criterion of .70 were *DSM–IV* for both versions, and the interview versions of the SOGS and ASI psychiatric symptoms. Further analysis of these items and scales showed that the low coefficient alphas were due primarily to a homogeneous sample with low scale variances. For example, each *DSM–IV* item was endorsed by most of the sample, and the average score was 8. The GTDQ/GTDI coefficient alphas ranged from .56 to .94. The GTDI scales with coefficient alphas below the criterion of .70 were Client Efforts at Recovery, Treatment Component Helpfulness, and Psychiatric Symptoms (past 30 days). GTFI coefficient alphas ranged from .77 to .91, and all were above the criterion of .70.

Temporal Stability Reliability

Temporal stability was measured with a 1-week test–retest procedure for both the GTAQ and GTAI; test–retest correlations are shown in Table 2. ICC coefficients ranged from .46 to .99 for the GTAQ and from .34 to .99 for the GTAI, and most items/scales exhibited an ICC above the criterion of .70. The SOC item had low test–retest correlations in both the GTAQ (ICC = .46) and GTAI (ICC = .34). Two gambling frequency items on the GTAI had low test–retest correlations: games of skill (ICC = .46) and slot machine play (ICC = .67). The most stable variables were alcohol, tobacco, and marijuana use; legal status; and gambling debt.

Validity

In terms of content validity, the scientific/clinical panels concurred with most of the content included in the GAMTOMS and

Table 1

Internal Consistency Reliability of Scales in the Gambling Treatment Admission, Discharge, and Follow-Up Questionnaire/Interview

Instrument/scale	Questionnaire	Interview
Gambling Treatment Admission, Questionnaire (*n* = 87) / Interview (*n* = 150)		
SOGS (20 items)	.79	.66
DSM–IV Diagnostic Criteria (10 items)	.59	.59
Financial Problems (23 items)	.78	.71
Psychiatric Symptoms, lifetime (9 items)	—	.69
Psychiatric Symptoms, past 30 days (9 items)	—	.66
Gambling Treatment Discharge, Questionnaire (*n* = 56) / Interview (*n* = 86)		
Client Efforts at Recovery (9 items)	—	.56
Treatment Component Helpfulness (10 items)	.80	.67
Client Satisfaction (14 items)	.94	.90
Psychiatric Symptoms, past 30 days (9 items)	—	.65
Gambling Treatment Follow-Up Interview (*n* = 54)		
DSM–IV (10 items)		.90
SOGS (20 items)		.91
Financial Problems (23 items)		.84
Psychiatric Symptoms, past 30 days (9 items)		.79
General Outcome (15 items)		.89

Note. Internal consistency was measured by using Cronbach's coefficient alpha. SOGS = South Oaks Gambling Screen; *DSM–IV = Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Dashes indicate that the item was not in the questionnaire version but was added to the interview version.

Table 2
One-Week Test–Retest Reliability of Items/Scales From the Gambling Treatment Admission Questionnaire (GTAQ) and Interview (GTAI)

Item/scale	GTAQ ($n = 72$)			GTAI ($n = 104$)		
	Test	Retest		Test	Retest	
	M (SD)	M (SD)	ICC	M (SD)	M (SD)	ICC
Gambling frequency by game[a]						
Cards in a casino	1.2 (1.5)	1.2 (1.6)	.84	1.2 (1.5)	1.2 (1.6)	.86
Lottery	1.4 (1.4)	1.3 (1.3)	.87	1.4 (1.4)	1.3 (1.3)	.85
Pull tabs	0.7 (1.0)	0.8 (1.1)	.85	0.7 (1.0)	0.8 (1.1)	.79
Sporting event	0.3 (0.8)	0.3 (0.8)	.78	0.3 (0.8)	0.3 (0.8)	.89
Games of personal skill for money (such as bowling)	0.2 (0.6)	0.1 (0.4)	.81	0.2 (0.6)	0.1 (0.4)	.46
Slot machine, poker machine, or other gambling machine	2.8 (1.3)	2.6 (1.3)	.91	2.8 (1.3)	2.6 (1.3)	.67
Bingo	0.4 (0.8)	0.3 (0.7)	.82	0.4 (0.8)	0.3 (0.7)	.91
Betting on horses, dogs, or other animals	0.2 (0.6)	0.2 (0.6)	.71	0.2 (0.6)	0.2 (0.6)	.89
Dice games	0.2 (0.6)	0.2 (0.6)	.72	0.2 (0.6)	0.2 (0.6)	.89
Keno	0.2 (0.7)	0.2 (0.8)	.89	0.2 (0.7)	0.2 (0.8)	.81
Gambling on high-risk stock	0.02 (0.1)	0.03 (0.2)	.98	0.02 (0.1)	0.03 (0.2)	.86
Gambling problem severity scales/items						
DSM–IV Diagnostic Criteria (0–10)	8.3 (1.4)	8.3 (1.5)	.74	8.0 (1.7)	8.3 (1.7)	.78
SOGS (0–20)	13.9 (3.0)	14.1 (2.9)	.90	12.7 (3.2)	12.8 (3.3)	.87
Financial Problems (0–23)	4.7 (3.2)	4.9 (2.9)	.91	5.0 (3.0)	4.9 (3.1)	.90
Total gambling debt[b]	35,685	36,323	.97	35,330	38,946	.97
Gambling debt accumulated in past 12 months[b]	16,453	18,576	.93	16,374	16,057	.96
Illegal Activities (0–9)	1.1 (1.2)	1.0 (1.2)	.92	0.9 (0.9)	0.9 (0.9)	.86
Arrests (0–9)	0.2 (0.4)	0.2 (0.5)	.73	0.2 (0.4)	0.2 (0.5)	.84
Legal status (none vs. probation/parole/pending)	0.15 (0.36)	0.14 (0.35)	.93	0.1 (0.3)	0.1 (0.3)	.99
Gambling-related variables						
Tobacco[a]	4.0 (2.8)	4.0 (2.7)	.99	3.0 (2.4)	3.0 (2.4)	.99
Alcohol[a]	2.2 (1.8)	2.1 (1.9)	.94	1.3 (1.2)	1.3 (1.3)	.90
Marijuana[a]	0.3 (1.1)	0.3 (1.1)	.97	0.2 (0.8)	0.2 (0.6)	.96
Other drugs[a]	0.7 (1.8)	0.5 (1.5)	.79	0.1 (0.5)	0.1 (0.4)	.93
Drinking or using drugs while gambling[c]	0.4 (0.7)	0.5 (0.8)	.83	1.2 (0.5)	1.2 (0.6)	.76
Stage of change (1 item)[d]	3.4 (0.8)	3.7 (0.8)	.46	3.7 (0.7)	3.7 (0.7)	.34
Psychiatric Symptoms, lifetime (0–9)		—		4.6 (2.1)	4.6 (2.0)	.83
Psychiatric Symptoms, past 30 days (0–9)		—		3.3 (1.8)	3.5 (1.7)	.75

Note. ICC = intraclass correlation coefficient; *DSM–IV* = *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.); SOGS = South Oaks Gambling Screen. Dashes indicate that the item was not in the questionnaire version but was added to the interview version.
[a] *0 = never, 1 = less than once/month, 2 = 1 to 3 days/month, 3 = 1 to 2 days/week, 4 = 3 to 6 days/week, 5 = daily.*
[b] Gambling debt test–retest coefficients were attenuated by one client who reported no debt at intake (test) and a debt of $1,000,000 at retest; this outlier was removed from the computation of gambling debt test—retest.
[c] *1 = never/seldom, 2 = sometimes, 3 = often, 4 = always.*
[d] *1 = precontemplation, 2 = contemplation, 3 = decision making, 4 = action, 5 = maintenance.*

also gave suggestions for adding content that they deemed important to outcome assessment. This content was then added prior to the onset of the administration of the GAMTOMS. Revisions included the addition of a TLFB measure that we revised for gambling (TLFB-RG) as a more detailed measure of gambling activity. Another addition was the ASI psychiatric items to provide a brief assessment of psychiatric symptoms. A third addition was the inclusion of items to measure the client's efforts at recovery. Given that (a) the expert panels found that the GAMTOMS contained most of the content relevant for outcome assessment, and (b) we added some minor content elements upon the panel's recommendation, we can conclude that the GAMTOMS demonstrated satisfactory content validity.

Criterion-related validity was examined by computing correlation coefficients between GAMTOMS scales and alternative measures. The *DSM–IV* PG scale contains the 10 diagnostic criteria paraphrased into questions; therefore, it should have been (and

was) correlated with the SOGS ($r = .63$, $r = .62$), Financial Problems scale ($r = .49$, $r = .48$), and BASIS-32 Impulsive/Addictive scale ($r = .29$, $r = .27$), questionnaire and interview versions, respectively. The highest level of gambling frequency over the past 12 months was moderately correlated with the TLFB-RG measure of days gambling during the past 4 weeks ($r = .55$, $r = .47$). The Financial Problems scale, which consists of a list of 23 gambling-related financial problems, was moderately correlated with *DSM–IV* ($r = .49$, $r = .48$) and the SOGS ($r = .71$, $r = .57$). The Psychiatric Symptoms scale (lifetime) was correlated with BASIS-32 Depression/Anxiety scale ($r = .50$) and Total score ($r = .47$). The Psychiatric Symptoms scale (past 30 days; Interview version) was also correlated with BASIS-32 Depression/Anxiety Scale ($r = .61$) and the Total (BASIS-32) score ($r = .56$).

Criterion-related validity for the GTSQ was measured by comparing the GTSQ data, which was completed by the counselor, with an independent chart review completed by research staff

regarding discharge status and types and numbers of treatment services. The comparison of the GTSQ with the chart review is shown in Table 3. The comparison indicated high rates of agreement between the GTSQ and the chart review, providing support for the validity of the GTSQ.

Discriminant validity was measured with correlations between gambling problem severity measures (*DSM–IV*, SOGS, gambling frequency, gambling-related financial problems, and gambling-related illegal activities) and variables that are not thought to be related to problem gambling (namely, demographic variables). These data are shown in Table 4 for both the questionnaire and interview. Correlations with demographic variables (gender, age, race, marital status, employment status, educational level, and income) ranged from $r = .00$ to $r = .32$, and almost all correlations were below the criterion of $r < .20$. Correlation coefficients between gambling problem severity measures and psychiatric measures (ASI Psychiatric Composite scale and BASIS-32 scales) provided more of a discriminant validity challenge, and these correlations ranged from $r = .04$ to $r = .42$. *DSM–IV*, SOGS, and financial problems were more highly correlated with BASIS-32 scales and the ASI scale than were gambling frequency and illegal activities.

In terms of the construct validity of the SOC item, it was hypothesized that the SOC would improve from the time of admission to discharge and again at follow-up. The SOC mean score increased from 3.5 at admission to 3.8 at discharge for the questionnaire version; this change was statistically significant, $t(54) = 3.0$, $p = .004$. The mean score change from admission to discharge was 0.35. For the interview version, the SOC mean score increased from 3.6 at admission, to 4.1 at discharge, and to 4.3 at 6 months follow-up; this change was also statistically significant, $F(2, 49) = 16.5$, $p < .001$. The mean score change from admission to discharge was 0.52, from discharge to follow-up was 0.23, and from admission to follow-up was 0.77. This improvement in SOC over time provides evidence of the validity of the SOC item.

It was hypothesized that the SOC item would be correlated with treatment completion and treatment outcome. Admission and discharge SOC scores were not correlated with treatment completion ($r = .20$, $r = .19$, respectively). Admission and discharge SOC scores were moderately correlated with outcome at discharge in the form of gambling frequency ($r = .38$, $r = .54$, respectively). Admission, discharge, and follow-up SOC scores were moderately correlated with follow-up gambling frequency ($r = .41$, $r = .30$,

and $r = .52$, respectively) and with follow-up SOGS ($r = .43$, $r = .30$, and $r = .49$, respectively). The SOC item was correlated with outcome but not treatment completion.

It was hypothesized that the Efforts at Recovery scale would be correlated with outcome, and Efforts at Recovery scores were correlated with gambling frequency at discharge ($r = -.40$) and gambling frequency at follow-up ($r = -.40$). These negative correlations indicated that the greater the client score on the Efforts at Recovery scale, the lower the gambling frequency at discharge and at follow-up, thus providing preliminary evidence of the validity of this new scale.

Construct validity of the GAMTOMS was examined by computing between-group tests comparing clinical versus nonclinical groups. Table 5 shows the between-group comparisons. There were large and statistically significant differences between the clinical and nonclinical samples on all GAMTOMS scales and variables. Therefore, evidence of the construct validity of the GAMTOMS was demonstrated in that GAMTOMS scales were able to discriminate between clinical and nonclinical samples.

Discussion

The purpose of this study was to examine the psychometric properties of the questionnaire and interview versions of the GAMTOMS. The reliability and validity of the GAMTOMS were at levels considered favorable from the perspective of psychometric standards for behavioral instruments (Allen & Yen, 1979; American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 1999; Nunnally, 1978). GAMTOMS scales demonstrated satisfactory internal consistency, with coefficient alphas ranging from .54 to .94; the majority of the scales were at or above the criterion of .70. The GTAQ scales exhibited satisfactory temporal stability with 1-week test–retest ICCs ranging from .46 to .99, with the majority falling at or above the criterion of .70. Likewise, the interview version (GTAI) scales exhibited satisfactory temporal stability with 1-week test–retest ICCs ranging from .34 to .99, with the majority falling at or above the criterion of .70. The temporal stability data support the view that GAMTOMS scales are stable over a short period of time. The only exception was the SOC item, which had a low test–retest coefficient (ICC = .46 and ICC = .34 for the GTAQ and GTAI, respectively); this was likely due to the

Table 3

Agreement Between the Gambling Treatment Service Questionnaire (GTSQ) and Independent Chart Review (Interview Sample; n = 90)

Item	GTSQ	Chart review	Agreement
Was client admitted? (yes/no; %)	100.0	100.0	100
Was client on psychoactive medications? (yes/no; %)	53.6	45.2	82
Discharge status (complete vs. incomplete; %)	60.7	62.9	98
Did significant other participate in treatment? (yes/no; %)	44.0	33.3	89
Was client assigned a Gamblers Anonymous sponsor, case manager, or aftercare contact person? (yes/no; %)	62.4	63.5	94
Total number of sessions (*M*)	22.0	22.7	0.86[a]
To what extent did client complete the treatment plan? (5-point scale; *M*)	4.1	4.2	0.92[a]

[a] Pearson product–moment correlation.

Table 4

Discriminant Validity of Gambling Problem Severity Variables for Both the Questionnaire and Interview

	Gambling problem severity variables									
	Questionnaire (*n* = 87)					Interview (*n* = 150)				
Discriminant variable	*DSM–IV*	SOGS	Gambling Frequency	Financial Problems	Illegal Activities	*DSM–IV*	SOGS	Gambling Frequency	Financial Problems	Illegal Activities
Gender	.02	.03	.13	.11	.06	.32*	.17	.12	.07	.01
Age	.04	.19	.22	.03	.20	.09	.15	.06	.05	.08
Race	.04	.13	.11	.10	.09	.06	.02	.00	.15	.02
Marital status	.03	.02	.00	.20	.05	.08	.01	.11	.23*	.13
Employment status	.10	.15	.05	.13	.01	.09	.01	.01	.03	.13
Educational level	.05	.13	.03	.12	.13	.03	.03	.16	.09	.04
Annual household income	.02	.11	.10	.09	.05	.17	.02	.02	.13	.16
ASI Psychiatric composite score			—			.38*	.16	.08	.22*	.14
BASIS-32 Relation to Self and Others	.40*	.41*	.30*	.34*	.10	.34*	.28*	.17	.20	.05
BASIS-32 Depression and Anxiety	.39*	.37*	.25	.42*	.15	.42*	.25*	.19	.29*	.18
BASIS-32 Daily Living and Role Functioning	.38*	.38*	.30*	.41*	.15	.38*	.30*	.18	.25*	.10
BASIS-32 Impulsive and Addictive Behavior	.29*	.41*	.27	.34*	.26	.27*	.23*	.14	.10	.04
BASIS-32 Psychosis	.25	.34*	.14	.32*	.07	.25*	.24*	.18	.15	.09
BASIS-32 Total score	.40*	.42*	.30*	.42*	.16	.41*	.30*	.20	.25*	.13

Note. DSM–IV = Diagnostic and Statistical Manual of Mental Disorders (4th ed.) Diagnostic Criteria; SOGS = South Oaks Gambling Screen; Gambling Frequency = highest level of gambling frequency; ASI = Addiction Severity Index. The dash indicates that the item was not in the questionnaire version but was added to the interview version.
* $p < .01$.

fact that readiness for change was amenable to fluctuations in a 1-week period of treatment.

The content, criterion-related, and construct validity evidence support the validity of the GAMTOMS. The content validity procedure supports the position that (a) items and scales are representative of the content domains that the battery was intended to cover and (b) the GAMTOMS includes important content domains as identified by both expert clinicians and scientists. GAMTOMS scales were moderately correlated with other gambling problem severity measures (ranging from $r = .33$ to $r = .65$), and all were above the criterion of $r > .30$. To further build the nomological net (Cronbach & Meehl, 1955) around gambling problem severity, GAMTOMS scales exhibited satisfactory discriminant validity in that they were not correlated with variables that they should not have been related to (namely, demographic

variables) and they were either not correlated or only modestly correlated with other psychiatric symptoms and scales. In terms of construct validity, GAMTOMS scales/items were able to discriminate clinical from nonclinical samples and yielded statistically significant differences in both the questionnaire and interview versions. The SOC item exhibited both construct and predictive validity by showing statistically significant increases in SOC scores over the course of treatment and follow-up, and the admission SOC score was correlated with treatment outcome at discharge and follow-up. These reliability and validity findings provide sufficient evidence at this time to warrant recommending the use of the GAMTOMS as a clinical assessment and treatment outcome tool.

There are some items and scales in the GAMTOMS that did not meet the minimum criterion. These included the ASI Psychiatric

Table 5

Validity of Gambling Treatment Admission Questionnaire and Interview: Comparison of Clinical and Nonclinical Samples

	Questionnaire				Interview			
Variable	Clinical (*n* = 86)	Nonclinical (*n* = 24)	*t*	*p*	Clinical (*n* = 150)	Nonclinical (*n* = 25)	*t*	*p*
Highest level of gambling (0–5)	4.6 (1.0)	1.2 (1.1)	15.0	< .001	3.4 (1.0)	1.8 (0.9)	8.0	< .001
Average level of gambling (0–5)	1.0 (0.6)	0.2 (0.3)	6.1	< .001	0.7 (0.3)	0.4 (0.3)	3.3	.001
DSM–IV Diagnostic Criteria (0–10)	8.1 (1.7)	0.5 (0.2)	21.1	< .001	8.0 (1.8)	0.4 (1.3)	20.5	< .001
SOGS (0–20)	13.6 (3.5)	1.6 (1.8)	15.2	< .001	12.6 (3.0)	1.6 (2.2)	17.8	< .001
Financial Problems (0–23)	4.4 (3.1)	0.1 (0.5)	6.3	< .001	5.0 (3.1)	0.3 (1.2)	7.5	< .001
Illegal Activities (0–9)	1.1 (1.3)	0.1 (0.3)	3.6	.001	1.0 (1.0)	0.2 (0.8)	3.7	< .001
Psychiatric Symptoms, lifetime (0–9)		—			4.7 (2.0)	2.0 (2.3)	6.1	< .001
Psychiatric Symptoms, past 30 days (0–9)		—			3.5 (1.9)	1.1 (1.5)	6.0	< .001

Note. Data are means (standard deviations). *DSM-IV = Diagnostic and Statistical Manual of Mental Disorders* (4th ed.); SOGS = South Oaks Gambling Screen. Dashes indicate that the item was not in the questionnaire version but was added to the interview version.

Symptom scale, which had coefficient alphas below .70. These modest estimates of internal consistency were likely due to the fact that the ASI items do not measure one psychiatric domain, but rather sample across a number of psychiatric disorders. This was also evident from the multifactorial findings from the PCA of the ASI at admission. The ASI items can be interpreted on an individual item basis; however, the ASI instructions were to combine the items in a scale. Furthermore, the ASI scale score was correlated with the BASIS-32 Depression/Anxiety scale and Total scores, which provides evidence of the validity of the ASI scale. The SOC had low 1-week test–retest reliability. Because the week between tests occurred early during the course of treatment, it is likely that the client's commitment to change may have fluctuated over the short term. However, the SOC item did show a gradual but consistent improvement from admission to discharge and from discharge to follow-up; the SOC item was correlated with treatment outcome, and these findings support the validity of this measure.

Limitations and Future Research Directions

There are limitations to these studies that need to be noted. First, the data were based on self-report, and there was no objective verification of the accuracy of this information. However, efforts were made to enhance the validity of self-report by informing respondents that their names and answers would be kept confidential. Nevertheless, the data were dependent on self-report, and further research needs to be conducted on the validity of self-report of gambling behaviors, including the addition of validity scales such as social desirability. Second, the results are based on relatively small samples of clients, nonclinical participants, and treatment providers. Therefore, continuing research on the reliability and validity of the GAMTOMS with larger and more diverse samples is needed. Because these studies involved only a small number of treatment agencies, it is unknown how other types of treatment will affect the reliability and validity of GAMTOMS instruments. Future research should look at psychometric properties on the GAMTOMS collected at more treatment agencies using a variety of treatment approaches. Other future research goals for the GAMTOMS should include collection of additional data for groups that were not adequately represented in these studies, such as non-White clients. Future research could also assess the interrater reliability of the interview version of the GAMTOMS.

Additional studies should also address the predictive validity of the GAMTOMS. Do GAMTOMS admission data predict treatment completion and treatment outcome? Do different patterns of GAMTOMS results predict differential responses to different types of treatments? How do various patterns of GAMTOMS results relate to treatment readiness and attrition from treatment? Also of interest is whether GAMTOMS scores can identify classifiable profiles or client types. If so, a wealth of problem gambling assessment and treatment studies utilizing these GAMTOMS client types should follow.

The GAMTOMS has a brief assessment of comorbid psychiatric symptoms, and this is another limitation, given that pathological gamblers have fairly high rates of comorbid mental disorders (Specker, Carlson, Edmondson, Johnson, & Marcotte, 1996). Therefore, future research needs to look at the inclusion of more thorough assessment of comorbid psychiatric symptoms. Although

the BASIS-32 was used as a validity instrument in these studies, it provided important information on client mental health and should be considered an important addition to gambling admission and outcome assessments.

Another future research direction is to develop individual intake assessment reports as well as treatment program outcome reports from GAMTOMS results. This could include the development of recommendations for treatment programs regarding dropout rates, the comparison of program outcome to the outcome of other programs/approaches, and the identification of therapeutic agents such as the most and least helpful treatment components as rated by clients. In the current GAMTOMS, the treatment approach, therapeutic orientation, and treatment processes are largely ignored (other than measuring types of services provided), and therefore it would be helpful to improve the measurement of these important treatment domains.

The GAMTOMS is currently in two formats: (a) self-administered paper-and-pencil questionnaire and (b) counselor-administered interview. Although the results of these two studies found that both forms of administration yielded fairly similar estimates of reliability and validity, it would be important to conduct a direct comparison of both forms of administration to examine their respective strengths and weaknesses. Furthermore, there are advantages to computer-administered and Web-based versions, such as ease of data collection, data aggregation, computerized scoring and interpretation or scales, and report generation. Thus, another priority for future work will be to develop and field-test a computerized version of the GAMTOMS.

The development of the GAMTOMS represents an important advance in the field of problem gambling assessment and treatment outcome evaluation. The GAMTOMS includes multiple content domains that are important for assessment and treatment evaluation. The GAMTOMS answers important clinical questions, including the following: Who comes to treatment? Who drops out of treatment and why? What treatment components do clients rate as helpful? Are clients satisfied with treatment? The GAMTOMS also answers important outcome questions, including the following: Who improves? Who stays the same? Who gets worse? How many clients improve following treatment? Do clients abstain from or reduce gambling involvement during the course of treatment, and do they maintain those changes after treatment? What efforts do clients make toward recovery? Do clients function better in the areas of social and vocational responsibilities following treatment? How many clients participate in aftercare and other posttreatment recovery services?

In summary, the GAMTOMS demonstrates satisfactory reliability and validity, but further research needs to be conducted on larger and more diverse samples. The GAMTOMS will facilitate progress in the study of adult problem gambling and will substantially advance the clinician's ability to diagnose and treat adults who suffer from this problem. A sign of a maturing scientific field is that the instruments used to measure phenomena of interest become more precise, and it is the intent of this research to improve the assessment of PG and of gambling treatment outcomes.

References

Allen, M. J., & Yen, W. M. (1979). *Introduction to measurement theory.* Monterey, CA: Brooks/Cole.

American Educational Research Association, American Psychological Association, National Council on Measurement in Education. (1999). *Standards for educational and psychological testing.* Washington, DC: American Psychological Association.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Bartko, J. J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin, 83,* 762–765.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56,* 81–105.

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6,* 284–290.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302.

Eisen, S. V., Dill, D. L., & Grob, M. C. (1994). Reliability and validity of a brief patient-report instrument for psychiatric outcome evaluation. *Hospital and Community Psychiatry, 45,* 242–247.

Eisen, S. V., Grob, M. C., & Klein, A. A. (1986). BASIS: The development of a self-report measure for psychiatric inpatient evaluation. *Psychiatric Hospital, 17,* 165–171.

Fry, E. (1977). Fry's Readability Graph: Clarifications, validity, and extension to Level 17. *Journal of Reading, 21,* 242–253.

Kallick, M., Suits, D., Dielman, T., & Hybels, J. (1979). *A survey of American gambling attitudes and behavior.* Ann Arbor: University of Michigan, Institute for Social Research, Survey Research Center.

Knapp, T. J., & Lech, B. C. (1987). Pathological gambling: A review with recommendations. *Advances in Behavior Research and Therapy, 9,* 21–49.

Larsen, D., Attkisson, C., Hargreaves, W., & Nguyen, J. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2,* 197–207.

Lesieur, H. R., & Blume, S. B. (1987). The South Oaks Gambling Screen (SOGS): A new instrument for the identification of pathological gamblers. *American Journal of Psychiatry, 144,* 1184–1188.

McConnaughy, E. A., Prochaska, J. O., & Velicer, W. F. (1983). Stages of change in psychotherapy: Measurement and sample profiles. *Psychotherapy: Theory, Research and Practice, 20,* 368–375.

McLellan, A. T., Alterman, A., Cacciola, J., Metzger, D., & O'Brien, C. (1992). A new measure of substance abuse treatment: Initial studies of the Treatment Services Review. *Journal of Nervous and Mental Disease, 180,* 101–110.

McLellan, A. T., Luborsky, L., Woody, G., & O'Brien, C. (1980). An improved evaluation instrument for substance abuse patients: The Addiction Severity Index. *Journal of Nervous and Mental Disease, 168,* 26–33.

Murray, J. B. (1993). Review of research on pathological gambling. *Psychological Reports, 72,* 791–810.

National Gambling Impact Study Commission. (1999). *National Gambling Impact Study Commission report.* Washington, DC: Author.

National Research Council. (1999). *Pathological gambling: A critical review.* Washington, DC: National Academy Press.

Nguyen, T. D., Attkisson, C. C., & Stegner, B. L. (1983). Assessment of patient satisfaction: Development and refinement of a service evaluation questionnaire. *Evaluation and Program Planning, 6,* 299–314.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw Hill.

Petry, N. M., & Armentano, C. (1999). Prevalence, assessment, and treatment of pathological gambling: A review. *Psychiatric Services, 50,* 1021–1027.

Prochaska, J., & DiClemente, C. (1992). Stages of change in the modification of problem behaviors. In M. Hersen, R. Eisler, & P. Miller (Eds.), *Progress in behavior modification* (pp. 184–214). Sycamore, IL: Sycamore Press.

Shaffer, H. J., Hall, M. N., & Vander Bilt, J. (1997). *Estimating the prevalence of disordered gambling behavior in the United States and Canada: A meta-analysis.* Boston: Harvard Medical School, Division of Addictions.

Sobell, L. C., Sobell, M. B., Maisto, S. A., & Cooper, A. M. (1985). Time-line follow-back assessment method. In D. J. Lettieri, M. A. Sayers, & J. E. Nelson (Eds.), *NIAAA treatment handbook series: Vol. 2. Alcoholism treatment assessment research instruments* (DHHS Publication No. 85–1380, pp. 530–534). Washington, DC: National Institute on Alcoholism and Alcohol Abuse.

Specker, S., Carlson, G., Edmondson, K., Johnson, P., & Marcotte, M. (1996). Psychopathology in pathological gamblers seeking treatment. *Journal of Gambling Studies, 12,* 67–81.

Stinchfield, R. (2002). Reliability, validity, and classification accuracy of the South Oaks Gambling Screen (SOGS). *Addictive Behaviors, 27,* 1–19.

Stinchfield, R. (2003). Reliability, validity, and classification accuracy of a measure of DSM-IV diagnostic criteria for pathological gambling. *American Journal of Psychiatry, 160,* 180–182.

Stinchfield, R., Govoni, R., & Frisch, R. G. (2004). Screening and assessment instruments. In J. E. Grant & M. N. Potenza (Eds.), *Pathological gambling: A clinical guide to treatment* (pp. 207–231). Washington, DC: American Psychiatric Association.

Stinchfield, R., Govoni, R., & Frisch, R. G. (2005). DSM-IV diagnostic criteria for pathological gambling: Reliability, validity, and classification accuracy. *American Journal on Addictions, 14,* 73–82.

Stinchfield, R., & Winters, K. (1996). *Effectiveness of six state-supported compulsive gambling treatment programs in Minnesota.* Saint Paul: Minnesota Department of Human Services.

Stinchfield, R., & Winters, K. C. (2001). Outcome of Minnesota's gambling treatment programs. *Journal of Gambling Studies, 17,* 217–245.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.

Viets, V. C. L., & Miller, W. R. (1997). Treatment approaches for pathological gamblers. *Clinical Psychology Review, 17,* 689–702.

Walker, M. B. (1993). Treatment strategies for problem gambling: A review of effectiveness. In W. R. Eadington & J. A. Cornelius (Eds.), *Gambling behavior and problem gambling* (pp. 533–566). Reno, NV: William R. Eadington & Judy A. Cornelius.

Weinstock, J., Whelan, J. P., & Meyers, A. W. (2004). Behavioral assessment of gambling: An application of the timeline followback method. *Psychological Assessment, 16,* 72–80.

(*Appendix follows*)

Appendix

Gambling Treatment Outcome Monitoring System (GAMTOMS) Measurement Points and
Content of Questionnaires/Interviews

Admission GTAQ/GTAI	Discharge GTDQ/GTDI and GTSQ	Follow-Up GTFQ/GTFI
Demographics		Demographics
Clinical and treatment history		
Stage of change	Stage of change	Stage of change
Gambling frequency, TLFB-RG	Gambling frequency, TLFB-RG	Gambling frequency, TLFB-RG
Preferred game, gambling debt, age at onset		Preferred game, gambling debt, age at onset
Gambling problem severity: *DSM-IV* and SOGS		Gambling problem severity: *DSM-IV* and SOGS
Financial problems		Financial problems
Illegal activity/arrests		Illegal activity/arrests
Tobacco/alcohol/drug use frequency		Tobacco/alcohol/drug use frequency
Psychiatric symptoms and conflict with family	Psychiatric symptoms and conflict with family	Psychiatric symptoms and conflict with family
	Treatment services, length of treatment, discharge status, psychiatric diagnoses, psychotropic medications, referrals	Posttreatment services; Gamblers Anonymous participation
	Client effort at recovery, client satisfaction and treatment component helpfulness	
		General outcome (in terms of family and occupational functioning)

Note. GTAQ/GTAI = Gambling Treatment Admission Questionnaire/Interview; GTDQ/GTDI = Gambling Treatment Discharge Questionnaire/Interview; GTSQ = Gambling Treatment Services Questionnaire; GTFQ/GTFI = Gambling Treatment Follow-up Questionnaire/Interview; TLFB-RG = Timeline Follow Back Revised for Gambling; *DSM-IV = Diagnostic and Statistical Manual of Mental Disorders* (4th ed.) Diagnostic Criteria; SOGS = South Oaks Gambling Screen (Lesieur & Blume, 1987).