Effects of Interactive Voice Response Self-Monitoring on Natural Resolution of Drinking Problems: Utilization and Behavioral Economic Factors

JALIE A. TUCKER, PH.D., M.P.H.,^{a,*} DAVID L. ROTH, PH.D.,^b JIN HUANG, M.D., M.S.,^b M. SCOTT CRAWFORD, M.P.H.,^c AND CATHY A. SIMPSON, PH.D.^a

^aDepartment of Health Behavior, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama

^bDepartment of Biostatistics, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama

^cDepartment of Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, Alabama

ABSTRACT. Objective: Most problem drinkers do not seek help, and many recover on their own. A randomized controlled trial evaluated whether supportive interactive voice response (IVR) self-monitoring facilitated such “natural” resolutions. Based on behavioral economics, effects on drinking outcomes were hypothesized to vary with drinkers’ baseline “time horizons,” reflecting preferences among commodities of different value available over different delays and with their IVR utilization. **Method:** Recently resolved untreated problem drinkers were randomized to a 24-week IVR self-monitoring program ($n = 87$) or an assessment-only control condition ($n = 98$). Baseline interviews assessed outcome predictors including behavioral economic measures of reward preferences (delay discounting, pre-resolution monetary allocation to alcohol vs. savings). Six-month outcomes were categorized as resolved abstinent, resolved nonabstinent, unresolved, or missing. Complier average causal effect (CACE) models examined IVR self-monitoring

effects. **Results:** IVR self-monitoring compliers ($\geq 70\%$ scheduled calls completed) were older and had greater pre-resolution drinking control and lower discounting than noncompliers ($< 70\%$). A CACE model interaction showed that observed compliers in the IVR group with shorter time horizons (expressed by greater pre-resolution spending on alcohol than savings) were more likely to attain moderation than abstinent resolutions compared with predicted compliers in the control group with shorter time horizons and with all noncompliers. Intention-to-treat analytical models revealed no IVR-related effects. More balanced spending on savings versus alcohol predicted moderation in both approaches. **Conclusions:** IVR interventions should consider factors affecting IVR utilization and drinking outcomes, including person-specific behavioral economic variables. CACE models provide tools to evaluate interventions involving extended participation. (*J. Stud. Alcohol Drugs*, 73, 686–698, 2012)

MOST PROBLEM DRINKERS DO NOT SEEK alcohol-focused clinical treatment. Many resolve problems on their own, and such “natural” resolutions are more common than treatment-assisted resolutions (Klingemann and Sobell, 2007; Sobell et al., 1996). The minority who enter treatment tend to have more serious problems, typically warranting abstinence, whereas the untreated majority tend to have less severe problems and to prefer and have more success with moderation drinking (Grant, 1996). Public health strategies to reduce drinking problems in the population thus depend on facilitating low-risk drinking,

in addition to abstinence, and on developing a spectrum of services ranging from self-change to specialty treatment to match the problem drinking population’s diverse needs and preferences (Tucker and Simpson, 2011).

Whether or not assisted by treatment, persons in the early months of a recovery attempt have higher relapse risk before new sober behavior patterns stabilize. Offering daily support for early behavior change may promote and help consolidate behaviors essential to longer-term resolution stability. Alcoholics Anonymous (AA) likely functions in this way and is readily accessible. Developing similar “low threshold” professional alternatives would expand recovery support options, and applications using interactive voice response (IVR) systems are promising.

IVR systems provide a computerized telephone-based platform for many health-related applications (Abu-Hasaballah et al., 2007) ranging from appointment reminders to monitoring health status, behavior patterns, and surrounding contexts. IVR systems can be made available over long intervals for cost-effective risk monitoring, relapse prevention, and rapid treatment reentry when needed (Tucker and Simpson, 2011). IVR systems typically yield high-quality data compared with interviews and questionnaires.

Received: July 13, 2011. Revision: January 31, 2012.

This research was supported by National Institutes of Health/National Institute on Alcohol Abuse and Alcoholism (NIH/NIAAA) Grant 2 R01 AA008972 and NIH/NIAAA American Recovery and Reinvestment Act supplement 3 R01 AA008972-16S1. Portions of the research were presented at the June 2009 meeting of the Research Society on Alcoholism, San Diego, CA, and at the June 2011 Mechanisms of Behavior Change satellite session of the meeting of the Research Society on Alcoholism, Atlanta, GA.

*Correspondence may be sent to Jalie A. Tucker at the Department of Health Behavior, School of Public Health, University of Alabama at Birmingham, 1665 University Boulevard, Birmingham, AL 35294, or via email at: jtucker@uab.edu.

IVR systems seem well suited to provide ongoing accessible support to recovering drinkers by offering daily self-monitoring opportunities for drinking-related behaviors. Brief feedback and recovery-focused messages also can be delivered via IVR systems (Helzer et al., 2008). Daily self-monitoring via paper or electronic means has been shown to result in beneficial reductions in the recorded behaviors, including drinking (Miller and Wilbourne, 2002). To date, alcohol-focused controlled trials have used IVR self-monitoring in conjunction with professional care but not as a natural resolution aid. Results have been mixed. For example, Mundt et al. (2006) implemented an IVR self-monitoring relapse prevention strategy following residential treatment and found better 6-month drinking outcomes among IVR self-monitoring participants than controls. Similar reductions were observed among risky-drinking, HIV-positive primary care patients who received brief motivational interviewing followed by IVR self-monitoring (Aharonovich et al., 2006). A larger trial of primary care patients with alcohol problems (Helzer et al., 2008) found no clear evidence of IVR self-monitoring efficacy, although IVR self-monitoring combined with monthly feedback on daily drinking was superior at 6 months relative to no feedback.

Thus, as aftercare or treatment adjunct, IVR self-monitoring appears to help maintain treatment gains and reduce relapse. However, IVR system applications are not universally appealing, and use varies (Helzer et al., 2008; Mundt et al., 2006). For example, following residential alcohol treatment, withdrawals from aftercare were greater among patients provided with IVR self-monitoring access compared with those without access, and frequent reminder calls undermined, rather than facilitated, IVR system utilization (Mundt et al., 2006).

The present study extended this work by using IVR self-monitoring to support natural resolutions among community-dwelling problem drinkers who had recently stopped high-risk drinking without treatment and who were abstaining or engaging in low-risk drinking. Compared with treatment-seeking samples, natural recovery samples tend to have less serious problems and greater likelihood of achieving stable moderation (Sobell et al., 1996), making them suitable for a lower intensity intervention such as IVR self-monitoring. A randomized controlled trial compared 6-month outcomes of IVR-assisted and unaided natural resolution attempts. Drinking outcome predictors were assessed at baseline. Participants received IVR system access for 24 weeks to report daily drinking and contextual factors, set weekly drinking goals, and hear recovery-supportive messages.

The study was guided by behavioral economic theory and research (Bickel and Marsch, 2001), which views substance misuse as persistent preference for shorter-term rewards like drinking and devaluation of larger, delayed rewards like the longer-term benefits of sobriety. Applied to the resolu-

tion process, this suggests that problem drinkers who are more sensitive to longer-term contingencies, even when drinking heavily, should have a better prognosis. Our earlier prospective studies supported this hypothesis based on pre-resolution patterns of monetary allocation to commodities available over shorter or longer delays (Tucker et al., 2008, 2009). Greater relative allocation to voluntary savings than to purchasing alcohol predicted resolution stability over 1- to 2-year follow-ups, and resolutions involving stable low-risk drinking were associated with the most balanced pre-resolution spending patterns. Thus, active problem drinkers with longer "time horizons" over which they organized monetary allocation were more likely to achieve stable resolutions, particularly stable low-risk drinking.

The present study evaluated how IVR self-monitoring may help stabilize positive change during early recovery and interact with behavioral economic outcome predictors. If engaging in IVR self-monitoring enhances the salience of and connection with highly valued, but delayed, sobriety outcomes (Rachlin, 1995), then IVR self-monitoring should be most beneficial to recovering drinkers with shorter baseline time horizons, especially those attempting a moderation resolution. This hypothesis would be supported by an interaction between baseline behavioral economic measures of reward preferences and 6-month resolution status showing selective benefits of IVR self-monitoring for drinkers with shorter initial time horizons. This effect would essentially offset our prior finding that, in the absence of IVR self-monitoring, stable resolution, particularly stable moderation, is predicted by longer time horizons.

Because utilization of extended IVR self-monitoring typically varies from 0% to 100% call compliance, evaluating intervention effects on outcomes requires consideration of utilization differences. Complier average causal effect (CACE) models (Frangakis and Rubin, 2002; Little and Yau, 1998; Yau and Little, 2001) take such differences into account by identifying predictors of utilization or compliance and evaluating differences in outcomes between those who complied with an intervention and the controls who would have complied had it been assigned to them. Conventional intention-to-treat (ITT) approaches, in contrast, analyze data for all participants as a function of random assignment to conditions and ignore utilization differences and attrition effects. ITT models result in overly conservative intervention efficacy estimates when applied to extensive interventions like IVR self-monitoring, which require active participation over time.

Method

Sample characteristics

The research received institutional review board approval and coverage by a federal certificate of confidentiality. To

FIGURE 1. Consolidated Standards of Reporting Trials (CONSORT) flow diagram for 6-months outcomes. BAC = blood alcohol concentration; IVR = interactive voice response.

recruit untreated problem drinkers in early recovery—our population of interest—respondents to media advertisements in Alabama, Georgia, and Mississippi were screened using the Michigan Alcoholism Screening Test (Selzer, 1971), Alcohol Dependence Scale (Skinner and Horn, 1984), and Drinking Problems Scale (Cahalan, 1970). Eligibility criteria

were the following: (a) legal drinking age (≥ 21 years), (b) problem drinking history > 2 years ($M = 15.8$, $SD = 10.9$), (c) no current other drug misuse (except nicotine), and (d) cessation of high-risk drinking in the past 3–16 weeks without alcohol-focused interventions ($M = 2.28$ months resolved, $SD = 1.01$). Resolution onset was defined

as the most recent date that participants began abstaining or drinking in a nonproblem manner without further high-risk drinking (Sobell et al., 1996) for 3 weeks or more. A pattern of moderation, or sustained consumption of limited amounts of alcoholic beverages without any associated problems, was defined as (a) no dependence symptoms on the Alcohol Dependence Scale, (b) no alcohol-related negative consequences on the Drinking Problems Scale, and (c) no daily consumption above risky drinking thresholds (five or more standard drinks/day for men, four or more drinks/day for women; National Institute on Alcohol Abuse and Alcoholism [NIAAA], 2005).

Participants had not received treatment for alcohol problems or attended more than three AA meetings within a year of resolution onset. Limited earlier help-seeking was not exclusionary if it occurred 2 or more years before the current resolution and was limited to medication from a nonspecialist physician, religious/pastoral counseling, or lifetime attendance of nine or fewer AA meetings. The sample was largely

treatment-naïve (91%) and had attended three or fewer AA meetings (90%).

Figure 1 presents the Consolidated Standards of Reporting Trials statement (Schulz et al., 2010) through the 6-month follow-up, the focus of this study. Of 1,606 respondents screened between January 9, 2006, and February 28, 2010, 77.9% were ineligible (e.g., recent help-seeking, resolution duration > 16 weeks). Of 355 screened eligible, 30.1% could not be interviewed (e.g., no-shows), and 17.2% were refusals, deemed ineligible in baseline interviews, or excluded because of procedural errors (e.g., staff errors with randomization procedures). The remaining 187 respondents gave informed consent, were interviewed, and were randomized to the IVR self-monitoring ($n = 88$) or assessment-only control ($n = 99$) condition. Urn randomization (Stout et al., 1994) used gender and race as balancing factors because behavioral economic differences (e.g., income) were expected as a function of these variables, and balance across IVR and control groups was desired. One participant in each group

TABLE 1. Sample characteristics at initial assessment for participants randomized to intervention and control conditions

Variable	IVR SM Intervention (<i>n</i> = 87)			Assessment-only control (<i>n</i> = 98)			Combined sample (<i>n</i> = 185)			Test statistic	<i>p</i>
	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%		
Demographic characteristics											
Male			63.2			62.2			62.7	$\chi^2(1) = 0.02$	N.S.
Female			36.8			37.8			37.3		
White			49.4			58.2			54.0	$\chi^2(1) = 1.42$	N.S.
Other race/ethnicity			50.6			41.8			46.0		
Married			39.5			31.3			35.2	$\chi^2(1) = 1.37$	N.S.
Employed full/part time			47.7			46.9			47.3	$\chi^2(1) = 0.01$	N.S.
Age, in years	46.5	11.9		44.2	12.6		45.3	12.3		$F(1, 183) = 1.73$	N.S.
Years of education	14.0	2.9		13.8	2.4		13.9	2.7		$F(1, 182) = 0.40$	N.S.
History of drinking problems											
Problem duration, years	16.6	11.0		15.1	10.8		15.8	10.9		$F(1, 166) = 0.76$	N.S.
Alcohol Dependence Scale (47)	18.3	9.4		16.6	8.2		17.4	8.8		$F(1, 180) = 1.58$	N.S.
Michigan Alcoholism Screening Test (25)	12.8	3.8		11.6	4.2		12.1	4.1		$F(1, 182) = 3.77$	.054
Drinking Problems Scale (40)	17.9	9.2		16.4	8.9		17.1	9.1		$F(1, 182) = 1.27$	N.S.
Pre-resolution year drinking practices, TLFB											
% days well functioning	36.2	36.4		38.9	35.5		37.6	35.9		$F(1, 182) = 0.27$	N.S.
<i>M</i> quantity per drinking day (ml ethanol)	192.9	125.8		197.7	142.8		195.4	134.7		$F(1, 182) = 0.06$	N.S.
Behavioral economic variables											
Discounting log <i>k</i> parameter	-5.1	2.2		-5.3	2.8		-5.2	2.6		$F(1, 174) = 0.27$	N.S.
Total income, \$	48,935.1	55,638.0		57,599.4	99,115.7		53,502.7	81,407.0		$F(1, 182) = 0.52$	N.S.
Total expenditures, \$	51,276.7	50,108.2		51,449.2	64,236.6		51,367.7	57,832.5		$F(1, 182) = 0.00$	N.S.
Discretionary expenditures (DE)	14,593.0	11,299.6		15,104.2	13,362.5		14,862.5	12,399.1		$F(1, 182) = 0.08$	N.S.
Expenditures on alcohol (A)	2,942.2	2,444.5		2,733.2	2,779.0		2,832.0	2,621.2		$F(1, 182) = 0.29$	N.S.
Money saved (S)	1,247.7	4,532.2		1,788.5	6,443.4		1,532.8	5,613.0		$F(1, 182) = 0.42$	N.S.
ASDE index	0.23	0.29		0.19	0.29		0.21	0.29		$F(1, 182) = 0.85$	N.S.
% days with alcohol expenditures	51.7	36.3		46.0	35.6		48.7	36.0		$F(1, 182) = 1.15$	N.S.
<i>M</i> dollars spent per alcohol expenditure day	20.2	29.7		19.6	16.3		19.9	23.5		$F(1, 182) = 0.03$	N.S.

Notes: Maximum possible scores (for scaled questionnaires) are given in parentheses after the variable name. Higher Michigan Alcoholism Screening Test and Drinking Problem Scale scores indicate greater alcohol-related problems; higher Alcohol Dependence Scale scores indicate greater alcohol dependence levels. Reward k parameters and total income log-transformed for analysis. P values are from independent group t tests or chi-square analyses comparing the interactive voice response (IVR) self-monitoring (SM) and assessment-only control groups; N.S. = nonsignificant $p > .10$; TLFB = Timeline Followback interview; days well functioning = abstinent days plus drinking days below risk thresholds (<4 drinks for women, <5 drinks for men); ASDE = Alcohol-Savings Discretionary Expenditure.

was later excluded because of reporting accuracy concerns (see data quality checks). Table 1 shows baseline characteristics of the final sample of 185. Gender distribution was similar to the U.S. problem drinker population, and problem histories were consistent with typical outpatient treatment samples (Miller and Munoz, 2005).

Overview of procedures

Procedures were the same across groups except that IVR participants were trained after baseline data collection to use the system and then had access for 24 weeks. In-person interviews (1.5–3.0 hours) were conducted at baseline, 12 months, and 24 months. Telephone interviews (30–45 minutes) were conducted at 6 and 18 months. Baseline data were analyzed, along with 6-month outcome status codes; longer-term outcomes and IVR self-monitoring data will be presented in future reports. At baseline, participants gave informed consent, and breath testing verified sobriety. An expanded Timeline Followback (TLFB) interview assessed drinking-related (Sobell and Sobell, 1992) and behavioral economic (Vuchinich and Tucker, 1996) variables during the pre-resolution year up to the interview. Participants completed a computerized delay discounting task (Rachlin et al., 1991). They were compensated with university-issued Visa gift cards or checks (\$50 per in-person interview, \$15 per phone interview).

Measures

Measurement procedures are summarized. Detailed psychometric information can be found in prior publications (Tucker et al., 2007, 2008, 2009). Measures were administered at all in-person interviews; as noted, a subset was phone administered.

Drinking practices and spending on alcohol. Participants reported past-year drinking practices on the TLFB (Sobell and Sobell, 1992). Reports of beer, wine, or distilled spirits consumed (in ounces) were converted to milliliters of 190-proof ethanol for analysis. Participants also reported daily spending on alcoholic beverages, regardless of whether beverages were consumed.

Monetary allocation. Participants reported their income and expenditures for the same intervals using U.S. federal consumer expenditure commodity classes (Vuchinich and Tucker, 1996). Information was recorded from any financial records that participants provided. Financial assessments were completed using TLFB interviewing techniques. Income (in U.S. dollars) was reported by source (e.g., wages, loans). Expenditures were reported in three general categories, each with subcategories: housing (e.g., mortgage, utilities), consumable goods (e.g., food, alcohol), and other (e.g., entertainment, elective savings). Expenditures were then summed separately into obligatory and discretionary

categories. Obligatory expenditures were essential, largely fixed costs of living (e.g., housing, utilities). Discretionary expenditures were less essential, intermittently purchased commodities (e.g., entertainment, alcohol, elective savings).

The Alcohol-Savings Discretionary Expenditure (ASDE) index was computed for analysis as the proportion of discretionary expenditures summed over the pre-resolution year spent on alcohol minus the proportion of pre-resolution discretionary expenditures put into savings. ASDE values ranged from 1.0 to -1.0. Higher scores indicated proportionally more money spent on alcohol and less put into savings.

Delay discounting. The discounting task involved repeated choices between hypothetical monetary rewards of a smaller amount now and a larger amount later across a series of delays (e.g., “Would you choose \$500 now or \$1,000 in 6 months?”). A titration procedure determined an equivalence point at each delay, at which the value of the larger later and the smaller immediate amounts were rated as equally preferred. The equivalence points were fit to Mazur’s (1987) discounting equation to determine the discount rate (k parameter) for each participant: $v = A_i / (1 + kD_i)$, with v , A_i , and D_i representing the equivalent present reward value (e.g., \$500), the fixed delayed amount (e.g., \$1,000), and delay in weeks to the reward, respectively. The k parameter is an individual difference variable that generally ranges from 0 to 1 and is proportional to a participant’s rate of reward discounting. Higher k parameters indicate greater discounting and stronger preference for more immediate rewards, typical of substance misusers (Bickel and Marsch, 2001). Because k parameters tend to cluster at the lower end of the possible range, the analyses used the natural logarithm of k .

Follow-up assessment. Phone interviews at 6 months after IVR implementation inquired about drinking and help-seeking since the baseline interview. If drinking was reported, a TLFB was conducted, and alcohol-related problems were assessed with the Alcohol Dependence Scale and Drinking Problems Scale.

Checks on data quality. Participant reports of help-seeking history, resolution status, and drinking practices were verified using brief collateral phone interviews. Participant reliability checks were conducted when collaterals were unavailable. Collaterals were nominated by 82% of participants; one or more interviews were completed for 43% of participants with collaterals. Participant data were excluded when collaterals failed to verify participant reports pertaining to eligibility or follow-up status ($n = 1$) or when participants gave discrepant reports over time of the same information ($n = 1$). As a further data quality check, TLFB and IVR reports of drinking were compared for IVR participants who provided 10 or more days of time-matched reports. Reports of the percentage of drinking days and mean quantity consumed per drinking day did not differ significantly across measures, which correlated $r = .854$ for drinking days and $r = .656$ for mean quantities ($ps < .0001$). This

adds to evidence supporting the reliability of TLFB reports for assessing drinking outcomes (Sobell and Sobell, 1992).

Interactive voice response self-monitoring intervention

IVR participants were assigned a personal identification number to access the IVR system at any time. The system was programmed using commercial software (SmartQ Version 5[5.0.141], Telesage, Inc., Chapel Hill, NC). A daily survey assessed ounces of beer, wine, and distilled spirits consumed; use of other drugs to "get high"; and dollars spent on alcohol and other drugs during the preceding day (defined as the 24-hour period midnight-to-midnight yesterday). When no substance use was reported, participants answered questions about other prior-day activities to balance call duration.

Four once-a-week surveys on Mondays through Thursdays assessed other relevant domains (e.g., strategies used to avoid/limit drinking, activities paired with drinking). On Mondays, participants also received verbal feedback concerning whether they met last week's drinking goals and set new goals for the coming week. On Fridays, they could listen to 1- to 2-minute IVR-delivered education modules prepared from evidence-based cognitive-behavioral treatments (Sobell and Sobell, 1993) and supportive of a typical course of recovery. Early topics included goal setting and relapse prevention; later topics included supports for stable resolution (e.g., developing nondrinking social networks, money management). Monthly feedback letters summarized calling and drinking patterns and praised high IVR utilization (Helzer et al., 2008). Daily and weekly surveys averaged 4.6 and 2.3 minutes, respectively.

To promote IVR utilization, participants accrued points for daily calls that were modestly reimbursed using an "electronic bank" (Searles et al., 1995). They received \$0.50 minimum for each daily call and \$1.00 after seven consecutive calls; bonus payments stopped when a call was missed, were reinstated after another seven consecutive calls, and so forth. Payments by check were made quarterly during the IVR period. Participants received up to \$164.50 for 100% compliance; among those who called at least once, mean earnings were \$33.79 ($SD = \52.56). Automated reminder calls were not used because of potential threats to confidentiality and previous findings that reminder calls might actually undermine long-term IVR utilization (Mundt et al., 2006).

Resolution status at follow-up

Those who abstained or drank without problems throughout the 6-month post-resolution period were considered to have maintained stable resolutions, either resolved abstinent (RA, $n = 68$; 36.8%) or resolved nonabstinent (RNA or moderation outcome, $n = 34$; 18.4%). Those who engaged in any

drinking exceeding moderation criteria were considered to have unstable resolutions (UR, $n = 27$, 14.6%), even if they later abstained or moderated. Most reports came from the 6-month phone interviews; when participants could not be reached by phone at 6 months, outcome data were obtained from the 12-month interview. A fourth outcome category comprised those with missing or insufficient drinking data to determine a 6-month outcome ($n = 56$, 30.3%). Attrition is common in longitudinal alcohol research, and problem drinkers with missing values are often assumed to have characteristics predictive of a poor prognosis. This category allowed all eligible members of the sample to be included in the analytic models.

TLFB reports of drinking during the 6-month follow-up showed good separation between the RNA and UR groups. RNA participants were within NIAAA (2005) gender-adjusted guidelines for low-risk drinking (women: $M = 35.9$ ml of 190-proof ethanol per drinking day, $SD = 8.9$; men: $M = 43.6$ ml, $SD = 19.3$); UR participants were above thresholds for high-risk drinking (women: $M = 108.3$ ml, $SD = 68.1$; men: $M = 94.4$ ml, $SD = 79.2$). Mean and median drinking days for RNA participants were 9.29 ($SD = 12.94$) and 4.0 days, respectively. Corresponding values for UR participants were 41.65 ($SD = 44.32$) and 20.0 days.

Statistical analysis

For interventions like extended IVR self-monitoring where utilization is variable, the CACE approach focuses directly on participants who complied with an intervention. The standard ITT approach ignores differences in compliance and technically tests only the effects of intervention assignment, rather than efficacy of the intervention itself (Little and Yau, 1998; Yau and Little, 2001). CACE models compare participants who complied with the intervention to a statistically identified or "latent" subset of control participants who would have complied with the intervention had it been assigned to them. Information on the outcome variable and a set of pre-intervention covariates are used to estimate the latent class variable of compliance in the control group, and estimated compliers in the control group are compared with observed compliers in the intervention group.

For the CACE analyses, the distribution of daily IVR utilization was examined to select a cut point for assigning cases to dichotomous high- and low-compliance groups. Call frequency ranged from 0% to 100%. Seventeen participants did not start IVR self-monitoring. For the 70 callers, the mean and median percentage of call days were 46.9% ($SD = 38.2$) and 33.9%, respectively. The call frequency distribution indicated a notable line of fracture between 64% and 74% of calls completed; therefore, a 70% call rate was selected empirically as our compliance cut point. IVR participants were classified as either infrequent ($n = 59$, 67.8%) or frequent ($n = 28$, 32.2%) callers based on whether

TABLE 2. Sample characteristics at initial assessment as a function of outcome status at 6 months

Variable	Resolved nonabstinent (<i>n</i> = 34)			Resolved abstinent (<i>n</i> = 68)			Unstable resolutions (<i>n</i> = 27)			Missing outcome (<i>n</i> = 56)			Test statistic	<i>p</i>
	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%		
Demographic characteristics														
Male			67.7			63.2			51.9			64.3	$\chi^2(1) = 1.78$	N.S.
Female			32.3			36.8			48.1			35.7		
White			61.8			52.9			59.3			48.2	$\chi^2(1) = 1.91$	N.S.
Other race/ethnicity			38.2			47.1			40.7			51.8		
Married			45.5			35.8			26.9			32.1	$\chi^2(1) = 2.54$	N.S.
Employed full/part time			60.6			46.3			42.3			42.9	$\chi^2(1) = 3.08$	N.S.
Age, years	49.2 ^a	13.8		46.2 ^{a,b}	11.2		46.3 ^{a,b}	12.7		41.2 ^b	11.7		$\chi^2(1) = 3.57$	.015
Years of education	15.0 ^a	3.0		14.0 ^{a,b}	3.0		13.4 ^{a,b}	2.2		13.3 ^b	2.0		$F(3, 180) = 3.26$	.023
History of drinking problems													$F(3, 164) = 0.30$	N.S.
Problem duration, years	16.9	13.8		16.4	10.5		15.1	9.6		15.0	9.9		$F(3, 178) = 2.26$	.084
Alcohol Dependence Scale (47)	14.1	7.3		17.4	8.9		19.2	10.1		18.4	8.6			
Michigan Alcoholism Screening Test (25)	11.1	4.0		11.9	4.3		11.8	3.7		13.2	3.8		$F(3, 180) = 2.19$	.091
Drinking Problems Scale (40)	13.1 ^a	8.4		18.2 ^b	9.5		16.9 ^{a,b}	8.5		18.3 ^b	8.7		$F(3, 180) = 2.94$	.034
Pre-resolution year														
drinking practices, TLFB														
% days well functioning	53.6 ^a	41.2		33.8 ^b	33.2		36.5 ^{a,b}	33.6		33.2 ^b	34.8		$F(3, 180) = 2.89$	.037
<i>M</i> quantity/drinking day, ml ethanol	142.4	129.3		203.3	137.8		201.3	128.4		215.4	132.3		$F(3, 180) = 2.31$	.078
Behavioral economic variables														
Discounting log <i>k</i> parameter	-5.6	3.1		-5.3	2.2		-5.5	1.9		-4.7	2.9		$F(3, 172) = 0.97$	N.S.
Total income, \$	82,891.7 ^a	99,769.4		59,645.9 ^{a,b}	103,652.3		41,506.7 ^{a,b}	43,444.3		33,769.3 ^b	32,968.4		$F(3, 180) = 2.99$	.033
Total expenditures, \$	70,018.5	66,924.6		50,499.8	55,298.1		50,060.2	79,565.4		41,705.4	39,072.2		$F(3, 180) = 1.73$	N.S.
Discretionary expenditures (DE), \$	18,012.9	13,347.3		14,561.5	10,831.5		10,905.7	5,854.2		15,152.3	15,244.8		$F(3, 180) = 1.66$	N.S.
Alcohol expenditures (A), \$	2,061.8	2,584.1		3,233.0	2,669.9		2,632.7	2,247.4		2,905.2	2,698.9		$F(3, 180) = 1.59$	N.S.
Money saved (S), \$	1,753.6	5,530.6		1,803.2	5,089.8		391.2	1,899.7		1,600.4	7,248.2		$F(3, 180) = 0.43$	N.S.
ASDE index	0.07 ^a	0.25		0.22 ^{a,b}	0.30		0.29 ^b	0.31		0.25 ^b	0.27		$F(3, 180) = 3.78$	.012
% days with spending on alcohol	32.6 ^a	35.1		55.7 ^b	34.2		46.3 ^{a,b}	34.8		50.9 ^{a,b}	36.9		$F(3, 180) = 3.39$	.019
<i>M</i> dollars spent per alcohol spending day	18.7	13.4		20.8	22.2		18.4	14.3		20.2	32.2		$F(3, 180) = 0.10$	N.S.

Notes: Maximum possible scores (for scaled questionnaires) are given in parentheses after the variable name. Higher Michigan Alcoholism Screening Test and Drinking Problem Scale scores indicate greater alcohol-related problems; higher Alcohol Dependence Scale scores indicate greater alcohol dependence levels. Reward *k* parameters and total income are log-transformed for analysis. *P* values are from independent group *t* tests or chi-square analyses comparing the four drinking status groups. Means with different superscripts were significantly different in pairwise comparisons using Tukey's honestly significant difference test. Means for behavioral economic variables are in U.S. dollars except for delay discounting and Alcohol-Savings Discretionary Expenditure (ASDE) index, which is computed as a difference of proportions: ASDE = [(A / DE) - (S / DE)]. Values could range from 1.0 to -1.0 (1.0 = all DE for alcoholic beverages and -1.0 = all DE were for saving money; 0 = equal proportions of DE for alcohol and savings). Days well functioning = abstinent days plus drinking days below risk thresholds (<4 drinks for women, <5 drinks for men); N.S. = nonsignificant, *p* > .10; TLFB = Timeline Followback interview.

they completed less than 70% or 70% or more of IVR calls, respectively.

Preliminary CACE models evaluated potential predictors of IVR utilization status from among established drinking outcome predictors (e.g., gender, age, drinking practices, alcohol dependence, ASDE index) to estimate the latent class variable of compliance in the control group. Three variables consistently and independently predicted IVR utilization in these models: age, delay discounting ($\log k$), and percentage of “days well-functioning” during the pre-resolution year [(abstinent days + drinking days below NIAAA [2005] risk thresholds) / 365 $\times$ 100]. These variables were included as predictors of utilization in the main CACE models that evaluated intervention-related effects. Moderation outcomes (RNA) served as the referent group because of our interest in evaluating the benefits of IVR self-monitoring to facilitate stable moderation. Polytomous logistic regressions were used for the four-level outcome category (RA, UR, missing, RNA) at the 6-month follow-up, resulting in three pairwise outcome contrasts (RA vs. RNA, UR vs. RNA, missing vs. RNA). Group (IVR vs. control) was the only independent variable in the primary CACE model. Subsequent CACE models included the group contrast and one additional baseline covariate in each model. A final set of models examined interactions between group and baseline behavioral economic variables (ASDE, $\log k$) to test the hypothesis about the selective benefits of IVR self-monitoring for drinkers with shorter initial time horizons.

All analyses were estimated using Mplus Version 6.1 (Muthén and Muthén, 1998–2010). Because of occasional missing data on some baseline covariates (1.3% missing overall), the analyses were based on data after these missing values were imputed using the Markov chain Monte Carlo algorithm available in Mplus. Ten imputed data sets were created, and the reported model estimates were obtained from the analyses across the imputed data sets. There were no missing data for IVR self-monitoring group assignment.

To compare the sensitivity of the CACE results with the more conservative ITT approach, a comparable set of ITT models was conducted on the four-category outcome variable. Group (IVR vs. control) was the only independent variable in the main ITT model. Subsequent models mirrored the organization of the CACE models and included group and one other predictor (e.g., a baseline behavioral economic variable), and Group $\times$ Baseline Behavioral Economic variable (e.g., ASDE, $\log k$) interactions.

Results

Tables 1 and 2 present the sample characteristics at baseline and as a function of 6-month outcomes, respectively. Urn randomization successfully equated group baseline characteristics. At 6 months, several significant differences were found across the four outcome categories, indicating baseline covariate effects on outcomes. Measures with outcome associations at the $p < .10$ level were further examined as covariates in the CACE models.

TABLE 3. Odds ratios (ORs) of single covariate complier average causal effect (CACE) analyses

Predictor	Missing vs. Resolved nonabstinent			Resolved abstinent vs. Resolved nonabstinent			Unstable resolution vs. Resolved nonabstinent		
	OR	[95% CI]	<i>p</i>	OR	[95% CI]	<i>p</i>	OR	[95% CI]	<i>p</i>
Group, IVR or control	undefined			1.560	[0.378, 6.437]	.539	3.007	[0.415, 21.814]	.276
Demographic characteristics									
Age, years	0.974*	[0.926, 1.025]	.319	0.992	[0.952, 1.034]	.707	0.989	[0.944, 1.037]	.665
Years of education	0.812*	[0.682, 0.967]	.020	0.896	[0.775, 1.036]	.136	0.789	[0.645, 0.965]	.022
History of drinking problems									
Alcohol Dependence Scale	1.053*	[0.993, 1.117]	.082	1.039	[0.993, 1.087]	.101	1.061	[1.002, 1.123]	.040
Michigan Alcoholism									
Screening Test	1.132*	[0.987, 1.298]	.078	1.040	[0.934, 1.158]	.474	1.030	[0.914, 1.161]	.627
Drinking Problems Scale	0.066*	[0.997, 1.140]	.059	1.067	[1.008, 1.130]	.027	1.049	[0.983, 1.119]	.143
Pre-resolution year									
drinking practices, TLFB									
% days well functioning	0.987*	[0.972, 1.003]	.114	0.985	[0.972, 0.999]	.033	0.988	[0.973, 1.004]	.125
Behavioral economic variables									
Discounting $\log k$ parameter	1.012	[0.811, 1.263]	.918	0.986	[0.833, 1.167]	.868	0.961	[0.790, 1.169]	.686
Total income, \$	0.559*	[0.343, 0.913]	.020	0.782	[0.533, 1.148]	.209	0.696	[0.433, 1.119]	.135
ASDE index	7.925*	[1.256, 50.019]	.028	5.789	[1.308, 25.627]	.021	11.370	[1.742, 74.197]	.011
% days with spending									
on alcohol	1.013	[0.995, 1.031]	.145	1.019	[1.003, 1.035]	.012	1.010	[0.992, 1.028]	.272

Notes: The effect of interactive voice response (IVR) versus control was tested in the primary univariate CACE model with resolved nonabstinent as reference group and 70% IVR utilization cut point. The missing versus resolved nonabstinent contrast could not be estimated because too few were missing among the compliers; see text. Other predictors were added individually into the primary model, and effects of the predictors were tested. Only ORs of covariates with $p < .05$ are shown in the table. Age, reward delay discounting $\log k$ parameter, % well functioning days were used as compliance predictors in CACE models. Reward k parameters and total income log transformed for analysis. CI = confidence interval; TLFB = Timeline Followback; ASDE = Alcohol-Savings Discretionary Expenditure.

* p for group effect $< .05$.

Table 3 summarizes the CACE findings. Given that RNA (moderation) was the referent group for all contrasts, odds ratios (ORs) significantly less than 1.0 indicated that higher values on that predictor were associated with more frequent RNA outcomes compared with the other outcome category examined (RA, UR, or missing). ORs significantly greater than 1.0 indicated that higher predictor values were associated with less frequent RNA outcomes relative to the other outcome category examined. In the first model with no outcome covariates, greater IVR compliance was significantly and independently predicted by older age (estimate = 0.077, $SE = 0.020$, $p < .0001$), lower delay discounting (estimate = - 0.223, $SE = 0.109$, $p = .040$), and more days well functioning (estimate = 0.014, $SE = 0.007$, $p = .038$). These three compliance predictors were included in all subsequent CACE models along with different outcome predictors.

The first line of Table 3 reports the IVR versus control comparisons before accounting for any baseline covariate effects on outcome. Because the CACE analyses focused only on compliers, and because relatively few compliers had missing 6-month outcome codes, some missing versus RNA contrasts could not be statistically estimated. There were no IVR versus control effects on the RA versus RNA and UR versus RNA contrasts. The subsequent lines in Table 3 summarize the findings when a single outcome covariate was added to the basic model. For each of these models, the UR versus RNA and RA versus RNA contrasts were examined as a function of the IVR versus control condition.

No significant IVR main effects were found in any analysis either before or after adjusting for covariates (all $ps > .20$). The covariate effect on outcome tested whether a particular baseline variable was a significant predictor of the outcome contrast of interest across both the IVR and control groups.

The CACE models that examined whether baseline time horizon variables ($\log k$, ASDE index) modified the impact of the IVR self-monitoring condition on outcomes among the compliers yielded a statistically significant Group $\times$ ASDE index interaction effect on the RA versus RNA contrast (estimate = - 3.071, $SE = 1.492$, $p = .040$). The Group $\times$ ASDE interaction effects for the other two outcome contrasts and the Group $\times$ Log k interactions for all three outcome contrasts were not significant. Figure 2 shows the CACE model that contains the significant Group $\times$ ASDE interaction effect and presents the model estimates. The path from the intervention variable to the outcome contrast is modified both by compliance status (as in the typical main-effects-only CACE model) and by the ASDE. However, this ASDE effect-modification path is itself modified by the compliance variable in that it is forced to be 0 for the noncompliers and allowed to be estimated for the compliers. Therefore, the two-way Group $\times$ ASDE interaction effect is conceptually a three-way Compliance Status $\times$ Group $\times$ ASDE interaction effect because it is only estimated and found to be significant for the compliers.

Other interesting findings shown in Figure 2 include the predictors of compliance and the effect that compliance has on outcomes. Age and delay discounting ($\log k$) did not have significant direct effects on this outcome contrast. However,

FIGURE 2. Path diagram of the Complier Average Causal Effect model including an Interactive Voice Response (IVR) Group $\times$ Alcohol-Savings Discretionary Expenditure (ASDE) Index interaction predicting abstinence (resolved abstinent [RA]) versus moderation (resolved nonabstinent [RNA]) among the resolved. % days well functioning = percentage of days during the pre-resolution year below thresholds for excessive drinking, including abstinent days; log k = natural logarithm of the k parameter from the hyperbolic delayed reward discounting function.

$^{\wedge}p < .10$; $*p < .05$; $**p < .01$.

FIGURE 3. Complier Average Causal Effect model predicting the probability of moderation from the resolved abstinent versus resolved nonabstinent contrast as a function of the Interactive Voice Response (IVR) Group $\times$ Alcohol-Savings Discretionary Expenditure (ASDE) Index interaction effect. prob = probability.

they did have significant indirect or mediated effects through the compliance variable.

Figure 3 illustrates the significant Group $\times$ ASDE interaction by plotting the predicted probability of moderation among resolved participants (RNA or RA) for observed compliers in the IVR group, for predicted compliers in the control group, and for all noncompliers. Higher ASDE values on the x axis indicate relatively more pre-resolution spending on alcohol and less on savings, and higher values on the y axis indicate a higher probability of moderation compared with abstinent resolutions for the indicated group. Higher ASDE values had relatively little impact on resolution type among compliers in the IVR self-monitoring condition, whereas index values had a strong impact among predicted compliers in the control group. Among control compliers, higher ASDE values were associated with more abstinent and less moderation outcomes.

In contrast, the corresponding analyses conducted using ITT models did not show any IVR-related intervention effects. None of the three pairwise outcome contrasts showed a significant effect of the IVR intervention in the basic model (all $ps > .06$) or in models that adjusted for each baseline covariate (RA vs. RNA contrast: all $ps > .385$; UR vs. RNA contrast: all $ps > .077$). No statistically significant Group $\times$ ASDE or Group $\times$ Log k effects were found in the ITT analyses that examined whether the baseline time horizon variables modified the impact of the intervention condition on outcomes (all $ps > .40$). Thus, the significant Group $\times$ ASDE interaction from the CACE analysis was specific to

compliers and did not represent a more general IVR assignment effect.

Although no IVR-related effects were observed in the ITT models, consistent with the CACE models and our prior research, moderation outcomes were associated with higher education, more pre-resolution days well functioning, lower alcohol dependence, fewer alcohol-related life-health problems, lower ASDE values, and fewer pre-resolution days involving spending on alcoholic beverages. Among behavioral economic predictors, the RNA group showed more balanced pre-resolution discretionary spending on alcohol relative to voluntary savings compared with the other categories.

Discussion

This investigation evaluated an extensive IVR self-monitoring intervention for alcohol problems using an innovative data analytic method that separated predictors of IVR utilization from IVR-related effects on 6-month drinking outcomes. Using the CACE approach, selective evidence supported a beneficial effect of IVR self-monitoring for the subset of participants who were attempting moderation, had shorter baseline time horizons as measured by pre-resolution discretionary spending patterns, and used the IVR frequently. There were no unqualified IVR main effects.

The CACE models further identified distinct predictors of IVR utilization versus drinking outcomes at the end of the IVR interval. Compliant callers were more likely to be older

adults with lower delay discounting rates and better control of pre-resolution drinking. These are generally “good prognosis” drinkers for whom lower-intensity interventions such as IVR self-monitoring may be most appropriate.

In contrast to the CACE findings, no IVR-related intervention effects were found in conventional ITT models that do not adjust for treatment compliance. It is important to note that the different IVR-related results using the two analytic methods were observed within a context of replication of established predictors of moderation (Tucker et al., 2009). In both approaches, moderation outcomes were associated with lower ASDE values and less severe pre-resolution drinking practices, problems, and alcohol dependence.

The CACE results showed that, among the subset of compliant callers, frequent IVR use offset the association found in our prior studies that increasing ASDE values were associated with decreasing probability of moderation (Tucker et al., 2009). In the absence of an active self-monitoring intervention, recovering problem drinkers with greater pre-resolution spending on alcohol than voluntary savings are less likely to succeed in maintaining moderation than abstinent resolutions. The present findings suggested that this effect can be attenuated if such participants with higher pre-resolution ASDE values use the IVR self-monitoring intervention frequently. In other words, IVR self-monitoring selectively supports moderation if it is used.

CACE applications in alcohol and drug intervention research are a very recent development, few in number, and focused on interventions with adolescents (Connell, 2009; Stanger et al., 2011). Previous applications in other content areas are largely limited to examining intervention main effects on continuous outcome variables. We extended the CACE methodology in several ways by using an IVR system intervention with many utilization choice points over time, examining intervention effects on a categorical outcome, and evaluating intervention-modifying interaction effects. Because drinking outcomes were determined over the same time interval that IVR self-monitoring compliance was observed, some ambiguity remains concerning the causal linkages between utilization and outcomes. It was not possible to distinguish whether using IVR self-monitoring promotes moderation during early recovery for participants with shorter ASDE-defined time horizons or whether this subset of moderation drinkers chose to use IVR self-monitoring frequently to regulate their post-resolution drinking. It is possible that both causal mechanisms are occurring simultaneously, and analyses of longer term follow-up data may help clarify causal linkages among time horizons, IVR utilization, and outcome stability. Resolving the temporal sequences of such linkages is not necessary, however, to offer IVR self-monitoring to problem drinkers who want to drink moderately.

Understanding how the appeal of IVR self-monitoring interacts with drinker characteristics, preferences, and out-

comes is vital for marketing and targeting consumers who will use and benefit from it. The present findings show that age and delay discounting ($\log k$) did not have significant direct effects on resolution outcomes but did have significant indirect or mediated effects through the compliance variable (Figure 2). Thus, younger, more impulsive drinkers with higher baseline discounting were less likely to use—and therefore less likely to benefit from—IVR self-monitoring, whereas those with higher ASDE values who used the system frequently and remained resolved had a greater probability of stable moderation.

Properly marketed, IVR self-monitoring holds promise to expand the spectrum of services and reach the underserved majority who will not use clinical services. IVR-based interventions may contribute longer-term telehealth support with stepped-care potential (Tucker and Simpson, 2011). Our findings suggest offering IVR self-monitoring to natural recovery samples generally and specifically to more impulsive problem drinkers who want to drink moderately and may find abstinence-oriented programs unappealing. Although they may be less likely to use IVR self-monitoring and may need incentives (e.g., lotteries, vouchers) to motivate use, such drinkers are common in the general population and may benefit from the self-regulation inherent to self-monitoring.

The study has limitations, and the findings raise questions for future research. First, the follow-up rate was about 70%. This suboptimal rate was partially addressed by including a “missing” category as an outcome code along with the three resolution outcomes so that the analyses included all enrolled participants. Participants with missing outcomes did not exhibit, as is often presumed, baseline characteristics associated with poor prognosis more so than participants with RA or UR outcomes. Rather, the group differences were accounted for by the unique good prognosis characteristics of RNA participants.

Second, our CACE approach required compliance to be represented as a dichotomous latent class variable. Modeling decisions must therefore be based on careful examination of the compliance metric and distribution in light of conceptual and empirical considerations. We chose 70% based on the natural line of fracture in the cumulative IVR utilization distribution and because this utilization level reflects reasonable “treatment exposure.” However, other useful predictors of compliance may be identified if a polytomous or continuous measure was used in regular regression analyses.

Third, despite their common behavioral economic roots, the ASDE index and k parameter showed different relationships with compliance and drinking outcomes and were not correlated with each other ($r = .01$, *n.s.*). The diverging results support suggestions that behavioral impulsivity is a complex multidimensional construct (Reynolds et al., 2006). The ASDE index is a measure of actual behavior in the natural environment over long intervals and captures “molar” behavioral allocation over different delays and contexts. The

discounting task seems more of an analogue “state” measure of preference that may fluctuate over time (Bickel et al., 2010).

Overall, the study advances knowledge about natural self-change and how to promote it among the majority of problem drinkers who will not cross the clinical threshold for care. It also breaks new ground using methods appropriate for addressing the fuzzy boundary between efficacy and effectiveness research questions when the intervention is a form of extensive care for temporally distributed behavior problems that are affected by participants’ motivation, resources, and context (Tucker and Roth, 2006). When used to model intervention outcomes that depend in part on such extra-therapeutic variables, the CACE approach offers a sophisticated method for handling individual variability in treatment engagement and retention, which is common in behavior change programs. In addition to modeling compliance, influences on attrition can be investigated by including a missing outcome category. CACE models thus can make use of all available data from the to-be-treated sample. In the present study, person-specific behavioral economic variables predicted IVR utilization and drinking outcomes and modified the IVR-related effects on outcomes. The results indicate that IVR system interventions should consider factors that affect utilization and behavioral outcomes.

Acknowledgment

The authors thank Susan Chandler and Jessica Stringer for expert help with data coding and manuscript preparation.

References

- Abu-Hasaballah, K., James, A., & Aseltine, R. H., Jr. (2007). Lessons and pitfalls of interactive voice response in medical research. *Contemporary Clinical Trials*, 28, 593–602.
- Aharonovich, E., Hatzenbuehler, M. L., Johnston, B., O’Leary, A., Morgenstern, J., Wainberg, M. L., . . . Hasin, D. S. (2006). A low-cost, sustainable intervention for drinking reduction in the HIV primary care setting. *AIDS Care: Psychological and Socio-medical Aspects of AIDS/HIV*, 18, 561–568.
- Bickel, W. K., & Marsch, L. A. (2001). Toward a behavioral economic understanding of drug dependence: Delay discounting processes. *Addiction*, 96, 73–86.
- Bickel, W. K., Yi, R., Landes, R. D., Hill, P. F., & Baxter, C. (2010). Remember the future: Working memory training decreases delay discounting among stimulant addicts. *Biological Psychiatry*, 69, 260–265.
- Cahalan, D. (1970). *Problem drinkers: A national survey*. San Francisco, CA: Jossey-Bass.
- Connell, A. M. (2009). Employing complier average causal effect analytic methods to examine effects of randomized encouragement trials. *American Journal of Drug and Alcohol Abuse*, 35, 253–259.
- Frangakis, C. E., & Rubin, D. B. (2002). Principal stratification in causal inference. *Biometrics*, 58, 21–29.
- Grant, B. F. (1996). Toward an alcohol treatment model: A comparison of treated and untreated respondents with DSM-IV alcohol use disorders in the general population. *Alcoholism: Clinical and Experimental Research*, 20, 372–378.
- Helzer, J. E., Rose, G. L., Badger, G. J., Searles, J. S., Thomas, C. S., Lindberg, S. A., & Guth, S. (2008). Using interactive voice response to enhance brief alcohol intervention in primary care settings. *Journal of Studies on Alcohol and Drugs*, 69, 251–258.
- Klingemann, H., & Sobell, L. C. (Eds.). (2007). *Promoting self-change from addictive behaviors: Practical implications for policy, prevention, and treatment*. New York, NY: Springer.
- Little, R. J., & Yau, L. H. Y. (1998). Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin’s causal model. *Psychological Methods*, 3, 147–159.
- Mazur, J. (1987). An adjusting procedure for studying delayed reinforcement. In M. Commons, J. Mazur, J. A. Nevin, & H. Rachlin. (Eds.). *Quantitative analysis of behavior (Vol. 5): The effect of delay and of intervening events on reinforcement value*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Miller, W. R., & Munoz, R. F. (2005). *Controlling your drinking: Tools to make moderation work for you*. New York, NY: Guilford.
- Miller, W. R., & Wilbourne, P. L. (2002). Mesa Grande: A methodological analysis of clinical trials of treatments for alcohol use disorders. *Addiction*, 97, 265–277.
- Mundt, J. C., Moore, H. K., & Bean, P. (2006). An interactive voice response program to reduce drinking relapse: A feasibility study. *Journal of Substance Abuse Treatment*, 30, 21–29.
- Muthén, L. K., & Muthén, B. O. (1998–2010). *Mplus user’s guide* (6th ed.). Los Angeles, CA: Authors.
- National Institute on Alcohol Abuse and Alcoholism. (2005). *Helping patients who drink too much: A clinician’s guide*. Rockville, MD: Author.
- Rachlin, H. (1995). Self-control: Beyond commitment. *Behavioral and Brain Sciences*, 18, 109–121.
- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55, 233–244.
- Reynolds, B., Ortengren, A., Richards, J. B., & de Wit, H. (2006). Dimensions of impulsive behavior: Personality and behavioral measures. *Personality and Individual Differences*, 40, 305–315.
- Searles, J. S., Perrine, M. W., Mundt, J. C., & Helzer, J. E. (1995). Self-report of drinking using touch-tone telephone: Extending the limits of reliable daily contact. *Journal of Studies on Alcohol*, 56, 375–382.
- Selzer, M. L. (1971). The Michigan Alcoholism Screening Test: The quest for a new diagnostic instrument. *American Journal of Psychiatry*, 127, 1653–1658.
- Schulz, K. F., Altman, D. G., & Moher, D., & the CONSORT Group. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *Journal of Clinical Epidemiology*, 63, 834–840.
- Skinner, H. A., & Horn, J. L. (1984). *Alcohol Dependence Scale (ADS) user’s guide*. Toronto, Ontario: Addiction Research Foundation.
- Sobell, L. C., Cunningham, J. A., & Sobell, M. B. (1996). Recovery from alcohol problems with and without treatment: Prevalence in two population surveys. *American Journal of Public Health*, 86, 966–972.
- Sobell, L. C., & Sobell, M. B. (1992). Timeline Followback: A technique for assessing self-reported alcohol consumption. In R. Litten & J. Allen (Eds.), *Measuring alcohol consumption* (pp. 41–72). Totowa, NJ: Humana Press.
- Sobell, M. B., & Sobell, L. C. (1993). *Problem drinkers: Guided self-change treatment*. New York, NY: Guilford Press.
- Stanger, C., Ryan, S. R., Fu, H., & Budney, A. J. (2011). Parent training plus contingency management for substance abusing families: A Complier Average Causal Effects (CACE) analysis. *Drug and Alcohol Dependence*, 118, 119–126.
- Stout, R. L., Wirtz, P. W., Carbonari, J. P., & del Boca, F. K. (1994). Ensuring balanced distribution of prognostic factors in treatment outcome research. *Journal of Studies on Alcohol, Supplement 12*, 70–75.
- Tucker, J. A., Foushee, H. R., & Black, B. C. (2008). Behavioral economic

- analysis of natural resolution of drinking problems using IVR self-monitoring. *Experimental and Clinical Psychopharmacology*, 16, 332–340.
- Tucker, J. A., Foushee, H. R., Black, B. C., & Roth, D. L. (2007). Agreement between prospective interactive voice response self-monitoring and structured retrospective reports of drinking and contextual variables during natural resolution attempts. *Journal of Studies on Alcohol and Drugs*, 68, 538–542.
- Tucker, J. A., & Roth, D. L. (2006). Extending the evidence hierarchy to enhance evidence-based practice for substance use disorders. *Addiction*, 101, 918–932.
- Tucker, J. A., Roth, D. L., Vignolo, M. J., & Westfall, A. O. (2009). A behavioral economic reward index predicts drinking resolutions: Moderation revisited and compared with other outcomes. *Journal of Consulting and Clinical Psychology*, 77, 219–228.
- Tucker, J. A., & Simpson, C. A. (2011). The recovery spectrum: From self-change to seeking treatment. *Alcohol Research & Health*, 33, 371–379.
- Vuchinich, R. E., & Tucker, J. A. (1996). The molar context of alcohol abuse. In L. Green & J. H. Kagel (Eds.), *Advances in behavioral-economics (Vol. 3): Substance use and abuse* (pp. 133–162). Norwood, NJ: Ablex.
- Yau, L. H. Y., & Little, R. J. (2001). Inference for the complier-average causal effect from longitudinal data subject to noncompliance and missing data, with application to a job training assessment for the unemployed. *Journal of the American Statistical Association*, 96, 1232–1244.