Journal of Consulting and Clinical Psychology
2011, Vol. 79, No. 3, 330–341

0022-006X/11/$12.00 DOI: 10.1037/a0023498

Internet Therapy Versus Internet Self-Help Versus No Treatment for Problematic Alcohol Use: A Randomized Controlled Trial

Matthijs Blankers
Academic Medical Center, University of Amsterdam, and Arkin

Maarten W. J. Koeter
Academic Medical Center, University of Amsterdam

Gerard M. Schippers
Academic Medical Center, University of Amsterdam, and Arkin

Objective: Problematic alcohol use is the third leading contributor to the global burden of disease, partly because the majority of problem drinkers are not receiving treatment. Internet-based alcohol interventions attract an otherwise untreated population, but their effectiveness has not yet been established. The current study examined the effectiveness of Internet-based therapy (therapy alcohol online; TAO) and Internet-based self-help (self-help alcohol online; SAO) for problematic alcohol users. ***Method:*** Adult problem drinkers (n = 205; 51% female; mean age = 42 years; mean Alcohol Use Disorders Identification Test score = 20) were randomly assigned to TAO, SAO, or an untreated waiting-list control group (WL). Participants in the TAO arm received 7 individual text-based chat-therapy sessions. The TAO and SAO interventions were based on cognitive–behavioral therapy and motivational interviewing techniques. Assessments were given at baseline and 3 and 6 months after randomization. Primary outcome measures were alcohol consumption and treatment response. Secondary outcome measures included measures of quality-of-life. ***Results:*** Using generalized estimating equation regression models, intention-to-treat analyses demonstrated significant effects for TAO versus WL (p = .002) and for SAO versus WL (p = .03) on alcohol consumption at 3 months postrandomization. Differences between TAO and SAO were not significant at 3 months postrandomization (p = .11) but were significant at 6 months postrandomization (p = .03), with larger effects obtained for TAO. There was a similar pattern of results for treatment response and quality-of-life outcome measures. ***Conclusions:*** Results support the effectiveness of cognitive–behavioral therapy/motivational interviewing Internet-based therapy and Internet-based self-help for problematic alcohol users. At 6 months postrandomization, Internet-based therapy led to better results than Internet-based self-help.

Keywords: randomized controlled trial, alcohol, Internet, self-help, therapy

This article presents the outcome of a three-arm open randomized controlled trial of Internet-based treatment for problematic alcohol users. Problematic alcohol use is the third leading contributor to the global burden of disease (Rehm, Taylor, & Room, 2006) and is considered to be the main cause of 3.8% of global mortality (Rehm et al., 2009). The consequences of problematic alcohol use impose large costs on societies. Estimates of the associated economic costs show that more than one percent of the gross domestic product in high- and middle-income countries is spent on cost attributable to alcohol consumption (Casswell & Thamarangsi, 2009).

The magnitude of this burden partly results from the fact that the majority of people suffering from alcohol-use disorders are not receiving any form of formal treatment. With an estimated 78.1% of cases untreated, alcohol abuse and dependence remain largely unaddressed, even though effective treatments exist (Kohn, Saxena, Levav, & Saraceno, 2004). In particular, those with recent, less severe alcohol abuse receive little attention. For example, in the Netherlands, where the present study was conducted, 85.9% of all alcohol abusers and 52.1% of alcohol-dependent people have not received any form of treatment during the last 12 months. For people who manage to stay gainfully employed, the proportion receiving treatment is even lower (de Graaf, ten Have, & van Dorsselaer, 2010). This difference between the prevalence of a

This article was published Online First May 2, 2011.

Matthijs Blankers and Gerard M. Schippers, Amsterdam Institute for Addiction Research, Department of Psychiatry, Academic Medical Center, University of Amsterdam, and Arkin, Amsterdam, the Netherlands; Maarten W. J. Koeter, Amsterdam Institute for Addiction Research, Department of Psychiatry, Academic Medical Center, University of Amsterdam.

The randomized controlled trial reported in this article was funded by Grant 31160006 from the ZonMw Addiction II Program (Risk Behavior and Dependency). The study was conducted in collaboration with the Jellinek Clinic, which is a division of Arkin, an Amsterdam-based public mental health and addiction treatment center. Arkin supported the trial and facilitated its development. The two interventions evaluated in this study were developed at the Jellinek Clinic. None of the funders of the project had any role in the design of the trial, selection of the measurement instruments, or preparation of the paper. We thank W. Miles Cox for editing the manuscript and Suzan C. C. Oudejans for her assistance with data collection. The trial is registered in the Dutch Trial Register (Cochrane Collaboration) and is traceable as NTR-TC1155.

Correspondence concerning this article should be addressed to Matthijs Blankers, Amsterdam Institute for Addiction Research, Department of Psychiatry, Room PA3.224, P.O. Box 22660, Academic Medical Center, University of Amsterdam, 1100 DD Amsterdam, The Netherlands. E-mail: m.blankers@amc.uva.nl/mblankers@gmail.com

disorder and the proportion of individuals affected by the disorder who are treated is referred to as the treatment gap. For problematic alcohol use, numerous reasons for the treatment gap have been suggested, including stigmatization (Room, 2005), restricted opening hours and availability of services, costs related to restricted insurance coverage, or a lack of motivation to undergo treatment. If problematic alcohol use in society is to be reduced, the treatment gap must be reduced (Kohn et al., 2004).

Among the possibilities for bridging this treatment gap is the development and use of innovative treatment options. These options should preferably focus on a currently insufficiently addressed population and should be effective in reducing problematic alcohol use at the lowest possible cost. One of these options is to use Internet-based interventions. Internet-based interventions are seen as attractive to drinkers with relatively mild alcohol-related difficulties, who constitute a "hidden" population (Blankers, Kerssemakers, Schramade, Nabitz, & Schippers, 2008; Cunningham, Wild, Cordingley, van Mierlo, & Humphreys, 2009; Postel, de Jong, & de Haan, 2005; Riper et al., 2008). For these individuals, an Internet-based intervention is often the first intervention that they attempt for resolving their drinking problems (Blankers et al., 2008). Internet-based interventions provide easy access to professional support in a secure, private environment (Tate & Zabinski, 2004) and with minimal cost for the problem drinker. Two such interventions were evaluated in the present study.

Although a wide spectrum of Internet-based interventions addressing problematic alcohol use is available, three core modalities are discernible: (a) self-assessment with automatized personalized feedback (e.g., Sinadinovic, Berman, Hasson, & Wennberg, 2010); (b) multisession, Internet-based self-help behavior change (e.g., Cunningham et al., 2009; Riper et al., 2008); and (c) individual therapy that uses interaction over the Internet. Self-assessment is the least intensive modality. The general aim is to provide the person in a single session with insight into his or her drinking behavior through automatized personalized feedback and comparison of his or her own behavior with drinking norms. Studies of Internet-based self-assessment for reducing problematic alcohol use have shown positive results (Bewick, Trusler, Mulhern, Barkham, & Hill, 2008; Riper et al., 2008; Williams, Herman-Stahl, Calvin, Pemberton, & Bradshaw, 2009), although effect sizes are generally small (Cohen's $d < 0.25$; Sinadinovic et al., 2010).

Internet-based self-help for problematic alcohol users is more intensive and is usually extended over multiple treatment sessions. Most self-help interventions are modular by design and are based on cognitive–behavioral therapy (CBT) and/or motivational interviewing (MI) techniques. This combination is widely used for the treatment of alcohol-use problems and is currently the method of choice for Internet-based self-help interventions for alcohol use (e.g., Cunningham et al., 2009; Hester & Delaney, 1997; Linke, Brown, & Wallace, 2004; Riper et al., 2008). Evidence for the effectiveness of Internet-based self-help for problematic alcohol users is accumulating; effect sizes are small to moderate (e.g., Cohen's $d = 0.40$; Riper et al., 2008). A shortcoming of current Internet-based self-help programs is that they are not tailored to the individual's needs. This could also explain low adherence rates; often, less than 50% of the initial participants complete a multisession self-help intervention (Christensen & Mackinnon, 2006; Eysenbach, 2005).

The third and most extensive intervention modality is individual Internet-based therapy. Internet-based therapy incorporates nonautomatized, personally tailored feedback with interaction with a therapist. In general, Internet-based therapy can better meet the drinker's expectations and encourage greater self-disclosure through interaction with the therapist and therapist-led guidance and advice. This interaction between participant and therapist can be either synchronous or asynchronous. In synchronous interaction, contact is in real time and is direct (e.g., through text-based chat), but contact by (Internet) telephone and videoconferencing can also be considered synchronous. In asynchronous interactions, the client and the therapist do not interact with each other in the present moment; rather, they correspond via successive text or e-mail messages.

Evidence for the effectiveness of Internet-based therapy for problematic alcohol use based on synchronous text interaction has until now been lacking. The study reported here is the first one to test the effectiveness of this kind of Internet-based therapy in a randomized controlled trial (RCT). In the last several years, the results of two randomized trials of Internet-based self-help for problem drinkers have been published. Bewick, Trusler, Barkham, et al. (2008) reviewed the evidence on the effectiveness of electronic screening and brief intervention for alcohol use and concluded that it was inconsistent. Kypri, Langley, Saunders, Cashell-Smith, and Herbison (2008) found that Internet-based alcohol screening and self-help in a primary care setting reduced hazardous drinking and that the effect lasted 12 months. Riper et al. (2008) published the results of an RCT of Internet-based self-help compared to an untreated control group and found an effect size of $d = 0.40$ after 6 months. In their RCT of an Internet-based self-help intervention, Cunningham et al. (2009) showed that problem drinkers in the intervention group drank 15.1 ($SD = 11.2$) standard drinks per week 3 months after baseline as compared to 18.4 ($SD = 12.3$) drinks per week in the no-intervention group, yielding an effect size of $d = 0.28$. Based on Spek et al.'s (2007) meta-analysis of Internet-based interventions for anxiety and depression, different effects for Internet-based self-help and Internet therapy can be expected. Spek et al. found that interventions that included therapist support yielded large effect sizes ($d = 0.75$–1.24), but interventions without therapist support had small-to-medium effect sizes ($d = 0.08$–0.44). Accordingly, in the present study a larger effect was expected for Internet-based therapy than for Internet-based self-help.

This study compared the effectiveness of Internet-based therapy (therapy alcohol online; TAO) with Internet-based self-help (self-help alcohol online; SAO) for problematic alcohol users. In a three-arm open RCT, TAO and SAO were evaluated against an untreated waiting list control group (WL). Both SAO and TAO are based on a CBT/MI conventional treatment manual, which is widely used in the Netherlands (de Wildt, 2000). In this study, problematic alcohol use was defined as reporting current drinking of more than 14 standard drinks while obtaining a score of 8 or above on the Alcohol Use Disorders Identification Test (AUDIT; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). Standard drinks in this study contain 10 g of ethanol (the European standard; in the United States and Canada, 13.6 g per standard drink is the convention). It was hypothesized that 3 months after participants had been randomly assigned to one of the interventions, TAO participants would have reduced their alcohol consumption more

and would have a better quality of life than either SAO or WL participants. It was also expected that SAO participants would have a larger reduction in alcohol consumption and better quality of life than WL participants. Finally, it was expected that 6 months after randomization, TAO participants would have maintained the improvements more than SAO participants (Blankers, Koeter, & Schippers, 2009).

Method

Participants

Participants were recruited through the website of Jellinek/Arkin, the collaborating substance abuse treatment center (SATC), which has 650,000 visitors annually. A section of the website containing all relevant information for interested participants was developed for the study. The stated goal of the interventions was "to reduce your alcohol intake or quit drinking." Website visitors who expressed an interest in Internet-based interventions for problematic alcohol users were referred to the pages with information about the study. There they could complete a screening instrument to determine whether they met the inclusion criteria. To be included, applicants had (a) to be between 18 and 65 years old, (b) to be a resident of the Netherlands with health care insurance, (c) to have Internet access at home, (d) to obtain a score greater than 8 on the AUDIT, and (e) to report drinking an average of more than 14 standard drinks per week. Potential participants were excluded if they had had prior substance abuse treatment, a history of alcohol delirium or a drug overdose, a severe coronary or intestinal disease, schizophrenia, epilepsy, or suicidal tendencies in the last 12 months; if they used cocaine or amphetamine for more than 4 days of during the last month or used cannabis for more than 9 days during the last month; or if they planned to travel extensively outside the Netherlands during the first 6 months of the study. These criteria were based on those used in a previous study (Riper et al., 2008) and were intended to correspond to those used for low-intensity outpatient treatment at the participating SATC.

Eligible participants who provided informed consent were randomly allocated in a 1/1/1 ratio to one of the three trial arms: TAO, SAO, or WL. Restricted randomization by minimization was used in order to prevent imbalances among the trial arms. Three variables that were prognostic of outcome were selected, and their variance among the trial arms was minimized: sex, AUDIT composite score, and years of alcohol problems. In a meta-analysis, Jarvis (1992) found sex differences in treatment outcome for alcohol dependence. Spek et al. (2007) and Riper et al. (2008) also found sex differences for Internet-based interventions. Duration of the alcohol problem and its severity at the start of an intervention have also consistently been found to predict treatment outcome (Bodin & Romelsjö, 2007; Conigrave, Saunders, & Reznik, 1995; McLellan et al., 1994; Moyer, Finney, Swearingen, & Vergun, 2002) and were, therefore, minimized. Participants were allocated to the arms with the Pocock and Simon (1975) variance method without factor weighting but with a computer-generated random component (chance of allocation to the optimal arm was 0.6). Allocation was based on the outcome of this minimization procedure, which was executed immediately after the baseline assessment had been completed. In the remainder of this article, this procedure is referred to as randomization. All participants received an automated e-mail message and automated feedback through the study's website immediately after the allocation. Allocation procedures (i.e., randomization, allocation, informing the participant of his or her group assignment) were automated and server based and involved no interaction with the participants. As a result, it was possible to conceal participants' allocation in advance from themselves, the research assistants, and the therapists.

A power analysis, which was based on an analysis of variance design using planned one-sided *t*-test contrasts, was performed with G*Power Version 3.0.5 software (Erdfelder, Faul, & Buchner, 1996; for additional details, see Blankers et al., 2009). A sample size of 159 was needed for detecting small-to-medium effects of the interventions ($f = 0.25$) on the primary outcome variable (alcohol consumption) with $\alpha = 0.05$ and $\beta = 0.80$. To adjust for an expected attrition rate of 30% and because it has been suggested that use of generalized estimating equations has inflated Type 1 error rates with group sizes smaller than 60 (Stiger, Kosinski, Barnhart, & Kleinbaum, 1998), we increased the sample size to 205.

Participant Flow

Between June 2, 2008, and June 2, 2009, interested participants could complete the screening instrument on the study's website. Inclusion was discontinued when, based on the power analysis, a sufficient number of participants had been included. On June 5, 2008, the first participant was randomly allocated to an arm; the last participant was allocated on May 30, 2009. Of the 1,720 people who completed the screening questionnaire, 832 were eligible to participate; 205 decided to participate and were included and randomly allocated to any of the three arms (see Figure 1). Compared to all 832 people who were eligible to participate, the 205 participants who were included reported somewhat higher baseline AUDIT scores, but according to a one-sample *t* test, this difference was not significant: ($M = 18.9$, $SD = 4.98$ vs. $M = 19.5$, $SD = 5.13$), $t(200) = 1.617$, $p = .11$. Sixty-nine applicants were ineligible to participate for various reasons (see Figure 1), for example, because they had plans to travel extensively during the study, did not provide a valid e-mail address, or did not have health care insurance. Ten applicants who provided informed consent were not included because they did not respond to the invitation to participate. Six people were excluded after they had taken the baseline assessment. They had already been randomly allocated to an arm, and according to the intention-to-treat principle (Hollis & Campbell, 1999), they should have been included in all of the analyses. However, these participants explicitly requested that they be excluded from the study, and this resulted in six protocol deviations. One of them had been allocated to the TAO arm, two of them to the WL arm, and three to the SAO arm.

Participant Characteristics

Baseline characteristics of the 205 participants are shown in Table 1. Half of them were female with a mean age of 42.2 years ($SD = 9.7$). Reported AUDIT scores in combination with drinking frequencies indicate that the participants drank in an unhealthy manner at baseline. Participants' scores on the global severity index of the Brief Symptom Inventory (de Beurs & Zitman, 2006; Derogatis & Melisaratos, 1983) were in the high range according

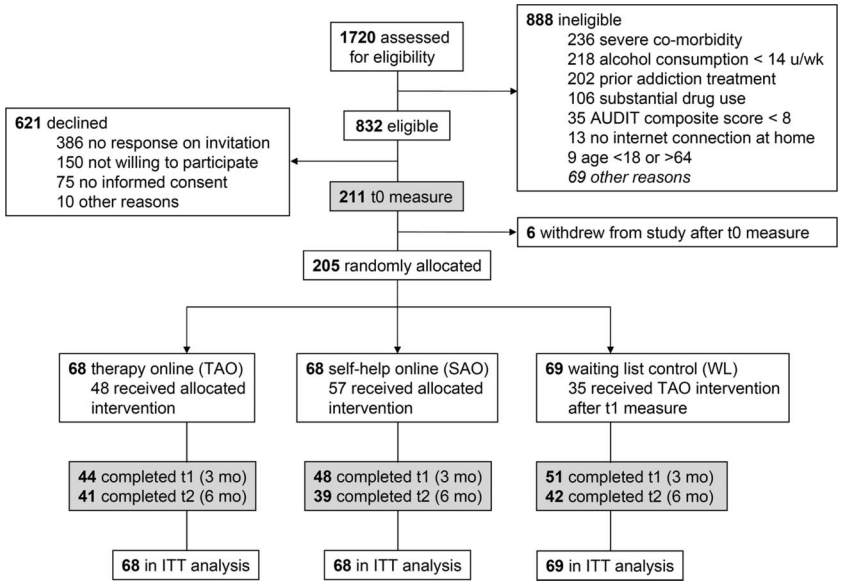

Figure 1. CONSORT flowchart of the recruitment and retention of participants. t0 = baseline; AUDIT = Alcohol Use Disorders Identification Test; ITT = intention-to-treat.

to the reference values for the Dutch general population (the high range is between 0.71 and 1.37 for women and between 0.56 and 1.34 for men; de Beurs, 2009). Most participants were employed; the most common job positions were in management, civil service, and information communications technology consultancy.

Measures

There were two primary outcome measures. The first was self-reported alcohol consumption (standard drinks) during the prior 7 days, measured at baseline, 3-month, and 6-month follow-up with

Table 1
Participant Characteristics

Variable	TAO (n = 68)	WL (n = 69)	SAO (n = 68)	F / Fisher's exact	p
Women	35 (51%)	34 (49%)	35 (51%)	0.106	.953
Age, years	41.9 (10.1)	43.7 (9.3)	41.1 (9.6)	1.284	.279
Education				5.569	.226
Low	2 (3%)	4 (6%)	7 (11%)		
Medium	24 (38%)	32 (49%)	30 (46%)		
High	38 (59%)	29 (45%)	29 (44%)		
Employed	58 (85%)	53 (78%)	55 (82%)	1.234	.539
Residential urbanization level				4.925	.295
Low	9 (13%)	7 (10%)	6 (9%)		
Medium	21 (31%)	32 (46%)	22 (32%)		
High	37 (55%)	30 (44%)	40 (59%)		
AUDIT composite score	18.8 (4.8)	20.1 (4.9)	19.6 (5.6)	0.096	.908
Years of alcohol problems	5.2 (5.7)	5.7 (5.7)	5.4 (5.7)	0.098	.907
Drinks per week	45.2 (26.3)	45.4 (28.9)	43.4 (24.0)	0.108	.898
Drinking days per week	6.0 (1.5)	5.9 (1.7)	5.6 (2.1)	1.129	.325
Cannabis lifetime use	29 (43%)	10 (15%)	21 (31%)	13.608	.001**
Cocaine lifetime use	17 (25%)	6 (9%)	11 (16%)	6.531	.042*
Amphetamine lifetime use	14 (21%)	6 (9%)	12 (18%)	4.150	.128
QOLS composite score	73.1 (14.4)	69.7 (16.7)	71.5 (20.0)	0.673	.511
EQ-5D score	0.79 (0.20)	0.76 (0.21)	0.80 (0.18)	1.048	.353
BSI Global Severity Index	0.81 (0.49)	0.80 (0.49)	0.77 (0.52)	0.159	.853
MCSDS Social Desirability scale	5.2 (2.6)	5.3 (2.2)	5.6 (2.5)	0.552	.557

Note. Presented data are counts (%) or M (SD). Education classification according to International Standard Classification of Education (UNESCO, 1997). TAO = therapy alcohol online; SAO = self-help alcohol online; WL = waiting-list control; AUDIT = Alcohol Use Disorders Identification Test; QOLS = Flanagan Quality of Life Scale; EQ-5D = the five-dimensional EuroQol instrument, score calculated with the MVH-A1 algorithm from Dolan (1997); BSI = Brief Symptom Inventory; MCSDS = Marlowe–Crowne Social Desirability Scale.

the timeline follow-back (TLFB) technique (Sobell & Sobell, 1992). The TLFB technique is a valid procedure for retrieving recent drinking histories (Carney, Tennen, Affleck, Del Boca, & Kranzler, 1998). The second primary outcome measure was treatment response (Blankers et al., 2009), which was defined as drinking within the British Medical Association's (1995) guidelines for safe drinking (a maximum of 14 standard drinks of alcohol/week for women, 21 standard drinks for men) and having less than a 10% deterioration on the AUDIT, the Flanagan Quality of Life Scale (QOLS; Flanagan, 1978), and the Global Severity Index (GSI) of the Brief Symptom Inventory (Derogatis & Melisaratos, 1983) between baseline and follow-up. Secondary outcome measures were the AUDIT total score and these quality-of-life measures: the QOLS and the EuroQol's EQ-5D (EuroQol Group, 1990). EQ-5D composite tariff scores were calculated with Dolan's (1997) algorithm. Tariff score valuations were obtained for EQ-5D health states from a representative sample of the UK population (Dolan, 1997). All primary and secondary outcome measures were given at baseline and at 3 months and 6 months postrandomization. Key demographic characteristics were also recorded at baseline. A 13-item version of the Marlowe–Crowne Social Desirability Scale (MCSDS; Crowne & Marlowe, 1960) was administered at baseline to evaluate the reliability of the self-reported data. The integrity of the CBT/MI that was delivered was assessed with the Yale Adherence and Competence Scale, Second Edition (YACS-II). The YACS-II is a validated instrument designed to assess therapist adherence and competence in the delivery of addiction treatment (Carroll et al., 2000). Two independent bachelor's degree psychologists (who were trained both in CBT and in how to score the YACS-II using 7-point Likert scales) rated all 48 chat-therapy transcripts. Both raters rated five of the transcripts so that intraclass correlation coefficients (ICC) could be calculated. The ICC for CBT competence was acceptable at 0.84; that for MI competence was low at 0.34. Mean competence scores for the CBT and MI items were calculated, resulting in a CBT competence score of $M = 4.32$ ($SD = 0.61$) and a MI competence score of $M = 4.30$ ($SD = 0.45$). The maximum possible rating on CBT/MI competence in the YACS-II is 7. A competence score of 3 or higher is considered acceptable (Carroll et al., 2000).

Procedure

Potential participants meeting the inclusion criteria were invited to participate. After providing informed consent, they were invited by e-mail to complete the online baseline assessment. All participants were invited for a follow-up assessment 3 and 6 months after randomization. After the baseline assessment had been completed, participants were allocated to one of the three trial arms. Participants allocated to TAO or SAO were introduced to the respective intervention. Participants allocated to the WL started with the TAO intervention when they completed the baseline assessment (3 months after they had been allocated to the WL).

Because attrition rates for Internet-based RCTs are often higher than for other kinds of RCTs (Eysenbach, 2005), extra effort was made to maximize response and retention rates. Suggestions for successful strategies to maximize response rates (e.g., Andrews, Nonnecke, & Preece, 2003; Kaplowitz, Hadlock, & Levine, 2004) were followed. For example, each follow-up questionnaire was made as short as possible, and participants were given feedback about their proceedings while they completed the questionnaire. All of the items of each measurement instrument were presented at once. Participants were sent a reminder if they had not responded within 5 days after they had been invited to complete the questionnaire. To encourage participants to feel that the project was worthwhile and to reward them for their time and effort, we sent them €15 in gift coupons (worth about $20) after they had completed the follow-up questionnaire. Participants who had not completed the questionnaire after being sent the reminder were contacted by telephone and encouraged to complete it. As a last resort, attempts were made to conduct interviews by telephone up to 75 days after the initial follow-up invitation. Master's-level psychology students working for a mental-health outcome evaluation call center conducted the interviews. This call center has been regularly collecting outcome data on outpatient addiction treatment for SATCs throughout the Netherlands since 2005; it is experienced in maximizing response rates in difficult-to-reach populations (Oudejans, Schippers, & Spits, 2006).

Interventions

SAO is a stand-alone, Internet-based, nontherapist involved, fully automated, self-guided treatment program that is based on a CBT/MI treatment protocol (de Wildt, 2000). It was developed in Adobe Flash, the" industry standard for interactive authoring and delivery of immersive experiences," and can be presented consistently "across personal computers, mobile devices, and screens of virtually any size and resolution" (http://www.adobe.com/products/flash/whatisflash). The information that is presented is text based, and feedback about alcohol consumption is provided with interactive graphs and tables. The intervention includes four tiers. The purpose of the first one is to monitor the person's alcohol consumption, help the person set drinking goals, and identify risky situations that might lead to relapse. By recoding his or her daily alcohol consumption and reporting the context cues and inner states (i.e., emotions, cognitions) related to it, the participant learns to recognize risky situations and to make them explicit. The purpose of the second tier is to provide feedback. It is presented in the form of graphs depicting the person's alcohol consumption and drinking-related contexts and inner states, and it compares the person's present consumption with the drinking goal that that person has set. The feedback helps participants gain insight into the process of goal striving and the fluctuations in their own drinking behavior. The third tier focuses on helping the person to acquire skills and knowledge about coping with craving, drinking lapses, and peer pressure and how to stay motivated in risky situations. Participants are encouraged to keep a personal online diary to apply the knowledge and skills acquired. The fourth tier consists of social support provided by other SAO participants through an Internet-based forum. In it, participants might discuss such things as their experiences while changing their drinking behavior, how to handle risky situations, and how to cope with lapses. Participants can access the self-help program any time they like, but it is suggested that they do so daily for at least 4 weeks. Earlier evaluations, however, indicate that only 5% of the participants actually spend this amount of time working with SAO (Blankers et al., 2008).

TAO is a synchronous online therapy that is based on the same CBT/MI treatment protocol (de Wildt, 2000) as SAO, and like

SAO, TAO was developed in Adobe Flash. It uses the same CBT treatment exercises as SAO, but they are extended and include up to seven synchronous text-based chat-therapy sessions lasting 40 min. each. Before each chat session, the participant works on a homework assignment. Each of the seven chat-therapy sessions has a different theme. Session 1 provides a general introduction to the TAO program. Session 2 discusses the pros and cons of alcohol use, how to monitor drinking behavior, and how to set goals. Session 3 teaches self-control. Session 4 focuses on risky situations; Session 5 introduces the concept of craving and how feelings can influence drinking; and Session 6 introduces the concepts lapse, relapse, and prolapse. In the final session, all the topics discussed during the previous sessions are reviewed. At the start of TAO, each participant is assigned to a therapist, on the basis of therapist availability. The therapists have a bachelor's or a master's degree in psychology, are supervised by PhD-level psychologists, and work for the collaborating SATC. They are trained in CBT and experienced in delivering protocolized, face-to-face CBT outpatient therapy to people with alcohol abuse or dependence. Each therapist received training and supervision in delivering CBT to problematic alcohol users over the Internet. Contact between the participant and the therapist was established in two ways: Asynchronous messaging (via secure e-mail) was used for making appointments, and synchronous messaging (via secure one-on-one chat sessions) was used for delivering the chat therapy. There was no other kind of contact between participants and therapists.

Statistical Analysis

Descriptive statistics were calculated for three kinds of participants: those who completed the initial screening but did not meet the inclusion criteria, those who were eligible to participate but decided not to do so, and those who were included in the study. Differences between the three groups were tested for significance with Fisher's exact test or one-way analysis of variance (ANOVA), as appropriate. Skewed distributions were log-transformed. Significant main effects in one-way ANOVAs were explored using post-hoc *t* tests with Bonferroni correction for multiple comparisons. Baseline characteristics of the participants included in the study were summarized using descriptive statistics, and balance across the three trial arms was tested using Fisher's exact test or one-way ANOVA, as appropriate. Again, skewness was reduced using log-transformations.

Analysis of the data was performed according to the intention-to-treat principle (Hollis & Campbell, 1999). The multiple imputation software Amelia-II was used to deal with missing data because it leads to the most optimal results for alcohol-consumption data. In a simulation study that used data collected in the pilot phase of the current RCT, Blankers, Koeter, and Schippers (2010) confirmed that the use of Amelia-II for imputing missing observations in longitudinal data sets that include data that are not normally distributed led to accurate results. Five instances of missing data in the original data set were imputed, and analyses were performed on each of these five data sets separately. In a final step, results from these five analyses were combined using Rubin's rules for combining estimates obtained from multiple imputed data sets (Rubin, 1987).

Effects of the interventions on the primary and secondary outcome variables were analyzed with generalized estimating equations (GEE) in a 3 (trial arm) $\times$ 3 (time) design. Covariates included (a) the minimized baseline variables (sex, AUDIT score, and years of problem drinking); (b) the variables on which the three trial arms differed significantly at baseline (lifetime cannabis and cocaine use); and (c) the MCSDS Social Desirability score. In all GEE analyses, an unstructured working correlation matrix was chosen. For alcohol consumption (standard drinks consumed in the prior week, one of the primary outcome measures), a negative binomial model with log link was used because it provides excellent fit for zero elevated alcohol-consumption data (Horton, Kim, & Saitz, 2007). For treatment response (another primary outcome measure), a binary logistic model was specified. For the secondary outcome variables, a normal model was chosen. Significance level was set at $\alpha = 0.05$, and all analyses were carried out with the R software environment for statistical computing (Version 2.7.0; R Development Core Team, 2008) and SPSS Version 17.0 (2008).

This trial was executed in compliance with the Helsinki Declaration and was approved by the Medical Ethics Committee of the University of Amsterdam, Academic Medical Center. All participants were provided with contact information for the collaborating SATC, the researchers, and an independent physician (in case they wanted additional treatment). The full trial protocol, which presents all study procedures, has been published elsewhere (Blankers et al., 2009).

Results

Response Analysis

Except for the proportion of participants reporting lifetime cannabis and cocaine use, the baseline characteristics of participants did not differ significantly among the three trial arms. Lifetime cannabis and cocaine use were, therefore, included as covariates in each of the GEE models.

After being invited for the 3-month postrandomization assessment, 51 participants responded within 5 days. After 5 days, 154 participants were sent a reminder by e-mail, to which 35 responded within a week; 119 of them did not respond and were contacted by the call center. Follow-up measures were collected from 143 of 205 (70%) of the randomly allocated participants, 3 months postrandomization. These measurements were collected at a median of 6 days ($M = 14.6$, $SD = 17.5$) after the first invitation had been sent. After being sent invitations for the 6-month postrandomization assessment, 43 participants responded within 5 days; 162 were sent a reminder by e-mail after 5 days, to which 27 responded within a week. The other 135 participants were contacted by the call center. Follow-up measures were collected from 122 of the original 205 participants (60%) a median of 7 days ($M = 18.4$, $SD = 23.7$) after the first invitation had been sent. A total of 156 participants (76%) completed at least one follow-up assessment. The proportion of participants who completed the 3-month (Fisher's exact test = 1.400, $p = .51$) and 6-month (Fisher's exact test = 0.219, $p = .94$) follow-up assessment did not differ among the three trial arms. Differences between assessment-responders and assessment-nonresponders on baseline characteristics (see Table 1) were explored. After Bonferroni correction for multiple post hoc comparisons (adjusted *p* value: 05/16 = .003), no differences were found between assessment responders and assessment nonresponders, 3 months postrandomization. Six months postrandomization, however, assessment responders and nonresponders

differed in age (M = 39.4 and M = 44.2 years, respectively), $t(202) = 3.622$, $p < .001$), but age was not included as a covariate in the model because the difference occurred after randomization. For most assessment responders (113/143, 79% at 3 months postrandomization; 108/122, 89% at 6 months postrandomization), the data were collected via Internet-based questionnaires. For a minority of them, data were collected via telephone interviews. After Bonferroni correction (adjusted p value: .05/16 = .003), there were no significant differences in participants' baseline characteristics between the two methods of data collection (self-report vs. telephone interview). At both the 3-month and the 6-month assessment, there were also no significant differences between the two methods of data collection in terms of the number of drinks that participants reported having consumed during the 7 days before the assessment: at 3 months (M = 30.3, SD = 24.7 vs. M = 28.0, SD = 22.0), $t(203) = 0.554$, $p = .58$; at 6 months (M = 31.2, SD = 25.2 vs. M = 28.2, SD = 22.2), $t(203) = 0.524$, $p = .60$.

GEE Model Parameters

Differential changes in the trial arms across time were modeled with GEE (see Table 2). For each of the four dependent variables that were predicted with GEE, the Trial Arm × Time interaction was significant, indicating differential effects of the different interventions on outcome. These interactions should be assessed in relationship to the hypotheses that were tested. Accordingly, in the next section the differences between the three trial arms at 3 months and between the SAO and TAO arms at 6 months are examined.

Primary Outcome Measures

In all three arms, participants reported less alcohol consumption at the 3-month follow-up than at baseline (see Figure 2). In the TAO arm, weekly consumption was reduced from M = 46.6 (SD = 26.4) standard drinks to M = 22.4 (SD = 21.3) standard drinks. In the SAO arm, consumption was reduced from M = 43.6 (SD = 23.8) standard drinks to M = 27.0 (SD = 24.8) standard drinks. In the WL arm, consumption was reduced from M = 47.2 (SD = 28.2) to M = 35.5 (SD = 23.6) standard drinks. To evaluate these changes in consumption, we ran a GEE negative binomial regression model in which alcohol consumption was the dependent variable; the predictor variables were trial arm (three levels) and time (three levels); the covariates were sex, duration of alcohol problems, AUDIT baseline scores, lifetime cannabis and cocaine use, and MCSDS Social Desirability scores. The model showed acceptable goodness of fit in an unstructured working correlation matrix according to the quasi-likelihood under independence model criterion (QIC): lowest = 531, highest = 566, for the five multiple imputed data sets. The overall test of the Trial Arm × Time interaction was significant ($p < .001$) across all of the imputed datasets, with Wald $\chi^2(8)$ ranging from 185 to 250.

Planned pairwise comparisons showed that participants in the TAO and SAO arms drank significantly less at the 3-months-postrandomization assessment than did participants in the WL arm: WL M = 35.5 vs. TAO M = 22.4, $t(135) = 3.15$, $p = .002$, one-tailed, $d = 0.59$; WL M = 35.5 vs. SAO M = 27.0, $t(135) = 2.04$, $p = .03$, one-tailed, $d = 0.36$ (see Table 3). The difference between TAO and SAO participants was not significant at 3 months postrandomization: TAO M = 22.4 vs. SAO M = 27.0, $t(134) = 1.24$, $p = .11$, one-tailed, $d = 0.20$. At 6 months postrandomization, the difference between TAO and SAO was significant, TAO M = 17.8 vs. SAO M = 26.2, $t(134) = 2.06$, $p = .03$, one-tailed, $d = 0.38$. TAO participants drank significantly fewer standard drinks than did SAO participants in the week before the 6-months-postrandomization assessment (see Table 4).

The second primary outcome measure was treatment response. In a logistic regression model to test its effects, treatment response at the 3-months-postrandomization assessment was the dependent

Table 2
Generalized Estimating Equations Predicting Primary and Secondary Outcome Variables at Baseline, 3 Months, and 6 Months Postrandomization for the Randomized Sample (n = 205)

Dependent	Variable	Average Wald χ^2	*df*	*p*
TLFB	(Intercept)	7,473.29	1	.0000
	Time	189.18	2	.0000
	Trial Arm	5.57	2	.0619
	Trial Arm × Time	27.32	4	.0000
AUDIT	(Intercept)	3,461.10	1	.0000
	Time	186.03	2	.0000
	Trial Arm	6.11	2	.0471
	Trial Arm × Time	30.84	4	.0000
QOLS	(Intercept)	6,802.48	1	.0000
	Time	87.87	2	.0000
	Trial Arm	11.11	2	.0039
	Trial Arm × Time	22.40	4	.0002
EQ-5D	(Intercept)	5,035.61	1	.0000
	Time	4.92	2	.0854
	Trial Arm	12.46	2	.0020
	Trial Arm × Time	16.86	4	.0021

Note. Covariates in each model were sex, AUDIT baseline score, years of drinking problems, lifetime cannabis use, lifetime cocaine use, and social desirability (Marlowe–Crowne Social Desirability Scale). *df* = degrees of freedom; TLFB = timeline follow-back; AUDIT = Alcohol Use Disorders Identification Test; QOLS = Flanagan Quality of Life Scale; EQ-5D = the five-dimensional EuroQol instrument.

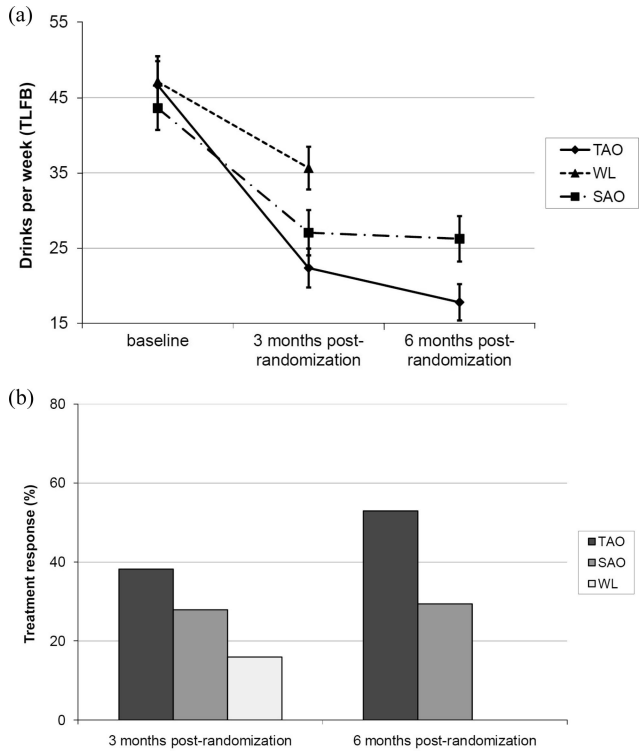

Figure 2. Primary outcome measures at 3 and 6 months postrandomization: (a) Means and standard errors of weekly alcohol consumption (standard drinks); (b) percentage of participants showing treatment response. TAO = therapy alcohol online; SAO = self-help alcohol online; WL = waiting-list control.

variable; trial arm was the predictor variable; and sex, duration of alcohol problems, AUDIT baseline scores, lifetime cannabis and cocaine use, and MCSDS were the covariates. The full model that included the predictor and all covariates (and a constant against a constant-only model across the five imputed data sets) reached borderline significance ($p < .10$) for three of the five imputed data sets, with $\chi^2(8)$ ranging from 9.8 to 21.3. The explained variance in treatment response was small, with Nagelkerke's R^2 ranging from 0.07 to 0.14 across the five data sets. Mean overall percentage of cases correctly predicted was 77%. According to the Wald criterion, participants in the TAO arm had significantly higher odds of being a treatment responder after 3 months than did participants in the WL arm, mean $\chi^2(1) = 9.0$, odds ratio (*OR*) = 3.5 (see Table 3). Number needed to treat (*NNT*) was calculated as 1/absolute risk reduction. Absolute risk reduction was calculated as the proportion of participants in the WL arm who did not show a treatment response and was reduced by the proportion who did not show a treatment response treatment response in the SAO arm. This yielded $NNT = 1/(58/69) - (42/68) = 5$, 95% CI [2.7, 12.7], pooled one-tailed, $p = .007$. Participants in the SAO arm had nonsignificantly higher odds of a treatment response 3 months postrandomization than did those in the WL arm, $\chi^2(1) = 2.8$, $OR = 2.1$, $NNT = 1/(58/69) - (49/68) = 9$, 95% CI [−3.9, 58.0], pooled one-tailed, $p = .06$. Participants in the TAO arm had higher odds of a treatment response than did participants in the SAO arm, but the difference was not significant according to the Wald criterion, $\chi^2(1) = 2.0$, $OR = 1.7$, $NNT = 1/(49/68) - (42/68) = 10$, 95% CI [−3.8, 18.4], pooled one-tailed, $p = .10$.

When we used a logistic regression model with the same predictors but with treatment response 6 months postrandomization as the dependent variable, the full model including all six predictors plus a constant against a constant-only model across the five imputed data sets reached borderline significance ($p < .10$) for four of the five imputed data sets, with $\chi^2(8)$ ranging from 8.5 to 30.8. The explained variance in treatment response was small to moderate, with Nagelkerke's R^2 ranging from 0.06 to 0.19 across the five data sets. Mean overall percentage of cases correctly predicted was 70%. According to the Wald criterion, participants in the TAO arm had nonsignificantly higher odds of being a treatment responder after 6 months than did participants in the SAO arm, mean $\chi^2(1) = 7.0$, $OR = 2.6$, $NNT = 1/(48/68) - (32/68) = 5$, 95% CI [2.5, 13.4], pooled one-tailed, $p = .06$ (see Table 4).

Table 3
Outcomes and Estimation 3 Months Postrandomization

Variable	TAO ($n = 68$)	WL ($n = 69$)	SAO ($n = 68$)	TAO vs WL	$t(135)$ / $\chi^2(1)$	p (one-tailed)
Drinks in last week	22.4 (21.3)	35.5 (23.6)	27.0 (24.8)	$d = 0.59$ [0.29, 0.90]	3.15	.002
Treatment response	26 (38%)	11 (16%)	19 (28%)	$OR = 3.54$ [1.33, 9.39]	9.0	.007
AUDIT score	13.7 (4.6)	16.4 (4.7)	14.8 (5.9)	$d = 0.59$ [0.29, 0.89]	3.41	.0006
QOLS score	84.9 (16.0)	77.0 (18.5)	83.9 (16.2)	$d = 0.46$ [0.17, 0.76]	2.48	.01
EQ-5D UK score	0.85 (0.33)	0.72 (0.33)	0.83 (0.24)	$d = 0.40$ [0.11, 0.70]	2.27	.02

Variable	SAO vs WL	$t(135)$ / $\chi^2(1)$	p (one-tailed)	TAO vs SAO	$t(134)$ / $\chi^2(1)$	p (one-tailed)
Drinks in last week	$d = 0.36$ [0.06, 0.65]	2.04	.03	$d = 0.20$ [−0.09, 0.50]	1.24	.11
Treatment response	$OR = 2.10$ [0.83, 5.32]	2.8	.06	$OR = 1.69$ [0.76, 3.77]	2.0	.10
AUDIT score	$d = 0.31$ [0.02, 0.60]	1.71	.05	$d = 0.21$ [−0.08, 0.51]	1.23	.11
QOLS score	$d = 0.40$ [0.11, 0.70]	2.29	.01	$d = 0.06$ [−0.23, 0.35]	0.34	.37
EQ-5D UK score	$d = 0.38$ [0.10, 0.68]	2.21	.02	$d = 0.07$ [−0.22, 0.36]	0.33	.38

Note. Presented data are counts (%) or *M* (*SD*) unless otherwise indicated. Cohen's *d* and 95% confidence interval were calculated with simulation (10,000 iterations). TAO = therapy alcohol online; SAO = self-help alcohol online; WL = waiting-list control; *SD* = standard deviation; *d* = Cohen's *d*; *OR* = odds ratio; data within brackets indicate the 95% confidence interval (CI) limits: [lower, upper]; AUDIT = Alcohol Use Disorders Identification Test; QOLS = Flanagan Quality of Life Scale; EQ-5D = the five-dimensional EuroQol instrument.

Table 4
Outcomes and Estimation 6 Months Postrandomization

Variable	TAO ($n = 68$)	SAO ($n = 68$)	TAO vs SAO	$t(134) / \chi^2(1)$	p (one-tailed)
Drinks in last week	17.8 (19.9)	26.2 (25.0)	$d = 0.38$ [0.09, 0.67]	2.06	.03
Treatment response	36 (53%)	20 (29%)	$OR = 2.62$ [0.87, 7.91]	7.0	.06
AUDIT score	12.6 (6.0)	15.0 (6.4)	$d = 0.39$ [0.11, 0.69]	2.26	.02
QOLS score	87.8 (17.5)	78.9 (23.4)	$d = 0.44$ [0.14, 0.73]	2.29	.02
EQ-5D UK score	0.89 (0.20)	0.78 (0.34)	$d = 0.40$ [0.12, 0.69]	2.86	.004

Note. Presented data are counts (%) or *M* (*SD*) unless otherwise indicated. Cohen's *d* and 95% confidence intervals were calculated with simulation (10,000 iterations). TAO = therapy alcohol online; SAO = self-help alcohol online; WL = waiting-list control; *SD* = standard deviation; *d* = Cohen's *d*; *OR* = odds ratio; data within brackets indicate the 95% confidence interval (CI) limits: [lower, upper]; AUDIT = Alcohol Use Disorders Identification Test; QOLS = Flanagan Quality of Life Scale; EQ-5D = the five-dimensional EuroQol instrument.

Secondary Outcome Measures

For the secondary outcome measures (scores on the AUDIT, QOLS, and EQ-5D), the pattern of results was similar to that for the primary outcome measures. A GEE linear regression model was run in which AUDIT, QOLS, or EQ-5D UK scores were the dependent variable; trial arm and time were predictors; and sex, duration of alcohol problems, AUDIT baseline scores, lifetime cannabis and cocaine use, and MCSDS Social Desirability scores were the covariates. The model yielded the following goodness-of-fit values according to the *QIC* in an unstructured working correlation matrix: (AUDIT, lowest = 12,988, highest = 13,807; QOLS, lowest = 159,143, highest = 168,090; EQ-5D, lowest = 29, highest = 31) for the five multiple imputed data sets. The Wald test of the Trial Arm × Time interaction was significant for the AUDIT, $\chi^2(8)$ between 192 and 247, $p < .001$; the QOLS, $\chi^2(8)$ between 100 and 149, $p < .001$; and the EQ-5D, $\chi^2(8)$ between 30 and 74, $p < .001$. Results of the pairwise comparisons for the three secondary outcome variables at 3 months are shown in Table 3; the results at 6 months are shown in Table 4. In general, differences between WL and the two interventions on the secondary outcome variables were significant at 3 months postrandomization. Differences between TAO and SAO did not reach significance at 3 months, but they were significant at 6 months postrandomization.

Discussion

In this study, participants who had received Internet-based therapy (TAO) had reduced their alcohol consumption 3 months after randomization more than had waiting-list (WL) controls. Three months after randomization, participants who received Internet-based self-help (SAO) had also reduced their alcohol consumption more than had WL controls. Although no support was found for the hypothesis that after 3 months participants receiving Internet-based therapy would show larger drinking reductions than those receiving Internet-based self-help, after 6 months the former participants did show larger reductions than the latter. The same differences between the groups at both 3 and 6 months were also found for reductions in alcohol-related problems and improvements in quality of life. Effect sizes for the primary and secondary outcome measures 3 months postrandomization indicated small effects for Internet-based self-help and small-to-medium effects for Internet-based therapy. The additional effects obtained for Internet-based therapy at 6 months postrandomization were small but statistically significant for alcohol consumption. In short, we found that both Internet-based therapy and Internet-based self-help are effective interventions for reducing problematic alcohol use, but larger effects were obtained for Internet-based therapy than for Internet-based self-help.

Generalizations of the study's results should be made in the light of the sample of problem drinkers who participated. Participation was not restricted to drinkers meeting diagnostic criteria for alcohol abuse or dependence. Rather, individuals who scored 8 or higher on the AUDIT were included. A cutoff score of 8 was chosen because previous studies using it have found very favorable sensitivity and usually lower, but still acceptable, specificity for ICD-10 (World Health Organization, 2005) diagnoses of current alcohol-use disorders and risk of future harm (Babor, Higgins-Biddle, Saunders, & Monteiro, 2001). Additionally, consistent with earlier research (Riper et al., 2008), both men and women who reported having drunk more than 14 standard drinks in the previous week were included in the study. The sample comprised relatively well-educated individuals, many of whom were employed full-time. These individuals represent a "new population" of problem drinkers who can be reached with Internet-based interventions (see Blankers et al., 2008; Cunningham et al., 2009; Postel et al., 2005; Riper et al., 2008). A potential threat to the generalizability of the current results could arise if drinkers who had been invited to participate but declined were markedly different from those who did participate. However, participants' AUDIT scores obtained during the screening phase indicated that this was not the case. Furthermore, the study was designed as a pragmatic clinical trial in that (a) clinically relevant alternative interventions were selected for comparison, (b) a diverse sample of participants was included, (c) participants from heterogeneous settings were recruited, and (d) a variety of health outcomes measures was collected (Tunis, Stryer, & Clancy, 2003). Based on these considerations, generalizability of the results to ordinary treatment contexts can be considered high.

The effect size obtained for Internet-based self-help compared to the waiting list control after 3 months ($d = 0.29$) is consistent with previous findings. To the best of our knowledge, this study was the first one to evaluate synchronous Internet-based therapy for problematic alcohol users in a randomized controlled design. For customary, face-to-face CBT/MI brief interventions for alcohol problems, short-term (6 months or less) effect sizes have ranged from $d = 0.14$ (Magill & Ray, 2009) to $d = 0.60$ (Vasilaki, Hosier, & Cox, 2006). For Internet-based interventions for depression and anxiety, interventions with therapist support show sub-

stantially larger effect sizes than Internet-based interventions without therapist support (i.e., $d = 1.00$ vs. $d = 0.24$; Spek et al., 2007). The results of the study presented in this paper show that although the effect size obtained for TAO compared with SAO was significant after 6 months ($d = 0.38$, $p = .03$), the differential effect size of the interventions with and without therapist involvement was modest when compared to Spek et al. (2007)'s results.

With assessment completion rates of 70% at 3 months and 60% at 6 months, participant attrition might be a limitation of this study. Attrition rates in eHealth studies (Eysenbach, 2005) and in brief face-to-face and self-help interventions (Miller & Wilbourne, 2002) tend to be relatively high. Addressing participant attrition and missing data is, therefore, of high importance. Dropout was minimized in the current study by supporting the data collection by a call center and rewarding participants with gift coupons for their participation. E-mail messages aimed at motivating participants were sent to those who had not responded to earlier invitations to complete an assessment. An intention-to-treat analysis was performed in which missing data were imputed with Amelia-II multiple imputation software for R (Honaker, King, & Blackwell, 2008). A secondary per-protocol analysis was performed in order to assess the sensitivity of the current data to Amelia-II's assumptions. The results led to comparable results and identical conclusions and are, therefore, not presented in this paper. This thorough approach to minimize the negative impact of participant attrition could be considered a strength of this study.

Data collection in this study was based solely on self-reports. Several studies have shown that the reliability and validity of self-report measures given over the Internet are acceptable (Del Boca & Darkes, 2003). Some studies have also validated instruments used in addiction research for use over the Internet. For example, Brodey et al. (2004) demonstrated the validity of Internet-based self-report administration of the Addiction Severity Index. Using noninvasive, self-report measures is the preferred method for obtaining information about participants' alcohol or other substance use at relatively low cost. In clinical trials that use self-selected research participants, biochemical tests and collateral informant reports do not add sufficiently to the accuracy of self-reports to warrant their routine use (Babor, Steinberg, Anton, & Del Boca, 2000). Nevertheless, some factors do influence their validity. An important influence on the accuracy of self-report data is participants' memory limitations. For example, how accurately can participants retrospectively recall the number of drinks they consumed at a particular time in the past? The TLFB methodology, which was used in this study to collect quantitative information about participants' alcohol consumption during the last 7 days, has been shown to be valid and reliable (e.g., Sobell, Maisto, Sobell & Cooper, 1979; Sobell, Sobell, Klajner, Pavan, & Basian, 1986; Tonigan, Miller, & Brown, 1997). The TLFB can be administered as an interview, self-administered, or administered by computer. This method involves asking clients to retrospectively estimate their daily alcohol consumption prior to the interview over a time period that can range from 7 days to 24 months. Important steps are taken to maximize the accuracy of the data: (a) A standard drink of alcohol is defined; (b) for days when the respondent did not drink, he or she is told to report 0 drinks rather than leaving the answer blank; and (c) memory aids are used to help the person remember the drinking pattern on the days when alcohol was consumed. Toll, Cooney, McKee, and O'Malley (2006) validated the use of the TLFB over a 7-day time horizon, which the current study also used. They compared reports of alcohol consumption using a daily interactive voice response system with the TLFB and found moderate-to-high correlations between the two ways of reporting drinking. In addition, a brief estimate of alcohol use over 7 days at baseline has been found to provide a representative assessment of number of drinks per drinking day in a comparison with time windows of 30 and 60 days (Toll, Leeman, McKee, & O'Malley, 2008). Another important safeguard is to try to prevent participants from answering in a socially desirable manner (Langenbucher & Merrill, 2001). Internet-based data-collection could have a positive influence in this regard. It has been shown, for instance, that in the absence of a research confederate, answers to survey questions collected using a computer interface are more reliable than when a confederate is present (Tourangeau, Steiger, & Wilson, 2002). In the current study, we included a social desirability scale in order to control statistically for the tendency to give socially desirable answers. A final consideration has to do with the interpretation of the MI competence score. The interpretation of this figure is difficult, given its low value of 0.34, as the reliability of the MI score is less than would be desirable.

Including an untreated waiting-list control group in this study was necessary for answering the research questions, but doing so was difficult from an ethical point of view. On the one hand, it was important to include an untreated control group in order to assess the full impact of Internet-based therapy. On the other hand, it would have been desirable to provide treatment with a minimal delay to all participants who needed treatment for their alcohol problem. These incompatible motivations were resolved by offering the control group treatment after the first follow-up assessment had been completed—a solution that Andersson et al. (2006) also used.

An alternative approach to using the TAO intervention would have been to include telephone contact as the communication medium between the participant and the therapist. There were, however, advantages to having written interaction between the two of them. One advantage was the high degree of perceived anonymity that it provided. Perceived anonymity helps to facilitate self-disclosure, openness, and disinhibition (Rochlen, Zack, & Speyer, 2004; Suler, 2004). It makes the problem behavior, which the participant is often ashamed of, easier to discuss, so that therapy can be delivered more directly. Another advantage of written communication is that it allows participants the possibility of rereading the interaction between themselves and their therapist. The therapeutic intervention is still available to the participant after the chat session has ended; this has been found to support the therapeutic process (Suler, 2004).

Based on this study and previous results, the use of Internet-based CBT/MI self-help and therapy for problematic alcohol use is backed by evidence on its effectiveness. When implemented in regular health care, such interventions are able to attract an insufficiently addressed population and will be ubiquitously available because they are Internet-based. The number of problem drinkers for whom CBT is available will grow because of these Internet-based interventions. Internet-based self-help reduces alcohol use at minimal or no costs to the participant. Internet-based therapy could make access to therapists more equitable by providing a service to clients in areas where therapist availability is currently limited (Kessler et al., 2009).

Some important questions were not addressed in the current study. Although some basic analysis of the text-based chat contacts has been performed and is presented, a more extensive analysis of these data is being planned. This is important for relating treatment integrity and treatment adherence to participant characteristics and clinical outcome. Overall, the results of the current study support the effectiveness of Internet-based therapy and Internet-based self-help for problematic alcohol users. After 6 months, the more intensive Internet-based therapy program with synchronous therapist contacts led to better outcome than did the less intensive self-help program. Internet-based interventions are able to attract a new population of problematic drinkers into treatment, including men and women who often are gainfully employed but have a clear need for assistance in tackling their drinking problems. Internet-based self-help is effective, but Internet-based therapy is more effective for reducing problematic alcohol use.

References

Andersson, G., Carlbring, P., Holmstrom, A., Sparthan, E., Furmark, T., Nilsson-Ihrfelt, E., . . . Ekselius, L. (2006). Internet-based self-help with therapist feedback and in vivo group exposure for social phobia: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 74,* 677–686.

Andrews, D., Nonnecke, B., & Preece, J. (2003). Electronic survey methodology: A case study in reaching hard-to-involve Internet users. *International Journal of Human–Computer Interaction, 16,* 185–210.

Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B., & Monteiro, M. G. (2001). *The Alcohol Use Disorders Identification Test: Guidelines for use in primary care* (2nd ed.). Geneva, Switzerland: World Health Organization.

Babor, T. F., Steinberg, K., Anton, R., & Del Boca, F. (2000). Talk is cheap: Measuring drinking outcomes in clinical trials. *Journal of Studies on Alcohol, 61,* 55–63.

Bewick, B. M., Trusler, K., Barkham, M., Hill, A. J., Cahill, J., & Mulhern, B. (2008). The effectiveness of web-based interventions designed to decrease alcohol consumption: A systematic review. *Preventive Medicine, 47,* 17–26.

Bewick, B. M., Trusler, K., Mulhern, B., Barkham, M., & Hill, A. J. (2008). The feasibility and effectiveness of a web-based personalized feedback and social norms alcohol intervention in UK university students: A randomized control trial. *Addictive Behaviors, 33,* 1192–1198.

Blankers, M., Kerssemakers, R., Schramade, M., Nabitz, U., & Schippers, G. M. (2008). Internetprogramm Selbsthilfe Alcohol: Erste Ergebnisse [Internet program self-help alcohol: First experiences]. *Sucht, 54,* 279–287.

Blankers, M., Koeter, M., & Schippers, G. M. (2009). Evaluating real-time Internet therapy and online self-help for problematic alcohol consumers: A three-arm RCT protocol. *BMC Public Health, 9,* 16.

Blankers, M., Koeter, M. W. J., & Schippers, G. M. (2010). Missing data approaches in eHealth research: Simulation study and a tutorial for non-mathematically inclined researchers. *Journal of Medical Internet Research, 12,* e54.

Bodin, M. C., & Romelsjö, A. (2007). Secondary outcomes: Group and individual change and relationships to drinking outcomes. *Addiction Research & Theory, 15,* 587–599.

British Medical Association. (1995). *Guidelines on sensible drinking.* London, England: Author.

Brodey, B. B., Rosen, C. S., Brodey, I. S., Sheetz, B. M., Steinfeld, R. R., & Gastfriend, D. R. (2004). Validation of the Addiction Severity Index (ASI) for Internet and automated telephone self-report administration. *Journal of Substance Abuse Treatment, 26,* 253–259.

Carney, M. A., Tennen, H., Affleck, G., Del Boca, F. K., & Kranzler, H. R. (1998). Levels and patterns of alcohol consumption using timeline follow-back, daily diaries, and real-time "electronic interviews." *Journal of Studies on Alcohol, 59,* 447–454.

Carroll, K. M., Nich, C., Sifry, R. L., Nuro, K. F., Frankforter, T. L., Ball, S. A., . . . Rounsaville, B. J. (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and Alcohol Dependence, 57,* 225–238.

Casswell, S., & Thamarangsi, T. (2009). Reducing harm from alcohol: Call to action. *Lancet, 373,* 2247–2257.

Christensen, H., & Mackinnon, A. (2006). The law of attrition revisited. *Journal of Medical Internet Research, 8,* e20.

Conigrave, K. M., Saunders, J. B., & Reznik, R. B. (1995). Predictive capacity of the AUDIT questionnaire for alcohol-related harm. *Addiction, 90,* 1479–1486.

Crowne, D., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting and Clinical Psychology, 24,* 349–354.

Cunningham, J. A., Wild, T. C., Cordingley, J., van Mierlo, T., & Humphreys, K. (2009). A randomized controlled trial of an Internet-based intervention for alcohol abusers. *Addiction, 104,* 2023–2032.

de Beurs, E. (2009). *Brief Symptom Inventory: Handleiding Addendum* [Brief Symptom Inventory: Manual addendum]. Leiden, the Netherlands: Psychologische Instrumenten Tests & Services.

de Beurs, E., & Zitman, F. G. (2006). The Brief Symptom Inventory (BSI): The reliability and validity of a brief alternative of the SCL-90. *Maandblad Geestelijke Volksgezondheid, 61,* 120–141.

de Graaf, R., ten Have, M., & van Dorsselaer, S. (2010). *De Psychische gezondheid van de Nederlandse bevolking: NEMESIS-2. Opzet en eerste resultaten* [Mental health in the Netherlands population: NEMESIS-2. Study design and first results]. Utrecht, the Netherlands: Trimbos Instituut.

Del Boca, F. K., & Darkes, J. (2003). The validity of self-reports of alcohol consumption: State of the science and challenges for research. *Addiction, 98,* 1–12.

Derogatis, L. R., & Melisaratos, N. (1983). The Brief Symptom Inventory: An introductory report. *Psychological Medicine, 13,* 596–605.

de Wildt, W. (2000). *Achilles leefstijl* 1 [Achilles Lifestyle 1]. Zeist, the Netherlands: Cure & Care.

Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care, 35,* 1095–1108.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavioral Research Methods, 28,* 1–11.

EuroQol Group. (1990). EuroQol: A new facility for the measurement of health-related quality of life. *Health Policy, 16,* 199–208.

Eysenbach, G. (2005). The law of attrition. *Journal of Medical Internet Research, 7,* e11.

Flanagan, J. C. (1978). A research approach to improving our quality of life. *American Journal of Psychology, 33,* 138–147.

Hester, R. K., & Delaney, H. D. (1997). Behavioral self-control program for Windows: Results of a controlled clinical trial. *Journal of Consulting and Clinical Psychology, 65,* 686–693.

Hollis, S., & Campbell, F. (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *British Medical Journal, 319,* 670–674.

Honaker, J., King, G., & Blackwell, M. (2008). Amelia 2: A program for missing data. R Package version 1.1–33 [Computer software]. Retrieved from http://gking.harvard.edu/amelia

Horton, N. J., Kim, E., & Saitz, R. (2007). A cautionary note regarding count models of alcohol consumption in randomized controlled trials. *BMC Medical Research Methodology, 7,* 9.

Jarvis, T. J. (1992). Implications of gender for alcohol treatment research: A quantitative and qualitative review. *British Journal of Addiction, 87,* 1249–1261.

Kaplowitz, M. D., Hadlock, T. D., & Levine, R. (2004). A comparison of

web and mail survey response rates. *Public Opinion Quarterly, 68,* 94–101.

Kessler, D., Lewis, G., Kaur, S., Wiles, N., King, M., Weich, S., . . . Peters, T. J. (2009). Therapist-delivered Internet psychotherapy for depression in primary care: A randomised controlled trial. *Lancet, 374,* 628–634.

Kohn, R., Saxena, S., Levav, I., & Saraceno, B. (2004). The treatment gap in mental health care. *Bulletin of the World Health Organization, 82,* 858–866.

Kypri, K., Langley, J. D., Saunders, J. B., Cashell-Smith, M. L., & Herbison, P. (2008). Randomized controlled trial of web-based alcohol screening and brief intervention in primary care. *Archives of Internal Medicine, 168,* 530–536.

Langenbucher, J., & Merrill, J. (2001). The validity of self-reported cost events by substance abusers: Limits, liabilities, and future directions. *Evaluation Review, 25,* 184–210.

Linke, S., Brown, A., & Wallace, P. (2004). Down your drink: A web-based intervention for people with excessive alcohol consumption. *Alcohol and Alcoholism, 39,* 29–32.

Magill, M., & Ray, L. A. (2009). Cognitive–behavioral treatment with adult alcohol and illicit drug users: A meta-analysis of randomized controlled trials. *Journal of Studies on Alcohol and Drugs, 70,* 516–527.

McLellan, A. T., Alterman, A. I., Metzger, D. S., Grissom, G. R., Woody, G. E., Luborsky, L., & O'Brien, C. P. (1994). Similarity of outcome predictors across opiate, cocaine, and alcohol treatments: Role of treatment services. *Journal of Consulting and Clinical Psychology, 62,* 1141–1158.

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.

Moyer, A., Finney, J. W., Swearingen, C. E., & Vergun, P. (2002). Brief interventions for alcohol problems: A meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction, 97,* 279–292.

Oudejans, S., Schippers, G. M., & Spits, M. (2006). *Benchmark Leefstijltraining in de Verslavingszorg* [Benchmark Lifestyle training in substance abuse treatment]. Amsterdam, the Netherlands: Amsterdam Institute of Addiction Research.

Pocock, S. J., & Simon, R. (1975). Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics, 31,* 103–115.

Postel, M. G., de Jong, C. A., & de Haan, H. A. (2005). Does e-therapy for problem drinking reach hidden populations? *American Journal of Psychiatry, 162,* 2393.

R Development Core Team. (2008). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rehm, J., Mathers, C., Popova, S., Thavorncharoensap, M., Teerawattananon, Y., & Patra, J. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. *Lancet, 373,* 2223–2233.

Rehm, J., Taylor, B., & Room, R. (2006). Global burden of disease from alcohol, illicit drugs and tobacco. *Drug and Alcohol Review, 25,* 503–513.

Riper, H., Kramer, J., Smit, F., Conijn, B., Schippers, G., & Cuijpers, P. (2008). Web-based self-help for problem drinkers: A pragmatic randomized trial. *Addiction, 103,* 218–227.

Rochlen, A. B., Zack, J. S., & Speyer, C. (2004). Online therapy: Review of relevant definitions, debates, and current empirical support. *Journal of Clinical Psychology, 60,* 269–283.

Room, R. (2005). Stigma, social inequality and alcohol and drug use. *Drug and Alcohol Review, 24,* 143–155.

Rubin, D. B. (1987). *Multiple imputation for survey nonresponse.* New York, NY: Wiley.

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption. *Addiction, 88,* 791–804.

Sinadinovic, K., Berman, A. H., Hasson, D., & Wennberg, P. (2010). Internet-based assessment and self-monitoring of problematic alcohol and drug use. *Addictive Behaviors, 35,* 464–470.

Sobell, L. C., Maisto, S. A., Sobell, M. B., & Cooper, A. M. (1979). Reliability of alcohol abusers' self-reports of drinking behavior. *Behaviour Research and Therapy, 17,* 157–160.

Sobell, L. C., & Sobell, M. B. (1992). Timeline follow-back: A technique for assessing self-reported ethanol consumption. In J. Allen & R. Z. Litten (Eds.), *Measuring alcohol consumption: Psychosocial and biological methods* (pp. 41–72). Totowa, NJ: Humana Press.

Sobell, M. B., Sobell, L. C., Klajner, D. F., Pavan, D., & Basian, E. (1986). The reliability of a timeline method for assessing normal drinker college students' recent drinking history: Utility for alcohol research. *Addictive Behaviors, 11,* 149–161.

Spek, V., Cuijpers, P., Nyklícek, I., Riper, H., Keyzer, J., & Pop, V. (2007). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: A meta-analysis. *Psychological Medicine, 37,* 319–328.

Stiger, T. R., Kosinski, A. S., Barnhart, H. X., & Kleinbaum, D. G. (1998). ANOVA for repeated ordinal data with small sample size? A comparison of ANOVA, MANOVA, WLS and GEE methods by simulation. *Communications in Statistics: Simulation and Computation, 27,* 357–375.

Suler, J. (2004). The online disinhibition effect. *Cyberpsychology & Behavior, 7,* 321–326.

Tate, D. F., & Zabinski, M. F. (2004). Computer and Internet applications for psychological treatment: Update for clinicians. *Journal of Clinical Psychology, 60,* 209–220.

Toll, B. A., Cooney, N. L., McKee, S. A., & O'Malley, S. S. (2006). Correspondence between interactive voice response (IVR) and timeline followback (TLFB) reports of drinking behavior. *Addictive Behaviors, 31,* 726–731.

Toll, B. A., Leeman, R. F., McKee, S. A., & O'Malley, S. S. (2008). A brief 7-day estimate of alcohol consumption for use in smoking cessation clinical trials. *Addictive Behaviors, 33,* 1606–1609.

Tonigan, J. S., Miller, W. R., & Brown, J. M. (1997). The reliability of Form 90: An instrument for assessing alcohol treatment outcome. *Journal of Studies on Alcohol, 58,* 358–364.

Tourangeau, R., Steiger, D. M., & Wilson, D. (2002). Self-administered questions by telephone: Evaluating interactive voice responses. *Public Opinion Quarterly, 66,* 265–278.

Tunis, S. R., Stryer, D. B., & Clancy, C. M. (2003). Practical clinical trials: Increasing the value of clinical research for decision making in clinical and health policy. *Journal of the American Medical Association, 290,* 1624–1632.

United Nations Educational, Scientific and Cultural Organization. (1997). *International standard classification of education (ISCED) 1997.* Paris, France: Author.

Vasilaki, E. I., Hosier, S. G., & Cox, W. M. (2006). The efficacy of motivational interviewing as a brief intervention for excessive drinking: A meta-analytic review. *Alcohol and Alcoholism, 41,* 328–335.

Williams, J., Herman-Stahl, M., Calvin, S. L., Pemberton, M., & Bradshaw, M. (2009). Mediating mechanisms of a military web-based alcohol intervention, *Drug and Alcohol Dependence, 100,* 248–257.

World Health Organization. (2005). *ICD-10: International statistical classification of diseases and related health problems, 10th revision* (2nd ed.). Geneva, Switzerland: Author.

Received May 26, 2010
Revision received January 25, 2011
Accepted February 23, 2011 ■