RESEARCH ARTICLE

ALCOHOL
CLINICAL & EXPERIMENTAL RESEARCH

The connection still matters: Therapeutic alliance with digital treatment for alcohol use disorder

Bryan Benitez | Tami L. Frankforter | Charla Nich | Brian D. Kiluk

Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut, USA

Correspondence
Bryan Benitez, Department of Psychiatry, Yale University, 40 Temple Street, Suite 6C, New Haven, CT 06510, USA.
Email: bryan.benitez@yale.edu

Abstract

Background: A strong cooperative bond between the patient and provider ("therapeutic alliance") is robustly associated with better alcohol use disorder (AUD) treatment outcomes. Although digital treatments for AUD have significant potential, the function of the alliance during digital programs is unclear. We compared the validity of patient-reported measures of the alliance with a digital treatment ("digital alliance") for AUD and the alliance with their clinician ("clinician alliance").

Methods: We used data from an 8-week, randomized clinical trial of a computerized cognitive behavioral therapy program (CBT4CBT) during outpatient AUD treatment. Treatment conditions included CBT4CBT with minimal clinical monitoring (CBT4CBT + monitor) or with treatment as usual (CBT4CBT + TAU). The digital alliance and clinician alliance were measured with similar versions of the Working Alliance Inventory (WAI). The WAI ratings were completed at the 2nd and 6th treatment sessions. A timeline followback calendar assessed daily alcohol use. Bayesian multilevel models compared the strength of the alliances and tested their associations with future alcohol use.

Results: Data from 43 participants were included (age M = 44; 65% male; 51% Black, 40% White, 9% other; 14% Hispanic). The digital alliance ratings had similar internal reliability as the clinician alliance ratings (ω's > 0.90). Differences between digital alliance and clinician alliance ratings were negligible in both treatment conditions (BF_{01} = 9 and 31). During treatment, within-person increases in the digital alliance and the clinician alliance predicted modest decreases in future drinking to a similar degree (BF_{01} = 15). Alliance ratings were not associated with future drinking when posttreatment follow-up drinking data were included (BF_{10} < 3).

Conclusions: The digital alliance with CBT4CBT was comparable to the clinician alliance. The digital alliance and clinician alliance had similar, albeit very small, associations with abstinence during treatment. Future research can explore how the digital alliance develops to improve AUD treatment efficacy.

KEYWORDS
alcohol use disorder, bonding, CBT, digital treatment, therapeutic alliance

INTRODUCTION

The therapeutic alliance ("alliance") is considered a core affiliative social process in psychiatric treatment. The alliance describes the cooperative bond between the patient and provider, including mutual trust and agreement with the treatment goals and tasks (Flückiger et al., 2018; Wampold, 2015). A stronger alliance is robustly associated with better treatment outcomes for substance use disorder (Connors et al., 2016; Meier et al., 2005; Miller & Moyers, 2015). Changes in the alliance during treatment often precede improvements in a variety of clinical outcomes, indicating that the alliance is temporally dynamic and potentially a vital treatment mechanism (Zilcha-Mano & Fisher, 2022). The relationship between the alliance and psychiatric symptoms is likely reciprocal, where increasing the alliance improves symptoms and improving symptoms increases the alliance (Falkenström et al., 2013; Flückiger et al., 2020). This reciprocal relationship may be reflected in the underlying neurobiology of addiction, since the same corticolimbic circuits are broadly implicated in both social processes and addiction (de Wit & Sayette, 2018; Goldman et al., 2010; Heilig et al., 2016; Venniro et al., 2018). Thus, multiple domains of research indicate that the alliance may be instrumental to the treatment of addictive behaviors, including alcohol use disorder (AUD).

However, conceptualizations of the alliance have been complicated by technological advances in treatment delivery (Lederman & D'Alfonso, 2021). Digital health tools for AUD, including mobile apps and web-based programs, have significant therapeutic potential (Carroll & Kiluk, 2017). But it is unclear how the alliance with a digital treatment program compares to the alliance with a human clinician. If the alliance is defined as requiring interpersonal interaction, then digital treatments may deliver specific treatment components (e.g., coping skills) without any alliance (Carroll & Kiluk, 2017). Alternatively, some researchers posit a "digital therapeutic alliance" to describe the quality of the therapeutic connection that patients form with digital treatments (D'Alfonso et al., 2020). A digital alliance consists of the patient's emotional engagement and agreement with the content portrayed by the digital treatment. For instance, Computer Based Training for Cognitive Behavioral Therapy (CBT4CBT) is a digital treatment that teaches CBT skills for avoiding substance use through didactic material and video vignettes (Carroll et al., 2008; Carroll & Kiluk, 2017). In these videos, professional actors portray protagonists with substance use disorders who face emotionally arousing situations often experienced by patients (e.g., interpersonal conflict; financial stressors). A patient's connection with the protagonists and trust that the CBT skills will reduce substance use could characterize an alliance with CBT4CBT. Therefore, a digital alliance is conceptually plausible as a variation of traditional formulations of the alliance.

Theoretical and empirical lines of evidence suggest that patients can form an alliance with digital treatments for AUD. The human capacity for affiliative bonding can extend to attachments with non-human species and socially significant objects (e.g., pets, religious symbols, and heirlooms), which can promote survival (Feldman, 2017; Moll & de Oliveira-Souza, 2009). Affiliative social experiences may also be elicited without any human interaction by simply viewing digital media that portray social content (Bortolini et al., 2021; Eslinger et al., 2021; Schilbach et al., 2013). Within the clinical domain, quantitative measures of the alliance with digital interventions have already been demonstrated in several studies (Kiluk et al., 2014; Lederman & D'Alfonso, 2021). Consequently, the plausibility of an alliance with digital interventions has been supported by both theory and empirical evidence.

An open question is whether the digital alliance functions similarly to the alliance with a clinician during AUD treatment. Similar functioning between the digital alliance and the alliance with a clinician could be demonstrated by comparable performance on empirical indicators of construct validity. A recent review suggested that the digital alliance may not be directly related to treatment outcomes, although few studies have explored this question (D'Alfonso et al., 2020). In our previous study evaluating participants' digital alliance with CBT4CBT, we did not find a statistically significant correlation between the digital alliance early in treatment and future cocaine abstinence (Kiluk et al., 2014). Notably, participants in that study self-reported a slightly lower bond with CBT4CBT compared to their bond with clinicians. Mean alliance ratings were somewhat stable over time; however, the association between outcomes and changes in the alliance over time was not directly tested. Investigating individual changes over time would be consistent with conceptualizing the alliance as a temporally dynamic treatment mechanism, rather than a fixed set of relational characteristics (Zilcha-Mano & Fisher, 2022). In the present study, we investigate how within-person longitudinal changes in the digital alliance relate to AUD treatment outcomes. We compared the reliability, mean ratings, and predictive accuracy of measures of the digital alliance with CBT4CBT and the alliance with clinicians during treatment for AUD.

MATERIALS AND METHODS

Analyses were conducted using data from a randomized clinical trial of treatments for AUD in an adult outpatient setting (Kiluk et al., 2016). All individuals seeking treatment for their alcohol use at an outpatient community facility that specialized in treatment for substance use disorders were invited to participate in the trial. Eligibility included being at least 18 years old, fluent in English, and meeting *DSM-IV* criteria for current alcohol abuse/dependence (American Psychiatric Association, 1994). Sixty-eight participants were randomly assigned to one of three treatment conditions delivered over an 8-week period: treatment as usual (TAU), CBT4CBT + TAU, or CBT4CBT + brief clinical monitoring ("monitor"). The "monitor" component consisted of a clinician's 10-minute weekly in-person review of the patient's functioning, safety, and use of CBT4CBT. TAU included weekly in-person individual and/or group therapy sessions with clinicians. Described in detail elsewhere (Kiluk et al., 2016), CBT4CBT consists of seven modules that each cover a core CBT skill (e.g., functional analysis; thought challenging; problem solving).

Participants were asked to complete one module per week on a computer in the clinic. Each module took about 35 minutes to complete. Modules included didactive instruction and emotionally arousing video vignettes portraying the use of CBT skills during high-risk situations for alcohol use.

Forty-six participants were randomized to one of the CBT4CBT conditions. The current analyses use data from 43 of these participants who completed the Working Alliance Inventory (WAI) at least once. In this subsample, 24 were assigned to CBT4CBT + monitor, and 19 were assigned to CBT4CBT + TAU. The mean age was 44 (SD = 12) and 65% were male. Most participants were Black (51%) or White (40%), and 14% identified as Hispanic.

Measures

Therapeutic alliance

Participants rated their alliance using the Working Alliance Inventory (Horvath & Greenberg, 1989). The WAI is comprised of 36 items that form 3 subscales measuring the strength of endorsement of the therapeutic tasks, goals, and bond with the clinician. Items are rated on a 7-point scale and averaged across items within each subscale. Since ratings on the 3 subscales are often highly correlated, they are typically combined into a mean total rating that is interpreted as the overall strength of the alliance (Falkenström et al., 2013; Horvath & Greenberg, 1989).

Participants completed separate WAI ratings for their alliance with the CBT4CBT program ("digital alliance") and their alliance with their clinician ("clinician alliance"). Both versions of the WAI were completed at sessions 2 and 6 during the 8-week treatment period. Participants in the CBT4CBT + monitor condition reported their alliance with their clinical monitor, since they did not have a formal therapist. Validation of the digital alliance version of the WAI has been described in a prior report (Kiluk et al., 2014). For the digital alliance version, the standard WAI was modified such that items identified the computer program rather than the participant's therapist (e.g., "I am confident in the computer program's ability to help me"), and instructions specified that items referred to the CBT4CBT program, not the computer itself or the research team. Results from a prior CBT4CBT trial showed that participants total ratings for their digital alliance were comparable to the total ratings for their clinician alliance (Kiluk et al., 2014). But the bond subscale of their digital alliance was rated lower than the bond subscale of their clinician alliance. Therefore, in this study, we first conducted analyses with the total WAI ratings and then repeated these analyses using only the WAI bond subscale ratings.

The WAI total ratings of the digital alliance and clinician alliance were strongly correlated in both treatment conditions at session 2 (CBT4CBT + monitor $r = 0.94$; CBT4CBT + TAU $r = 0.78$) and session 6 (CBT4CBT + monitor $r = 0.77$; CBT4CBT + TAU $r = 0.68$). Similarly, the WAI bond ratings of the digital alliance and clinician alliance were strongly correlated in both treatment conditions at session 2

(CBT4CBT + monitor $r = 0.86$ vs. CBT4CBT + TAU $r = 0.59$) and session 6 (CBT4CBT + monitor $r = 0.72$ vs. CBT4CBT + TAU $r = 0.67$).

Alcohol use

Daily alcohol use was self-reported using a timeline followback calendar at baseline, each week in therapy, and posttreatment follow-ups at 1 month, 3 months, and 6 months (Sobell et al., 1988). Each day was coded as either a drinking day (= 1) or not a drinking day (= 0). If an assessment was missed, then the next followback calendar assessment would also cover the missed days. The daily drinking data were linked to the most recently completed alliance assessment. That is, each alliance assessment was matched to drinking data that were reported only after the date when the alliance was measured. This data structure established temporal precedence, such that alliance ratings were always predicting future drinking (i.e., session 2 ratings predicted drinking between sessions 2 and 6; session 6 ratings predicted drinking between sessions 6 and 8). To evaluate the durability of any associations between alliance ratings and drinking outcomes, separate analyses were conducted that included drinking data collected during the follow-up period (i.e., session 6 alliance ratings predicted drinking between session 6 and the end of the 6-month follow-up period). Drinking data collected between treatment initiation and the first alliance assessment at session 2 were used as a covariate in the models predicting drinking outcomes. At baseline, all participants reported consuming alcohol at least once during the 28 days prior to study enrollment. The mean number of days participants had consumed alcohol during the 28 days prior to enrollment was 13 (SD = 9; Med = 12).

Alcohol use disorder symptoms

Baseline AUD severity was measured using the Alcohol Use Disorder Identification Test (Babor et al., 2001). The mean AUDIT score was 18.37 (SD = 7.97; Med = 17), indicating that most participants had moderate or severe AUD symptoms.

Statistical analyses

Analyses were conducted using the *R* statistical software (R Core Team, 2021). Internal consistency estimates for the digital alliance and the clinician alliance were generated using McDonald's ω and average interitem correlations, $\bar{r}_{ij}$. We then compared mean differences in the strength of endorsement between the digital alliance and the clinician alliance using Bayesian linear multilevel models. Since participants completed both WAI versions at two timepoints, the models specified that the alliance ratings were nested within participants. Predictors in the models included type of alliance (digital vs. clinician), treatment condition (CBT4CBT + TAU vs. CBT4CBT + monitor), timepoint, age, gender, AUDIT score, and proportion of days

drinking during the treatment period. An interaction between type of alliance and treatment condition was specified, since the alliance may depend on the extent of adjunctive clinical contact.

To test whether digital alliance ratings were negligibly lower than clinician alliance ratings, one-sided directional hypotheses were specified. One-sided hypothesis tests are preferred in Bayesian inference when there is scientific rationale for testing a specific alternative hypothesis against a null hypothesis (Keysers et al., 2020; van Doorn et al., 2021). We defined a negligible difference as 0.10 or less on the 7-point WAI scales, which roughly reflects a difference of 0.10 SD units in the WAI scales (see Table 1). A standardized difference of 0.10 is half the magnitude of the "small" Cohen's *d* effect size (Kruschke, 2018). Bayes factors compared the posterior odds favoring negligible differences given the data relative to the prior odds, BF_{01} (Makowski, Ben-Shachar, Chen, & Lüdecke, 2019). Larger BF_{01} values indicate evidence that favors negligible differences in *b* values ($b \in [-0.10, 0]$). We interpreted all Bayes factors using Jeffreys' evidence categories, which suggest that values greater than 3 substantially support a hypothesis (Wetzels et al., 2011). The prior distributions were informed by the maximum differences between the digital alliance and clinician alliance found in our previous study (Kiluk et al., 2014). Normal prior distributions were specified for the differences between mean alliance ratings for the total scale N (M = 0.24, SD = 2.50) and for the bond subscale N (M = 0.56, SD = 2.50). The means of these prior distributions reflect slightly higher mean clinician alliance ratings than mean digital alliance ratings, as found in the previous study. The standard deviations indicate that differences greater than 5.00 are very unlikely since the maximum possible difference on the WAI scales is 6.00. For all other covariates, we specified normal distributions with M = 0 and SD = 2.50. These models directly tested the evidence supporting negligible differences between the strength of the digital alliance and the clinician alliance.

To compare the predictive validity of the digital alliance and clinician alliance, daily drinking data during treatment were matched to the most recently completed alliance ratings. To ensure alliance ratings predicted only future drinking, drinking data between treatment initiation and the first alliance assessment were included as a covariate in the models, but not in the outcome variable. Bayesian logistic multilevel models nested daily drinking observations within participants. Covariates included type of alliance, treatment condition, weekend (Friday and Saturday vs. other days), days since study enrollment, age, gender, AUDIT score, and the proportion of days drinking between treatment initiation and the first alliance assessment. A logistic regression on dichotomized drinking data was used because 73% of the timeline followback calendar days recorded 0 drinks consumed, indicating a non-normal distribution of drinking counts. Alliance ratings were included in the model as both between-person and within-person predictors. The between-person alliance predictor was the participant's mean rating across both timepoints. The within-person predictor was the alliance rating at each timepoint subtracted from the participant's mean rating. For example, a participant who rated a 4 at the first timepoint and a 7 at the second timepoint would have a between-person alliance rating of 5.50 at both timepoints. Their within-person alliance ratings would be −1.50 for the first timepoint and 1.50 for the second timepoint. Both the between-person and within-person alliance predictors were included in the same models, with each being modeled as an interaction with the type of alliance. These interactions compared the associations between drinking and alliance ratings by the type of alliance (digital vs. clinician).

To determine if the digital alliance was less predictive of drinking than the clinician alliance, one-sided directional hypotheses were specified. We first tested for negative associations between each type of alliance and future drinking. These hypotheses were tested with Bayes Factors comparing the proportion of the posterior distribution that was negative versus positive (BF_{10}; $b < 0$ vs. $b > 0$). Larger BF_{10} values indicate evidence that favors negative *b* values (i.e., a negative association between the alliance and future drinking). Along with the BF_{10} values, we report regression coefficients and two-sided, equal-tailed 95% credibility intervals for the associations between alliance ratings and drinking. If the directional hypothesis test indicated an association, then we conducted two-sided hypothesis tests to assess the magnitude of the association between the alliance and drinking. These two-sided hypotheses were tested with Bayes Factors comparing

TABLE 1 Descriptive statistics and internal consistency for alliance ratings by session and type.

Scale	Session	Type	*n*	M	SD	$\bar{r}_{ij}$	ω_t
WAI Total	2	Clinician	41	5.83	1.01	0.42	0.97
		Digital	43	5.66	0.91	0.38	0.96
	6	Clinician	22	5.95	1.02	0.47	0.97
		Digital	26	5.78	1.05	0.50	0.98
WAI Bond	2	Clinician	41	5.77	1.01	0.43	0.91
		Digital	43	5.50	1.02	0.39	0.90
	6	Clinician	22	5.97	0.86	0.40	0.91
		Digtal	26	5.62	1.08	0.48	0.93

Abbreviations: Clinician, patient-reported alliance with their clinician; Digital, patient-reported alliance with CBT4CBT; $\bar{r}_{ij}$, average interitem correlation; WAI, Working Alliance Inventory; ω_t, McDonald's omega total coefficient (similar to Cronbach's α).

the posterior odds relative to the prior odds favoring a negligibly small association between alliance ratings and future drinking, BF_{01} (Makowski, Ben-Shachar, Chen, & Lüdecke, 2019). Larger BF_{01} values indicate evidence that favors negligibly small log odds coefficients ($b \in [-0.18, 0.18]$). Log odds values of −0.18 and 0.18 correspond to Cohen's d values of −0.10 and 0.10, which are half the magnitude of the frequently cited "small" effect size benchmarks of $d = 0.20$ and −0.20. To assess whether the association between the digital alliance and future drinking was weaker than the association between the clinician alliance and future drinking, we defined a negligible difference between these coefficients as a log odds difference of −0.18 or less. Bayes factors for these directional hypotheses compared the posterior odds favoring negligible differences in b values relative to the prior odds, BF_{01}. Larger BF_{01} values indicate evidence that favors negligibly small differences in log odds coefficients ($b \in [-0.18, 0]$). The prior distributions were informed by previous research showing small to moderately sized correlations between alliance ratings and substance use outcomes (Kiluk et al., 2014; Miller & Moyers, 2015). Normal prior distributions were specified for the within-person and between-person alliance regression coefficients: N (M = −0.50, SD = 2.50). The means of these prior distributions reflect a modest negative association between alliance ratings and drinking ($r = -0.14$, converted to −0.50 log odds). The standard deviations indicate that extremely strong associations (r's > 0.80 and < −0.80) are very unlikely. For all other covariates, we specified normal distributions with M = 0 and SD = 2.50.

To check the robustness of our results, we reanalyzed the data using alternative models with different prior distributions. Normal prior distributions with M = 0 and SD = 1 were specified for all the predictors, and numeric variables were standardized as z-scores (Lemoine, 2019). We report in the results one instance when these alternative standardized models produced a different conclusion from the initial models. Intraclass correlation coefficients (ICC) generated from unconditional models assessed the proportion of variance in each outcome variable attributable to within-person changes. Potential bias due to missing data was addressed using multilevel multiple imputation (Grund et al., 2018). The number of imputed datasets was determined by the percentage of missing data. Separate imputation datasets were generated for each type of alliance and then merged to preserve any potential interaction effects. All covariates from the Bayesian models were included in the imputation process. Predictors were included as fixed effects. Random intercepts were specified in the models predicting alliance ratings. Random intercepts and random slopes for the number of days since study enrollment were specified in the models predicting drinking.

Bayesian analyses were used because they can directly test whether the data supports "null" findings (e.g., negligible differences, practical equivalence, etc.). Simulation research indicates that Bayes factor testing of the region of practical equivalence has more statistical power for detecting negligible differences than a frequentist method (Linde et al., 2021). The Bayesian models were implemented in R with the $brms$ package (Version 2.17) (Bürkner, 2017). Negligible differences were tested in these models using the $emmeans$ (Version 1.8.3) and the $bayestestR$ package (Version 0.12.1) (Lenth, 2016; Makowski, Ben-Shachar, & Lüdecke, 2019). Model fit comparisons were made using the widely applicable information criterion ($WAIC$) estimates from the loo package (Version 2.5.1) (Vehtari et al., 2017). ICC estimates were generated using the $performance$ package (Version 0.7.3) (Lüdecke et al., 2021). Multiple imputation was conducted with the $mice$ package (Version 3.14) (van Buuren & Groothuis-Oudshoorn, 2011). The internal consistency estimates were obtained using the $psych$ package (Version 2.2.9) (Revelle & Condon, 2019).

RESULTS

Prior to multiple imputation, a total of 132 alliance assessments and 2369 daily timeline followback data points (9145 including data during follow-up period) were available for analyses. Twenty-two participants completed ratings of their clinician alliance at both timepoints, 19 completed the ratings only once, and 2 did not complete the ratings at either timepoint. Twenty-six participants completed ratings of their digital alliance at both timepoints, and 17 completed the ratings only once. Descriptive statistics and internal consistency estimates for the alliance measures showed similarities between the digital alliance and clinician alliance (Table 1). Both alliance measures demonstrated high internal consistency. Within-person changes accounted for 28% of the variance in the alliance total scale ratings ($ICC = 0.72$), 39% of the variance in the alliance bond subscale ratings ($ICC = 0.61$), and 29% of the variance in drinking ($ICC = 0.71$). Diagnostic measures suggested that the parameters had converged across all the models ($\hat{r}$ values < 1.04).

For the alliance total scale, the strength of the digital alliance and clinician alliance was negligibly different in both treatment conditions (Figure 1; CBT4CBT + monitor, $BF_{01} = 9$; CBT4CBT + TAU, $BF_{01} = 31$). For the alliance bond subscale, the strength of the digital alliance and clinician alliance was also negligibly different in both treatment conditions (Figure 1; CBT4CBT + monitor, $BF_{01} = 7$; CBT4CBT + TAU, $BF_{01} = 11$). The alternative model did not substantially support negligible differences in the strength of the bond subscales for the CBT4CBT + monitor condition ($BF_{01} = 2$). However, the alternative models did substantially support all the other negligible differences in the strength of the alliances ($BF_{01} > 3$).

During the days between the first alliance assessment and the end of treatment, alcohol use was reported on 27% of the days. Figure 2 displays bubble charts of the raw data showing that increases in alliance ratings were associated with decreased drinking during the 8-week treatment period. A model with a random slope for the number of days since study enrollment showed improved fit over a random intercept only model ($elpd\_diff = -207$, SE = 19). For the alliance total scale, future alcohol use was negatively associated with within-person increases in the digital alliance (Figure 3; $b = -0.38$, 95% CI[−0.92, 0.10], $BF_{10} = 14$) and the clinician alliance
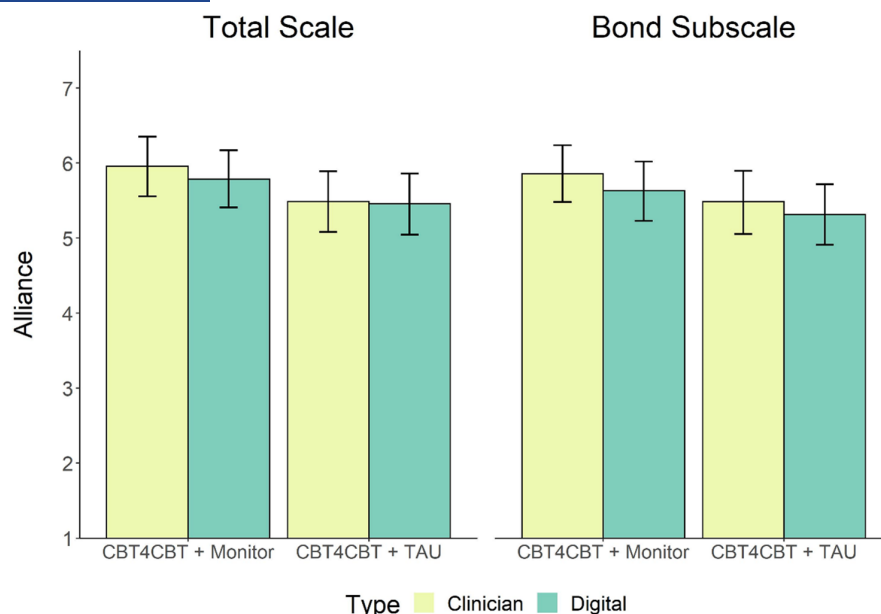
FIGURE 1 Results of the Bayesian linear multilevel models comparing mean levels of alliance ratings by treatment condition and alliance type. For the total scale and bond subscale, differences between the digital alliance and clinician alliance were negligible in both treatment conditions (compare light and dark bars). Error bars represent 95% Bayesian credibility intervals. CBT4CBT=Computer Based Training for Cognitive Behavioral Therapy; TAU=treatment as usual with a clinician; Monitor=clinician who provided 10 minutes of minimal clinical monitoring per week; Clinician=patient-reported alliance with their clinician; Digital=patient-reported alliance with CBT4CBT.
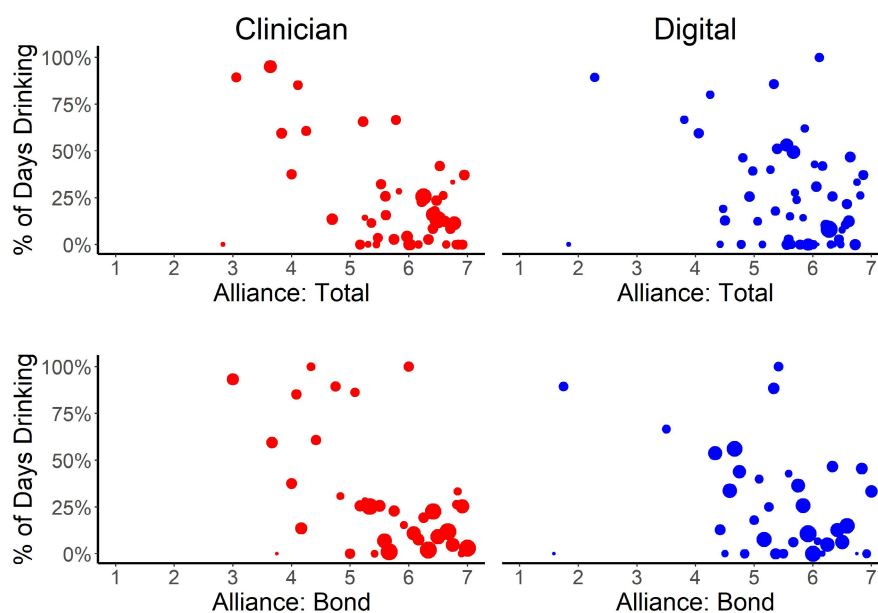

FIGURE 2 Bubble charts of future daily drinking proportions during treatment by alliance scale ratings (total, bond) and alliance type (clinician, digital). The sizes of the bubbles are proportional to the amount of data that contributed to the calculation of that proportion (i.e., the denominator); larger bubbles indicate that more data contributed to that proportion. The proportions were generated from disaggregated, unimputed raw data. Clinician=patient-reported alliance with their clinician; Digital=patient-reported alliance with CBT4CBT.

($b=-0.27$, 95% CI[−0.99, 0.23], $BF_{10}=5$). However, the evidence suggested that the associations with future drinking were negligibly small for both the digital alliance ($BF_{01}=7$) and the clinician alliance ($BF_{01}=10$). Furthermore, the difference in the magnitudes of these associations between drinking and within-person changes in the digital and clinician alliance total scales was negligible ($BF_{01}=15$).

For the alliance total scale, future alcohol use was not negatively associated with between-person differences in the digital alliance (Figure 3; $b=-0.06$, 95% CI[−0.59, 0.44], $BF_{10}=1$) and the clinician alliance ($b=-0.11$, 95% CI[−0.64, 0.39], $BF_{10}=2$). There was a negligible difference in the magnitude of these associations between drinking and between-person differences in the alliance total scale

FIGURE 3 Results of Bayesian logistic multilevel models comparing the probability of future daily drinking during treatment by alliance total ratings and type of alliance. As within-person alliance total ratings increased, future drinking decreased. Between-person alliance total ratings were not substantially associated with future drinking. Differences between the digital alliance and clinician alliance were negligible (compare solid lines and dotted lines). Shaded regions represent 95% Bayesian credibility intervals. Clinician=patient-reported alliance with their clinician; Digital=patient-reported alliance with CBT4CBT.
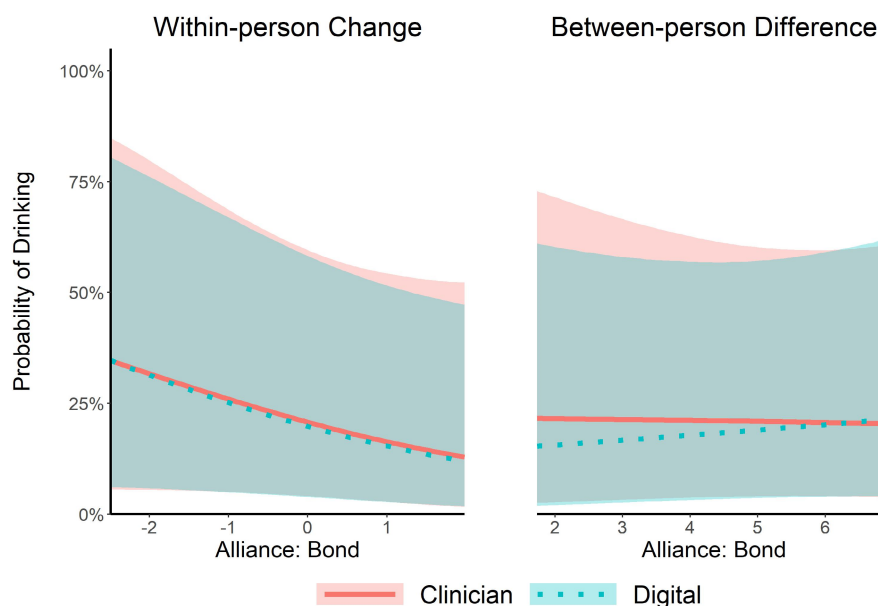

FIGURE 4 Results of Bayesian logistic multilevel models comparing the probability of future daily drinking during treatment by alliance bond ratings and type of alliance. As within-person alliance bond ratings increased, future drinking decreased. Between-person alliance bond ratings were not substantially associated with future drinking. Differences between the digital alliance and clinician alliance were negligible (compare solid lines and dotted lines). Shaded regions represent 95% Bayesian credibility intervals. Clinician=patient-reported alliance with their clinician; Digital=patient-reported alliance with CBT4CBT.

($BF_{01}=41$). The alternative models substantially supported all the same negligible differences ($BF_{01}>3$).

The results for the alliance bond subscale mirrored the results for the alliance total scale. For the alliance bond subscale, future alcohol use was negatively associated with within-person increases in the

digital alliance (Figure 4; $b=-0.31$, 95% CI[−0.75, 0.11], $BF_{10}=12$) and the clinician alliance ($b=-0.30$, 95% CI[−0.95, 0.19], $BF_{10}=6$). However, the evidence suggested that the associations with future drinking were negligibly small for both the digital alliance ($BF_{01}=9$) and the clinician alliance ($BF_{01}=9$). Furthermore, the difference in

the magnitudes of these associations between drinking and within-person changes in the digital and clinician alliance total scales was negligible ($BF_{01} = 15$). For the alliance bond subscale, future alcohol use was not negatively associated with between-person differences in the digital alliance (Figure 4; $b = 0.08$, 95% CI[−0.26, 0.42], $BF_{10} = 0.48$) and the clinician alliance ($b = −0.01$, 95% CI[−0.39, 0.36], $BF_{10} = 1$). There was a negligible difference in the magnitude of these associations between drinking and between-person differences in the alliance bond subscales ($BF_{01} = 22$). The alternative models substantially supported all the same negligible differences ($BF_{01} > 3$).

When including all drinking data collected during the treatment period and the 6-month follow-up period, alcohol use was reported on 25% of the days. For the alliance total scale, future alcohol use was no longer negatively associated with within-person increases in the digital alliance ($b = −0.04$, 95% CI[−0.35, 0.20], $BF_{10} = 1$) nor the clinician alliance ($b = −0.05$, 95% CI[−0.34, 0.17], $BF_{10} = 2$). There was a negligible difference in the magnitude of these associations between drinking and within-person changes in the alliance total scale ($BF_{01} = 41$). For the alliance bond subscale, future alcohol use was no longer negatively associated with within-person increases in the digital alliance ($b = −0.07$, 95% CI[−0.30, 0.14], $BF_{10} = 2$) nor the clinician alliance ($b = −0.06$, 95% CI[−0.33, 0.26], $BF_{10} = 2$). There was a negligible difference in the magnitude of these associations between drinking and within-person changes in the alliance bond subscale ($BF_{01} = 38$). Between-person differences in alliance total ratings and bond ratings were not negatively associated with future drinking ($BF_{10} < 3$). The alternative models also did not support negative associations between alliance ratings and future drinking when drinking data from the follow-up period were included ($BF_{10} < 3$).

DISCUSSION

We compared the internal reliability, mean ratings, and predictive validity of patient-reported measures of their alliance with clinicians and with a digital treatment for AUD (CBT4CBT). Participants rated the strength of their digital alliance at a comparable level to their clinician alliance. Ratings of the digital alliance and clinician alliance also had comparably high levels of internal consistency. During the treatment period, increases in the digital alliance over time predicted modest decreases in future drinking to a similar degree as increases in the clinician alliance over time. When the bond subscale of the alliance measure was analyzed separately, the results mirrored those of the total alliance scale. However, the weak associations between alliance ratings (both digital and clinician) and future drinking were not maintained when drinking data collected during a 6-month follow-up period were included in the analyses. These results suggest that the digital alliance may function similarly to the alliance with clinicians during AUD treatment, although the relationship to drinking outcomes is modest.

Our findings provided some evidence that a digital alliance may be a significant component of ongoing AUD treatment. When participants reported a stronger alliance with CBT4CBT, they also reported drinking less in the following weeks during treatment. A 1-point within-person increase in the digital alliance was prospectively associated with a 32% decrease in the odds of drinking (log odds $b = −0.38$). Notably, the association between future drinking and the digital alliance was comparable to that of the alliance with clinicians (24% decrease in the odds of drinking). However, the results suggested overall that the magnitude of the associations between drinking during the treatment period and alliance ratings were negligibly small for both the digital alliance and the alliance with a clinician. Furthermore, the small negative associations between alliance ratings and future drinking were not maintained when drinking data during the treatment follow-up period were included in the analyses. Therefore, the digital alliance may have a similar role as the alliance with clinicians during the treatment of AUD, although the importance of this role for promoting abstinence was modest in the current study.

Results from the alliance bond subscales were also notable. Participants rated their bond formed with digital treatment at a similar magnitude to their bond formed with their clinician. The strength of the digital bond and the bond with clinicians also predicted future drinking to a similar, albeit negligibly small, degree during the treatment period. The alliance bond subscale consists of trust (e.g., "I trust what _____ is doing for me"), confidence (e.g., "I am confident in _____'s ability to help me"), and positive affect toward the clinician or CBT4CBT program (e.g., "My interactions with _____ are positive"; "I feel uncomfortable with _____"). For the digital alliance version, participants were explicitly instructed to rate the CBT4CBT program itself, not the computer they used or any other aspects of the clinical research setting. The digital bonding experience with CBT4CBT may be due to emotionally engaging with the social content in the video vignettes, which portray familiar life circumstances similar to those encountered by persons with AUD (Eslinger et al., 2021; Feldman, 2017; McGaugh, 2018; Moll & de Oliveira-Souza, 2009; Montoya et al., 2008, 2017; Phelps, 2006). However, more research is needed to directly test how the bond portion of a digital alliance may develop and influence treatment outcomes (Zilcha-Mano & Fisher, 2022).

Although most of the variance in the digital alliance ratings was attributed to between-person differences, these between-person differences did not prospectively predict drinking in our study. In contrast, within-person changes in the digital alliance prospectively predicted decreases in drinking during treatment, albeit to a modest degree. These findings highlight the distinction between viewing the alliance as a static relational characteristic versus a temporally dynamic process that may influence outcomes (Connors et al., 2016; Zilcha-Mano & Fisher, 2022). Our findings suggest the potential utility of differentiating between static and temporally dynamic perspectives on the digital alliance.

A primary limitation of our study is the lack of experimental control, which is common in therapeutic alliance research (Flücki-ger et al., 2018). Although the alliance temporally preceded drinking in our statistical design, we did not randomly assign participants to treatment components that may have differentially influenced the

digital alliance. Therefore, we cannot rule out possible extraneous variables that simultaneously increase the alliance and decrease drinking. We did, however, include potentially important confounders as covariates in our models (e.g., drinking levels early in treatment; alcohol use disorder severity at baseline). Another limitation is that participants may have rated their digital alliance based on the overall support they received from the research study, although written instructions on the questionnaire emphasized rating their experience with only the CBT4CBT program and not the staff or setting. The strong correlations between the ratings of the digital alliance and clinician alliance may be evidence that it was difficult to rate each component separately. Similarly, it is possible that the alliance with CBT4CBT directly influenced the alliance with the therapist or clinical monitor, and vice-versa. Another notable limitation is that we did not detect any associations between alliance ratings and future drinking when the analyses included drinking data collected during the follow-up period. This finding may suggest that the alliance's influence on drinking is limited to near-term outcomes. However, this finding may also be explained by the primary results of the clinical trial, which showed that abstinence rates increased over time during the 8-week treatment period, but not the follow-up period (Kiluk et al., 2016). The plateauing of abstinence rates after the treatment period suggests that reductions in drinking stabilized for most participants, potentially restricting the variance in drinking rates that could be predicted in our models. Another methodological limitation to our study is the amount of uncertainty in the statistical estimates due to the modest sample size. Larger sample sizes could increase the precision of the estimated association between the digital alliance and future drinking. Since both digital alliance and clinician alliance ratings were necessary for meaningful within-person comparisons, data from the TAU only condition were excluded from our analyses. The alliance with clinicians in the TAU only condition may have differed from the treatment conditions that included CBT4CBT, limiting the generalizability of our findings to standard TAU. Given these limitations, our findings offer only preliminary evidence for the role of an alliance in digital treatments for AUD.

Future transdisciplinary research may explore more precisely how the alliance functions during digital treatment for AUD. Understanding how patients cognitively process digital social content would be especially useful, since social information is endowed with intrinsic value due to its relevance for survival (Norris & Cacioppo, 2007; Padilla-Coreano et al., 2022; Phelps, 2006). For CBT4CBT, this may include identifying which video vignettes are most memorable, emotionally arousing, personally relevant, and motivative of treatment activities (e.g., homework compliance). Investigating how engagement with digital social content influences neural pathways may provide further insights into the mechanisms of digital treatments (Magill et al., 2020; Reese et al., 2021). Culturally adapted digital treatments, such as the Spanish-language version of CBT4CBT, could also demonstrate how the digital alliance may depend upon what social information is valued within the patient's milieu (Paris et al., 2018). These future directions also necessitate more frequent measurement of the digital alliance, as they

all assume the alliance is temporally dynamic throughout treatment (Connors et al., 2016; Zilcha-Mano & Fisher, 2022).

Our study indicates that dynamic changes in the strength of the alliance may play an important role during digital treatments for AUD, as well as in traditional treatment modalities. Future research should explore how the digital alliance with computerized treatment develops and functions to change motivation for drinking.

CONFLICT OF INTEREST STATEMENT
Brian D. Kiluk is a consultant to CBT4CBT LLC, which makes CBT4CBT available to qualified clinical providers and organizations on a commercial basis. The conflict is managed through Yale University.

DATA AVAILABILITY STATEMENT
The data and R code that support the findings of this study are openly available in the Open Science Framework (OSF) repository at https://doi.org/10.17605/OSF.IO/XKARY.

ORCID
Bryan Benitez https://orcid.org/0000-0002-5759-5362
Brian D. Kiluk https://orcid.org/0000-0002-1175-7979

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

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*. Geneva: World Health Organization.

Bortolini, T., Melo, B., Basilio, R., Fischer, R., Zahn, R., de Oliveira-Souza, R. et al. (2021) Striatal and septo-hypothalamic responses to anticipation and outcome of affiliative rewards. *NeuroImage*, 243, 118474.

Bürkner, P.C. (2017) Brms: an R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, 80, 1–28.

Carroll, K.M., Ball, S.A., Martino, S., Nich, C., Babuscio, T.A., Nuro, K.F. et al. (2008) Computer-assisted delivery of cognitive-behavioral therapy for addiction: a randomized trial of CBT4CBT. *American Journal of Psychiatry*, 165, 881–888.

Carroll, K.M. & Kiluk, B.D. (2017) Cognitive behavioral interventions for alcohol and drug use disorders: through the stage model and back again. *Psychology of Addictive Behaviors*, 31, 847–861.

Connors, G.J., Maisto, S.A., Schlauch, R.C., Dearing, R.L., Prince, M.A. & Duerr, M.R. (2016) Therapeutic alliances predict session by session drinking behavior in the treatment of alcohol use disorders. *Journal of Consulting and Clinical Psychology*, 84, 972–982.

D'Alfonso, S., Lederman, R., Bucci, S. & Berry, K. (2020) The digital therapeutic alliance and human-computer interaction. *JMIR Mental Health*, 7, e21895.

de Wit, H. & Sayette, M. (2018) Considering the context: social factors in responses to drugs in humans. *Psychopharmacology*, 235, 935–945.

Eslinger, P.J., Anders, S., Ballarini, T., Boutros, S., Krach, S., Mayer, A.V. et al. (2021) The neuroscience of social feelings: mechanisms of

adaptive social functioning. *Neuroscience and Biobehavioral Reviews*, 128, 592–620.

Falkenström, F., Granström, F. & Holmqvist, R. (2013) Therapeutic alliance predicts symptomatic improvement session by session. *Journal of Counseling Psychology*, 60, 317–328.

Feldman, R. (2017) The neurobiology of human attachments. *Trends in Cognitive Sciences*, 21, 80–99.

Flückiger, C., Del, A.C., Wampold, B.E. & Horvath, A.O. (2018) The Alliance in adult psychotherapy: a meta-analytic synthesis. *Psychotherapy*, 55, 316–340.

Flückiger, C., Rubel, J., Del Re, A.C., Horvath, A.O., Wampold, B.E., Crits-Christoph, P. et al. (2020) The reciprocal relationship between alliance and early treatment symptoms: a two-stage individual participant data meta-analysis. *Journal of Consulting and Clinical Psychology*, 88, 829–843.

Goldman, M.S., Darkes, J., Reich, R.R. & Brandon, K.O. (2010) Anticipatory processing as a transdisciplinary bridge in addiction. In: Ross, D., Kincaid, H., Spurrett, D. & Collins, P. (Eds.) *What is addiction?* Cambridge, MA: The MIT Press, pp. 290–334.

Grund, S., Lüdtke, O. & Robitzsch, A. (2018) Multiple imputation of missing data for multilevel models: simulations and recommendations. *Organizational Research Methods*, 21, 111–149.

Heilig, M., Epstein, D.H., Nader, M.A. & Shaham, Y. (2016) Time to connect: bringing social context into addiction neuroscience. *Nature Neuroscience Reviews*, 17, 592–599.

Horvath, A.O. & Greenberg, L.S. (1989) Development and validation of the working Alliance inventory. *Journal of Counseling Psychology*, 36, 223–233.

Keysers, C., Gazzola, V. & Wagenmakers, E. (2020) Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, 23, 788–799.

Kiluk, B.D., Devore, K.A., Buck, M.B., Nich, C., Frankforter, T.L., LaPaglia, D.M. et al. (2016) Randomized trial of computerized cognitive behavioral therapy for alcohol use disorders: efficacy as a virtual stand-alone and treatment add-on compared with standard outpatient treatment. *Alcoholism, Clinical and Experimental Research*, 40, 1991–2000.

Kiluk, B.D., Serafini, K., Frankforter, T., Nich, C. & Carroll, K.M. (2014) Only connect: the working alliance in computer-based cognitive behavioral therapy. *Behaviour Research and Therapy*, 63, 139–146.

Kruschke, J.K. (2018) Rejecting or accepting parameter values in Bayesian estimation. *Advances in Methods and Practices in Psychological Science*, 1, 270–280.

Lederman, R. & D'Alfonso, S. (2021) The digital therapeutic Alliance: prospects and considerations. *JMIR Mental Health*, 8, e31385.

Lemoine, N.P. (2019) Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, 128, 912–928.

Lenth, R.V. (2016) Least-squares means: the *R* package lsmeans. *Journal of Statistical Software*, 69, 1–33.

Linde, M., Tendeiro, J.N., Selker, R., Wagenmakers, E.-J. & van Ravenzwaaij, D. (2021) Decisions about equivalence: a comparison of TOST, HDI-ROPE, and the Bayes factor. *Psychological Methods*, 28, 740–755.

Lüdecke, D., Ben-Shachar, M., Patil, I., Waggoner, P. & Makowski, D. (2021) Performance: an R package for assessment, comparison and testing of statistical models. *Journal of Open Source Software*, 6, 3139.

Magill, M., Tonigan, J.S., Kiluk, B., Ray, L., Walthers, J. & Carroll, K. (2020) The search for mechanisms of cognitive behavioral therapy for alcohol or other drug use disorders: a systematic review. *Behaviour Research and Therapy*, 131, 103648.

Makowski, D., Ben-Shachar, M. & Lüdecke, D. (2019) bayestestR: describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, 4, 1541.

Makowski, D., Ben-Shachar, M.S., Chen, S.H.A. & Lüdecke, D. (2019) Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, 10, 1–14.

McGaugh, J.L. (2018) Emotional arousal regulation of memory consolidation. *Current Opinion in Behavioral Sciences*, 19, 55–60.

Meier, P.S., Barrowclough, C. & Donmall, M.C. (2005) The role of the therapeutic alliance in the treatment of substance misuse: a critical review of the literature. *Addiction*, 100, 304–316.

Miller, W.R. & Moyers, T.B. (2015) The forest and the trees: relational and specific factors in addiction treatment. *Addiction*, 110, 401–413.

Moll, J. & de Oliveira-Souza, R. (2009) "Extended attachment" and the human brain: internalized cultural values and evolutionary implications. In: *The moral brain*. Dordrecht: Springer Netherlands, pp. 69–85.

Montoya, R.M., Horton, R.S. & Kirchner, J. (2008) Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships*, 25, 889–922.

Montoya, R.M., Horton, R.S., Vevea, J.L., Citkowicz, M. & Lauber, E.A. (2017) A re-examination of the mere exposure effect: the influence of repeated exposure on recognition, familiarity, and liking. *Psychological Bulletin*, 143, 459–498.

Norris, C.J. & Cacioppo, J.T. (2007) I know how you feel: social and emotional information processing in the brain. In: *Social neuroscience: integrating biological and psychological explanations of social behavior*. New York, NY: The Guilford Press, pp. 84–106.

Padilla-Coreano, N., Tye, K.M. & Zelikowsky, M. (2022) Dynamic influences on the neural encoding of social valence. *Nature Reviews. Neuroscience*, 23, 535–550.

Paris, M., Silva, M., Añez-Nava, L., Jaramillo, Y., Kiluk, B.D., Gordon, M.A. et al. (2018) Culturally adapted, web-based cognitive behavioral therapy for Spanish-speaking individuals with substance use disorders: a randomized clinical trial. *American Journal of Public Health*, 108, 1535–1542.

Phelps, E.A. (2006) Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology*, 57, 27–53.

R Core Team. (2021) *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reese, E.D., Kane, L.F., Paquette, C.E., Frohlich, F. & Daughters, S.B. (2021) Lost in translation: the gap between neurobiological mechanisms and psychosocial treatment research for substance use disorders. *Current Addiction Reports*, 8, 440–451.

Revelle, W. & Condon, D.M. (2019) Reliability from α to ω: a tutorial. *Psychological Assessment*, 31, 1395–1411.

Schilbach, L., Timmermans, B., Reddy, V., Costall, A., Bente, G., Schlicht, T. et al. (2013) Toward a second-person neuroscience. *Behavioral and Brain Sciences*, 36, 393–414.

Sobell, L.C., Sobell, M.B., Leo, G.I. & Cancilla, A. (1988) Reliability of a timeline method: assessing normal drinkers' reports of recent drinking and a comparative evaluation across several populations. *British Journal of Addiction*, 83, 393–402.

van Buuren, S. & Groothuis-Oudshoorn, K. (2011) Mice: multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45, 1–67.

van Doorn, J., van den Bergh, D., Böhm, U., Dablander, F., Derks, K., Draws, T. et al. (2021) The JASP guidelines for conducting and reporting a Bayesian analysis. *Psychonomic Bulletin & Review*, 28, 813–826.

Vehtari, A., Gelman, A. & Gabry, J. (2017) Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27, 1413–1432.

Venniro, M., Zhang, M., Caprioli, D., Hoots, J.K., Golden, S.A., Heins, C. et al. (2018) Volitional social interaction prevents drug addiction in rat models. *Nature Neuroscience*, 21, 1520–1529.

Wampold, B.E. (2015) How important are the common factors in psychotherapy? An update. *World Psychiatry*, 14, 270–277.

Wetzels, R., Matzke, D., Lee, M.D., Rouder, J.N., Iverson, G.J. & Wagenmakers, E.J. (2011) Statistical evidence in experimental psychology: an empirical comparison using 855 t tests. *Perspectives on Psychological Science*, 6, 291–298.

Zilcha-Mano, S. & Fisher, H. (2022) Distinct roles of state-like and trait-like patient–therapist alliance in psychotherapy. *Nature Reviews Psychology*, 1, 194–210.