MEASURING THE IMPACT OF PSYCHOLOGICAL THERAPY
AN EXPERIMENTAL APPROACH

Intermediate Report – Impact on Psychological Health

Morten Hesse,[†] Bastien Michel,[‡§*] Kristine Rømer Thomsen,[†] Marianne Simonsen[§]

[†] Department of Psychology and Behavioural Sciences, Aarhus University
Tuborgvej 164, 2400 København NV, Denmark
Bartholins Allé 10, 8000 Aarhus C, Denmark
[‡]Paris School of Economics
48 boulebard Jourdan, Paris, France
[§]Department of Economics and Business Economics, Aarhus University
Fuglesangs Allé 4, 8210 Aarhus V, Denmark

*To whom correspondence should be addressed: bastien.michel@psemail.eu

September 2021

TRYGFONDENS BØRNEFORSKNINGSCENTER

AARHUS BSS
SCHOOL OF BUSINESS AND SOCIAL SCIENCES
AARHUS UNIVERSITET

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

SUMMARY

In Denmark, alcohol and drug abuse problems are widespread issues and, as a consequence, a large number of individuals are growing up or have grown up in families in which such problems exist or existed. For instance, an estimated 632,000 people are growing up or have grown up in families with alcohol problems, with serious consequences for their present and future lives. Unfortunately, there is still little robust evidence showing the causal impact of interventions designed to help individuals growing up or who have grown up in an environment characterized by a substance abuse problem.

The objectives of this project are manifold: 1) to measure the impact of a therapy-based intervention targeting individuals who have grown up in such an environment and are seeking help; 2) to document the indirect effects that such an intervention may have on the family members of the recipients; and 3) to investigate the potentially harmful consequences of the long waiting time that individuals often face (as is the case in Denmark) before they can begin such an intervention.

We provide evidence on these questions through the evaluation of a free therapy-based intervention, taking the form of individual counselling. The intervention is specifically designed to help young individuals who have grown up in a family with an alcohol or drug problem in Denmark. The intervention was offered by TUBA (*Therapy & Counselling for young people, who are Children of Alcohol and Drug abusers*), a large non-profit organization that offers specialized help from 34 centers scattered throughout the country.

To measure the impact of the intervention, we conducted a randomized controlled trial in which people who had signed-up to participate in TUBA's intervention were offered to enroll in our study and randomly selected to start the intervention either immediately (the "treatment" group) or one year later (the "control" group). From an ethical point of view, this design was deemed acceptable because

we focused on people aged 25 to 35 years who registered in one of the four most populated cities in Denmark (Copenhagen, Odense, Aalborg, and Aarhus) where the current waiting time exceeds one year. In total, 358 people participated in the study. It should be noted that because individuals were randomly assigned to the treatment and control groups, their characteristics are comparable at enrolment and, therefore, any differences observed between these two groups after the start of the intervention can be attributed to the effect of TUBA's intervention.

The purpose of this intermediate report is to document the impact of TUBA's intervention on participants' psychological health one year after randomization, that is, one year after the start of the intervention for treated individuals and just before the start of the intervention for control individuals. To do this, we use survey data collected from 95.5% of them (attrition is balanced between groups).

We reach several conclusions:

- Confirming that there is a high demand for the type of interventions offered by TUBA in our target population, we find that participation in the intervention is extremely high: 98.1% of treatment individuals who could be reached at follow-up participated in TUBA's intervention. The intensity of the intervention is relatively limited as treatment individuals report having received only 13.6 one-hour sessions in the past 12 months.

- Despite the fact that a large proportion of the individuals in the control group received some form of assistance (48.5% of them report having received another form of help in the past 12 months and, in particular, 21.9% of them report having received help from a psychologist), we find that TUBA's intervention has a relatively large positive effect on respondents' psychological health. In particular, we find that well-being is 0.5 standard deviation higher in the treatment group than in the control group. Similarly, scores of depression, psychological distress and post-traumatic stress disorder are 0.29 to 0.36 standard deviation lower in the

treatment group than in the control group. As a result, the intervention increases the proportion of individuals without depression from 35.5% to 57.4% and the proportion of individuals without current mental well-being problems from 33.7% to 52.7%.

ACKNOWLEDGMENTS

Financial support from Det Obelske FamilieFond, Lundbeck Fonden, Novo Nordisk Fonden, and TrygFonden is gratefully acknowledged. We would like to thank TUBA (and, in particular, Henrik Appel, Alex Kastrup Nielsen, Marc Krabbe-Sørensen, and Helle Lindgaard) for making this impact evaluation possible. We are also grateful to Sanne Dalgaard Toft who provided outstanding project management. Finally, we would also like to thank seminar participants at Aarhus University for their useful comments. Last but not least, we would like to thank all participants for taking part in the study. The study received IRB approval from Aarhus University (serial number 2019-17) and was registered in the AEA RCT Registry (AEARCTR-0005776). The usual disclaimer applies.

1. INTRODUCTION

In Denmark, alcohol and drug abuse problems are widespread issues and, as a consequence, a large number of individuals are growing up or have grown up in families in which such problems exist or existed. For instance, using survey data from the Danish Health Interview Survey 2005, Hansen et al. (2011) estimated that in 2005, 20% of the Danish population were heavy drinkers (between 672,002 and 1,195,069 individuals), 14% had harmful alcohol use (between 439,221 and 944,992) and 3% were dependent drinkers (between 118,196 and 188,384). The extent of the problem is such that 632,000 individuals in Denmark are growing up or have grown up in families with alcohol problems, 122,000 of which are young individuals between 0-18 years old currently growing up in such families (Kristiansen et al., 2008).[1]

Unfortunately, growing up in a family with alcohol and/or drug problems has severe consequences for individuals' current and future life. For instance, existing research shows that growing up in families with an alcohol problem is correlated with an increased risk of anxiety, depression, suicidal behavior, low self-esteem, eating disorders, self-abuse, and post-traumatic stress disorder (Christensen and Bilenberg, 1999; Lindgaard, 2002, 2005), which in turn translate into self-destructive behaviors (e.g. attempted suicide or substance use disorders) and increased mortality (Christoffersen et al., 2003). Furthermore, the indirect cost of the problem for the society is also believed to be high. For instance, about 40% of individuals in alcohol treatment grew up in families with an alcohol problem (Nielsen et al., 2000) and individuals with alcohol problems make use of the health system up to four times more than the general population (Lennox, 1992). Also, 40% of

[1] A recent study finds that survey data underestimate the true prevalence of substance use disorders among parents (Frederiksen et al., 2021), suggesting that actual prevalence rates may be higher than those reported in the body of the text.

children placed outside the family home by Child Protective Services are there due to parental alcohol or drug problems according to public records (Harwin et al., 2010).

Despite the high prevalence of the problem and its dramatic consequences, there is still little robust evidence showing the causal impact of interventions designed to help individuals growing up or who have grown up in such an environment. Consequently, there is little evidence in favor or against government subsidization of this type of psychological health support interventions. There is also limited evidence of the potentially harmful consequences of the substantial waiting time individuals often face (as is the case in Denmark) before they can start such an intervention. Finally, there is also limited robust evidence of any indirect effects these interventions may have on the recipients' family members, which, in turn, may lead to a potentially large underestimation of these interventions' true impact.

Through this research project, we provide new evidence on these questions through the evaluation of a free therapy-based intervention, taking the form of individual counselling. The intervention is specifically designed to help young individuals who have grown up in a family with an alcohol or drug problem in Denmark. The intervention was offered by TUBA (*Therapy & Counselling for young people, who are Children of Alcohol and Drug abusers*), a large non-profit organization that offers specialized help from 34 centers scattered throughout the country. Information on the content of the intervention offered by TUBA is provided in *Section 5. Materials and Methods*.

As there were strong reasons to expect substantial selection with respect to the type of individuals who would enroll in the program studied (as well as in the type of individuals who would enroll earlier or later), we implemented a randomized controlled trial to measure its impact on recipients. Our experiment combines the features of phase-in and rotation designs in that four batches of clients were recruited over a 13-month period and their therapy start date randomized – with treatment

respondents starting immediately and control respondents starting a year after. This ensured that, within our sample, individuals receiving the intervention were comparable to those not receiving it at the time of the draw and that, therefore, any differences found between the two groups of individuals at follow-up could be interpreted as the impact of the intervention. From an ethical point of view, this design was deemed acceptable because we focused on people aged 25 to 35 years who registered in one of the four most populated cities in Denmark (Copenhagen, Odense, Aalborg, and Aarhus) where the current waiting time exceeds one year. It was also made possible thanks to several foundations which, through their donation, allowed the recruitment of extra therapists, thus increasing the number of TUBA recipients without altering the waiting time for non-participants or the quality of the therapy. The project time is summarized in *Fig. 1* and additional information on the experimental design is provided in *Section 5. Materials and Methods*. We measure the impact of the intervention on participants' psychological health one year after treatment respondents began therapy and just before those in the control group begin theirs.

2. DATA

Data was collected on both control and treatment respondents at two different points in time. Respondents were first surveyed when they enrolled to receive TUBA's intervention. At that point, we collected background information and assessed their psychological health (this round of data collection is hereafter referred to as "*baseline*"). Second, respondents were surveyed again one year after treatment respondents began therapy, immediately prior to the time at which those in the control group would begin their therapy (this round of data collection is hereafter referred to as "*follow-up*"). Here, we collected information on their current psychological health, as well as on a limited set of related outcomes (e.g. alcohol consumption, occupation, family status, COVID-related anxiety, etc.).

Both at baseline and follow-up, current psychological health was measured using a range of internationally validated self-report questionnaires. First, we measure respondents' well-being using the World Health Organization-Five Well-Being Index (hereafter referred to as "WHO-5") which is a short measure designed to capture overall current mental well-being (WHO, 1988; Topp et al., 2015). The index comprises five items. Respondents reply to these items by indicating how often (all the time, most of the time, more than half the time, less than half the time, some of the time, never) they experienced various feelings in the past two weeks (e.g. "I felt happy and in a good mood", “I have felt active and vigorous”, etc.). The total score ranges from 0 to 100 (each question contributing equally), with lower scores indicating more severe mental well-being problems.

Second, we measure respondents’ depressive symptoms using the Major Depression Inventory (hereafter referred to as "MDI"), which is used to identify the presence of major depression according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) and the ICD-10, and the degree of depression severity (Bech et al. 2001; Cuijpers, 2007; Olsen et al., 2003; Konstantinidis et al., 2011; Bech et al., 2015). The scale comprises 12 questions combined into 10 items. The MDI asks how often individuals have experienced various negative feelings and events (e.g., "feeling low in spirit or sad," "suffered decreased appetite," etc.). The set of possible responses and recall period are identical to those used by the WHO-5. The total score ranges from 0 to 50 (each item contributes equally). Higher scores on the MDI indicate more severe depressive problems.

Third, we measure respondents’ psychological distress with the help of the Clinical Outcomes in Routine Evaluation-Outcome Measure (hereafter referred to as "CORE") which measures the degree of psychological distress (Evans et al, 2000; Evans et al., 2002; Barkham et al., 2006). The scale comprises 34 items. Clients respond by indicating how often (most or all the time, often, sometimes, only occasionally, not at all) they have experienced various feelings in the past week (e.g. "I have felt terribly alone and isolated", “I have felt overwhelmed by my problems”, etc.). The total score ranges

from 0 to 40 (each question contributing equally), with higher scores indicating more severe psychological distress.

Finally, we measure whether or not respondents exhibit post-traumatic stress disorder (PTSD) using the International Trauma Questionnaire (hereafter referred to as "ITQ") (Hyland et al., 2017; Cloitre et al., 2018; Hansen et al., 2021). The scale comprises six questions. The ITQ begins by asking respondents to identify the experience that troubles them the most and to answer how much (not at all, a little, moderately, quite a bit, extremely) respondents have been bothered by various issues in the past month as a result of the mentioned troubling experience (e.g. "having upsetting dreams that replay part of the experience or are clearly related to the experience", “Avoiding external reminders of the experience (for example, people, places, conversations, objects, activities, or situations)”, etc.). The total score ranges from 0 to 24 (each question contributing equally), with higher scores indicating more severe psychological distress as a result of a traumatic experience.

For each of these tools, the combination of total score and validated thresholds can be used to diagnose psychological health problems. For example, the MDI score classifies individuals into the following four categories: 1) not suffering from depression, 2) suffering from mild depression, 3) suffering from moderate depression, 4) suffering from severe depression.

In *Table S.1 (Supplementary Materials)*, we show that these measures, although related, do not capture exactly the same information. For example, 23% of respondents with PTSD have no signs of depression and 48% of people without PTSD have some form of depression.

3. RESULTS

We now turn to the impact of the intervention. As a preliminary step, we describe respondents' actual participation in counselling interventions (including TUBA and others) by randomization status.

Then, we analyze the impact of randomization on respondents' psychological health using the different tools described above. Finally, we investigate whether or not the intervention had any impact on individuals' daily life.

As described in *Section 5. Materials and Methods*, we were able to successfully survey 342 out of 358 respondents as part of the follow-up survey, *i.e.* 95.5% of the respondents included in our sample (with attrition balanced between groups). Estimates of intervention impact were obtained by regressing the variables of interest on a constant, a binary variable indicating whether or not individuals were selected to receive the intervention immediately (treatment individuals), and a set of strata fixed effects (*Panel A*). We also present the coefficients obtained when selected sets of covariates (identified by a double-lasso procedure) are added to the estimated equation (*Panel B*). Because the two sets of results are very similar (demonstrating that treatment and control individuals were comparable before the intervention began), we comment only on those presented in *Panel B.*

The main results are illustrated by histograms placed in the body of the text. The associated regression tables are reported below in the Tables section.

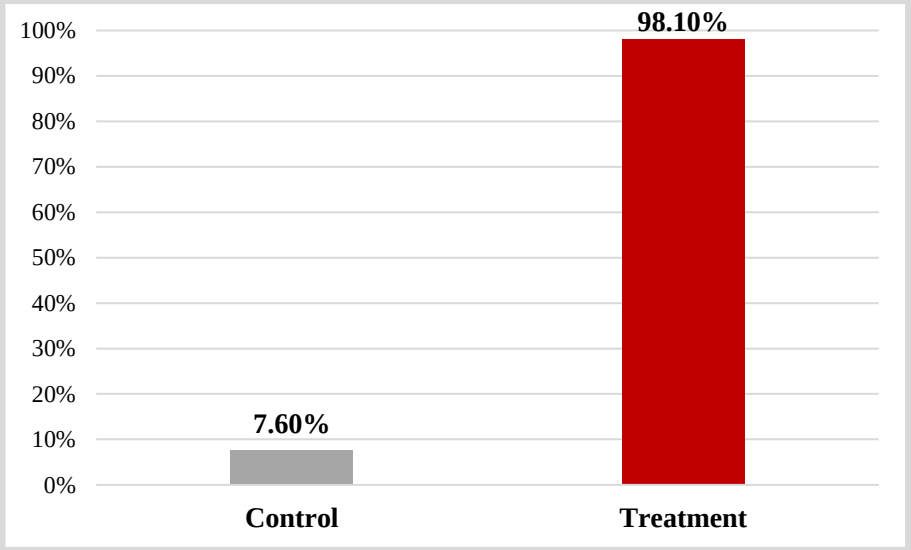
3.1 Participation in counselling interventions

The first key finding is that, within the target population, demand is remarkably high for the type of specialized help offered by TUBA. As displayed in *Table 2*, 98.1% of treated individuals who could be reached at follow-up participated in the intervention offered by TUBA. Interestingly, the intensity of the intervention is not that high, as treatment respondents received only 13.6 one-hour sessions over the 12-month period. Also note that, at follow-up, 69.5% of the treatment respondents were not done with their therapy and were still receiving help from TUBA.

As expected, the proportion of respondents in the control group who actually received the TUBA intervention is much lower (7.6%). However, it is important to note for the interpretation of the results

discussed later that, as displayed in *Table 3*, almost half of the controls (48.5%) received some form of formal psychological help (other than TUBA) in the 12 months before follow-up. The most commonly used psychological help was from a psychologist (21.9% of controls) or a psychotherapist (7.1% of controls). The corresponding shares are significantly lower in the treatment group (32.2% any help, 9.1% from a psychologist, and 1.1% from a psychotherapist).

Hist. 1: Participation in TUBA's intervention

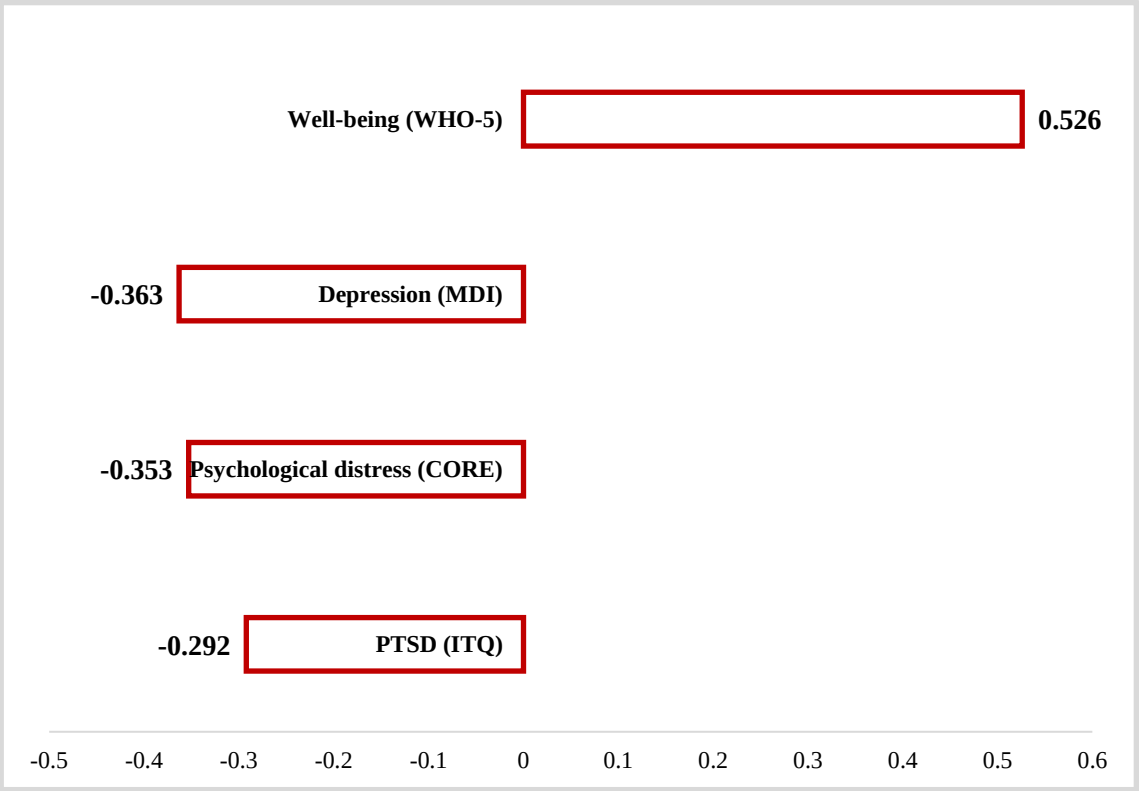


Notes: In this histogram, we present the share of respondents who report having participating in TUBA's intervention in the past 12 months in the treatment and control groups. The share of respondents who participated in TUBA's intervention is remarkably high in the treatment group suggesting that, within the target population, demand is high for the type of intervention offered by TUBA.

3.2 Psychological health

The second key result is that TUBA's intervention had a relatively large positive effect on respondents' psychological health at follow-up (i.e. 12 months after randomization). As displayed in *Table 4*, we find that respondents' well-being is 0.5 standard deviation higher in the treatment group than in the control group. Similarly, scores of depression, psychological distress and post-traumatic stress disorder are between 0.29 and 0.36 standard deviation lower in the treatment group, all suggesting that the intervention offered by TUBA significantly improves the psychological health of recipients.

Hist. 2 : Impact on respondents' psychological health (standardized effects)



Notes: In this histogram, we report on the impact of the intervention on respondents' psychological health as measured by four internationally validated self-report questionnaires. World Health Organization-Five Well-Being Index (WHO-5) captures overall current mental well-being. The Major Depression Inventory (MDI) identifies the presence of major depression. The Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE) measures the degree of psychological distress. The International Trauma Questionnaire (ITQ) measures whether or not respondents exhibit Post-Traumatic Stress Disorder (PTSD). Note that while lower WHO-5 scores indicate more severe mental well-being problems, lower MDI, CORE, and ITQ scores indicate less severe psychological health problems. These results suggest that TUBA's intervention had a relatively large positive effect on respondents' psychological health at follow-up.

Next, for each tool, we then combine scores and validated cutoffs to study the impact of the intervention on the prevalence of each possible diagnosis. As displayed in *Table 5*, at follow-up, the proportion of control group respondents with current poor mental well-being is 66.3%, at least mild depression is 64.5%, at least mild psychological distress is 88.4%, and post-traumatic stress disorder is 40.7%. The intervention of TUBA allowed to substantially decrease these figures in the treatment group. They are 47.3%, 42.6%, 73.1%, and 28.0% respectively.

In *Table 6*, we investigate the heterogeneity of the intervention's impact between subsets of respondents: (a) individuals under age 30 at baseline vs. older individuals; (b) male vs. female respondents; and (c) half of the respondents with the lowest mental well-being at baseline vs. half of

the respondents with the highest mental well-being at baseline. We find no difference between groups, suggesting that the intervention works equally well for all groups compared.

These results are encouraging and indicate that the intervention offered by TUBA has a significant positive effect on its recipients. We emphasize that the effects we measure do not capture the impact of benefiting from the intervention offered by TUBA compared to a situation in which non-recipients would do nothing. Instead, our coefficients measure the impact of the intervention relative to what non-recipients do when they cannot immediately benefit from the intervention – and, as we described in the previous subsection, nearly half of them end up receiving some alternative form of assistance.

3.3 Other impacts

These positive effects on mental health already appear to be having a positive impact on the daily life of these individuals. Although we will have to wait until we have access to the administrative (i.e. register based) data for a more comprehensive analysis of the consequences of the intervention on the lives of the recipients (beyond that on their psychological health), a few questions asked in the follow-up questionnaire already allow us to provide some initial answers.

As *Table 7* shows, while we find no effect of the intervention on our very crude measure of respondents' main source of income, we do find that the improvement in recipients' mental health appears to have already had significant impacts on their daily lives. First, we observe that individuals in the treatment group are less likely to have a problem with alcohol (as measured by the CAGE-C test). Interestingly, we find that the change is not in the amount of alcohol consumed – which is 4.3 drinks in the last 30 days[2] in the control group and 4.8 in the treatment group – but rather in how comfortable they are with the amount of alcohol they consume. In particular, they are less likely to

[2] Note that a "drink" was defined as a can or bottle of beer, a glass of wine or a wine cooler, a shot of liquor, or a mixed drink with liquor in it. We did not take into account times when respondents only had a sip or two from a drink.

report feeling embarrassed by anyone criticizing their drinking or that they have ever felt bad or guilty about their drinking.

Second, although this result should be taken with caution and will need to be confirmed by administrative data, the survey data indicate that the intervention increased the likelihood of having children, suggesting that better psychological health or its prospect increases willingness to have children.

Finally, we find that individuals in the treatment group experienced less stress due to the COVID pandemic. Indeed, while only 11.7% of individuals in the control group report not being anxious at all about the COVID pandemic, the corresponding share is 22.8% in the treatment group – an increase of 94,9%.

4. NEXT STEPS

In the coming months, survey data will be linked to administrative data, which will allow us to measure the impact of the intervention on respondents on a broader range of outcomes. These include health use outcomes (such as the number and duration of hospitalizations, etc.), labor outcomes (such as employment status, number of days worked, number of absences from work, earnings and the types and amounts of social transfers received, etc.) and crime outcomes (number of charges, crimes, and punishments – as well as type). We will use similar data to explore the impact of the intervention on respondents' partners and children, for whom the evidence is currently limited (Brummer et al., 2021).

In a few years, respondents will also be surveyed again with the purpose of measuring the impact of long waiting lists/earlier intervention. In March 2022, all participants will have received the intervention, but those in the treatment group will have received it one year earlier. We will also use

administrative data to measure the impact of receiving the intervention early vs. waiting one year on participants' health, labor market, and criminality, as well as that of their partners and children.

5. MATERIALS AND METHODS

5.1 TUBA's intervention

TUBA's approach is holistic, client-centered, and focuses on resolving the problems faced by the individual that prevent them from living a satisfying life. This is defined as a life with a well-being comparable to that of the normal population and an ability to function in daily life with regard to education, work, intimate relationships, etc.

Assessment session

Individuals seeking help from TUBA begin by participating in an assessment session. This session consists of an interview in which the client's family background, education, current problems and reasons for seeking help are discussed. At the end of the session, the client answers a number of questions about his or her history, substance use, etc., which helps the therapist understand the individual's needs. At the end of the assessment session, the therapist and client agree on the type of intervention that would be most appropriate and the client is placed on a waiting list.

Treatment process

At the beginning of the intervention, the client's goals are explored, discussed and agreed upon in order to establish a direction for the treatment. A theory of change, from both the therapist's and the client's perspective, is explored, discussed, negotiated and finally agreed upon. Next, specific methods that may be used in therapy are explained by the therapist for acceptance by the client. Finally, the goals, the theory of change and the methods are combined in writing in a treatment plan. If new goals emerge during the therapy, or if goals are changed, they are documented in the treatment plan.

Individuals meet regularly with their therapist on a mutually agreed upon schedule. The meetings are in person and take place in the TUBA center. Exceptionally, some of these meetings were held online during the height of the COVID pandemic.

5.2 Experimental Design

Study population

Inclusion criteria: Potential participants were eligible for the study if they sought help from one of TUBA's four largest centers (Copenhagen, Odense, Aalborg, and Aarhus), were between 25 and 35 years old, and would have been put on a waitlist for treatment (minimum one year), had they not been able to participate in the study.

Exclusion criteria: Potential participants were excluded, if they a) did not consent to participate in the study; b) were unable to wait one year to receive treatment; c) were not looking for a full course of therapy but rather a few counselling sessions; d) were found not to have sufficient psychological strain or functional impairment to be in need of treatment; e) had alcohol or drug consumption usage that was so problematic that they could not benefit from therapeutic treatment until they had received help to reduce their substance abuse; f) took so much medicine that they could not benefit from therapeutic counselling; g) were currently psychotic.

For both ethical, logistical, and statistical reasons, the study only focused on individuals aged between 25 and 35 who sought help in one of TUBA's four largest centers (Copenhagen, Odense, Aalborg, and Aarhus) because the time they had to wait to benefit from this intervention exceeded a year there. Together, they represent 56% of TUBA's clients. Note that those under 25 years of age could not be included in the study sample because their waiting time was significantly reduced thanks to specific government grants.

Procedure: Following their registration, potential participants met with one of TUBA's therapists who decided on their eligibility and informed them of the purpose of the study (orally and in writing). Participants were informed that study participation is voluntary and that they can withdraw at any time. In total, 358 individuals were included in the study sample.

Characteristics of respondents: In *Table 1*, we describe the average baseline characteristics of the individuals included in our sample who were also surveyed at follow-up (342 out of 358 respondents, *i.e.* 95.5% of the sample studied in this article). The average age is 29.5 years and 77.8% of respondents are women. As expected, our sample is overwhelmingly constituted of young adults who grew up in a difficult family environment: 99.7% of them grew up in a home with an alcohol abuse problem, and 42.4% of them grew up in a home with a drug abuse problem. In both cases, the problem originated, in order of importance, firstly from the biological father, secondly from the biological mother, and lastly from another person in the household. In total, 78.2% report to have experienced violence during childhood and 15.2% report to have attempted suicide. The vast majority of them had already sought psychological help, with 80.7% of them stating that they had already received psychological help ever and 44.8% of them stating that they had received help in the last 12 months.

Trial design

For both financial and logistical reasons, the intervention was rolled out progressively over a period of 13 months. The first batch consisted of eligible individuals currently on the waiting list with an anticipated waiting time over a year. Subsequent batches included eligible individuals who had signed up for TUBA's intervention within a 4-month period on the basis of "first-come, first-enrolled in the study." The first batch included 150 individuals, batches 2 and 3 70 individuals each, and batch 4 68 individuals. Once eligible individuals agreed to participate in the study, they were invited to register by completing our baseline questionnaire.

Upon registration, eligible individuals were randomly allocated to either the treatment or the control group in order to guarantee the comparability of the two groups prior to the intervention roll-out. Randomization was performed on computer by the research team. In all, four random draws were made (one for each batch). These were stratified by center. (Copenhagen, Odense, Aalborg, or Aarhus).

As expected, due to randomization, the characteristics of the individuals included in our sample are not correlated with whether respondents were offered enrollment immediately ("T" group for treatment) or in one year ("C" group for control), as also displayed in *Table 1*. Coefficients displayed in the "*Diff. T. vs. C."* column are obtained by estimating equation (1) using successively each of the baseline characteristics displayed in the left column of the table as the dependent variable. We do so using all observations for which baseline information are available. The point estimates associated with the treatment variables remain small and non-significant, suggesting again that respondents' randomization status is uncorrelated with their baseline characteristics.

Also, as it is not possible for TUBA to deny eligible individuals treatment for an indefinite period of time, the RCT followed a rotation design, according to which only the treatment start date varied across individuals at random. While individuals in the treatment groups immediately had access to the intervention, those in the control groups had to wait for a year. The draw was stratified by treatment site.

Importantly, the workload arising from the implementation of the intervention was shared by all therapists in the four study sites. In turn, this ensured that the study results are representative of TUBA's work and not merely of a few therapists.

5.3 Statistical Analysis

Specification: In order to assess the impact of the intervention on an outcome (y_{it}) measured at time *t*, we estimate the following equations:

$$y_i = \alpha + \beta_1 T_i + \mu_i + \varepsilon_i \qquad \textbf{Eq. (1)}$$

$$y_i = \alpha + \beta_1 T_i + \mu_i + X_i\beta + \varepsilon_i \qquad \textbf{Eq. (2)}$$

T_i is a dummy variable indicating whether or not respondent i was randomly selected to have immediate access to the intervention. μ_i is a vector of strata fixed effects, which were obtained by interacting of a set of dummy variables indicating the site where respondents sought treatment (Copenhagen, Odense, Aalborg, or Aarhus) and dummy variables indicating their batch number (1, 2, 3, or 4). A vector of baseline covariates X_i containing respondents' pre-randomization background information was also selected using a double lasso procedure and added to the regression (Belloni et al., 2014).

To better understand our main results, we also investigate the impact of the intervention on different subgroups of respondents: (a) individuals under age 30 at baseline vs. older individuals; (b) male vs. female respondents; and (c) half of the respondents with the lowest mental well-being at baseline vs. half of the respondents with the highest mental well-being at baseline. For each of these subgroups, we measure the impact of the intervention estimating the following equation:

$$y_i = \alpha + \beta_{1,1}\,(Group_{1,i} * T_i) + \beta_{1,2}\left(Group_{2,i} * T_i\right) + \gamma Group_{1,i} + \mu_i + X_i\beta + \varepsilon_i \qquad \textbf{Eq. (3)}$$

In this equation, $Group_{1,i}$ and $Group_{1,i}$ dummy variables indicating whether respondent *i* belongs to *group 1* or *group 2* (e.g. female and male respondents). Again, X_i is a vector of baseline covariates selected using a double-lasso procedure.

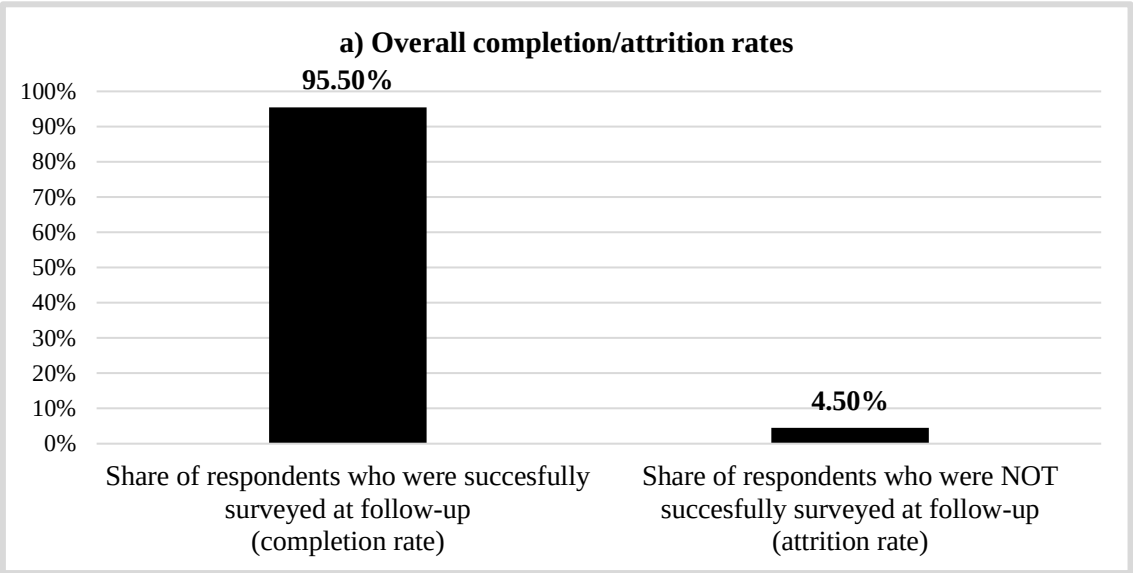
Standard errors: We compute robust standard errors.

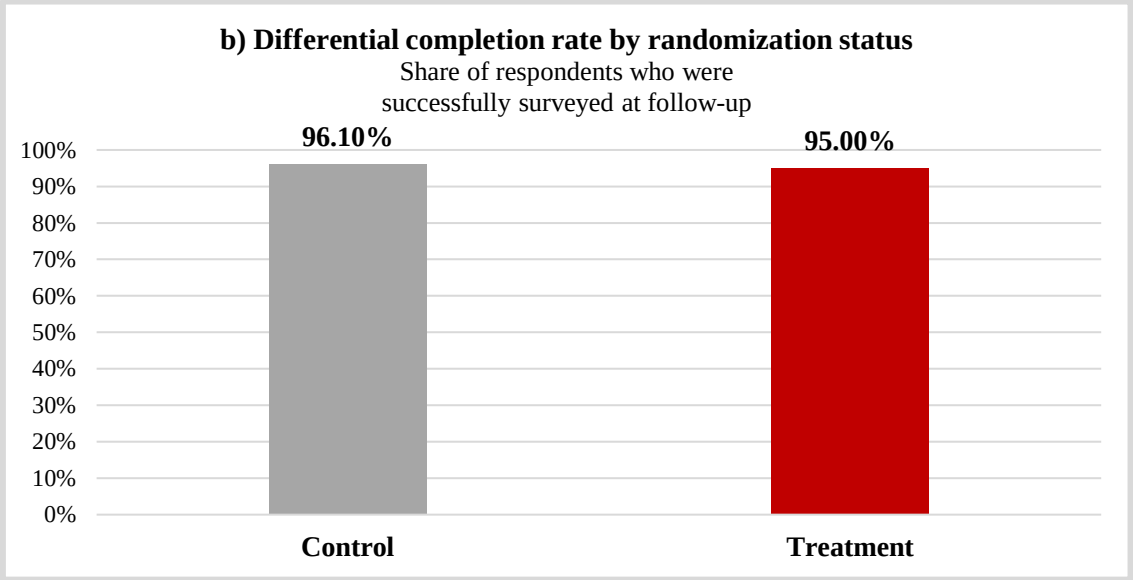
Procedures for dealing with missing data: We did not perform any imputation for missing data from item non-response.

5.4 Attrition

As part of the follow-up survey, we were successfully able to survey 342 out of 358 respondents, *i.e.* 95.5% of the respondents included in our sample. We estimate the balancing of the attrition rate across groups estimating equation (1). We make this comparison first for the sample as a whole (Panel A) and then for each batch separately (Panel B). In all cases, attrition appears independent of respondents' treatment status, as detailed in *Table S.2 (Supplementary Materials)*: coefficients on treatment dummies are small and non-significant.

Hist. 3: Completion and attrition rates





Notes: In these histograms, we present the share of respondents who were successfully surveyed at follow-up (the completion rate) and the share who were not (the attrition rate), as well as the completion rates in the treatment and control groups. The overall completion rate is remarkably high and similar in both groups, suggesting that attrition problems do not affect the representativeness of the sample or the comparability of the treatment and control groups.

REFERENCES

Barkham, M., Mellor-Clark, J., Connell, J., & Cahill, J. (2006). A core approach to practice-based evidence: A brief history of the origins and applications of the CORE-OM and CORE System. *Counselling and Psychotherapy Research*, *6*(1), 3-15.

Bech, P., Rasmussen, N. A., Olsen, L. R., Noerholm, V., & Abildgaard, W. (2001). The sensitivity and specificity of the Major Depression Inventory, using the Present State Examination as the index of diagnostic validity. *Journal of affective disorders*, *66*(2-3), 159-164.

Bech, P., Timmerby, N., Martiny, K., Lunde, M., & Soendergaard, S. (2015). Psychometric evaluation of the Major Depression Inventory (MDI) as depression severity scale using the LEAD (Longitudinal Expert Assessment of All Data) as index of validity. *BMC psychiatry*, *15*(1), 1-7.

Belloni, A., Chernozhukov, V. and Hansen, C. (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies*, *81*(2), pp.608-650.

Brummer, J., Hesse, M., Frederiksen, K. S., Karriker-Jaffe, K. J., & Bloomfield, K. (2021). How do register-based studies contribute to our understanding of alcohol's harms to family members? A scoping review of relevant literature. *Journal of Studies on Alcohol and Drugs*, *82*(4), 445-456.

Christensen, H. B., & Bilenberg, N. (2000). Behavioural and emotional problems in children of alcoholic mothers and fathers. *European child & adolescent psychiatry*, *9*(3), 219-226.

Christoffersen, M. N., & Soothill, K. (2003). The long-term consequences of parental alcohol abuse: a cohort study of children in Denmark. *Journal of substance abuse treatment*, *25*(2), 107-116.

Cloitre, M., Shevlin, M., Brewin, C.R., Bisson, J.I., Roberts, N.P., Maercker, A., Karatzias, T. & Hyland, P., (2018). The International Trauma Questionnaire: Development of a self-report measure of ICD-11 PTSD and complex PTSD. *Acta Psychiatrica Scandinavica*, *138*(6), 536-546.

Cuijpers, P., Dekker, J., Noteboom, A., Smits, N., & Peen, J. (2007). Sensitivity and specificity of the Major Depression Inventory in outpatients. *BMC psychiatry*, *7*(1), 1-6.

Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: clinical outcomes in routine evaluation. *Journal of Mental Health*, *9*(3), 247-255.

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: Psychometric properties and utility of the CORE–OM. *The British Journal of Psychiatry*, *180*(1), 51-60.

Frederiksen, K. S., Hesse, M., Grittner, U., & Pedersen, M. U. (2021). Estimating perceived parental substance use disorder: Using register data to adjust for non-participation in survey research. *Addictive Behaviors*, *119*, 106897.

Harwin, J. P., Madge, N. P., & Heath, S. (2010). Children affected by parental alcohol problems (ChAPAPs).

Hyland, P., Shevlin, M., Brewin, C.R., Cloitre, M., Downes, A.J., Jumbe, S., Karatzias, T., Bisson, J.I. and Roberts, N.P. (2017). Validation of post-traumatic stress disorder (PTSD) and complex PTSD using the International Trauma Questionnaire. *Acta Psychiatrica Scandinavica*, *136*(3), pp.313-322.

Konstantinidis, A., Martiny, K., Bech, P., & Kasper, S. (2011). A comparison of the Major Depression Inventory (MDI) and the Beck Depression Inventory (BDI) in severely depressed patients. *International Journal of Psychiatry in Clinical Practice*, *15*(1), 56-61.

Kristiansen, L., & Ekholm, O., & Grønbæk, M., & Tolstrup, J. (2008). Danskernes alkoholvaner (Drinking in Denmark) København: Sundhedsstyrelsen

Lennox, R. D., Scott-Lennox, J. A., & Holder, H. D. (1992). Substance abuse and family illness: Evidence from health care utilization and cost-offset research. *The journal of mental health administration*, *19*(1), 83-95.

Lindgaard, H. (2002). *Voksne børn fra familier med alkoholproblemer: mestring og modstandsdygtighed*. Center for Rusmiddelforskning, Aarhus Univ..

Lindgaard, H. (2005). Adult children of alcoholics: Are they any different? Nordisk Psykologi, 57(2): 107-129.

Hansen, M., Vægter, H. B., Cloitre, M., & Andersen, T. E. (2021). Validation of the Danish International Trauma Questionnaire for posttraumatic stress disorder in chronic pain patients using clinician-rated diagnostic interviews. *European journal of psychotraumatology*, *12*(1), 1880747.

Nielsen, A. S., Nielsen, B., & Petersen, P. (2000). Behandlingsbehov hos alkoholmisbrugere, hvis forældre havde et alkoholmisbrug. *Ugeskrift for læger*, *162*(34), 4540-44.

Olsen, L. R., Jensen, D. V., Noerholm, V., Martiny, K., & Bech, P. (2003). The internal and external validity of the Major Depression Inventory in measuring severity of depressive states. *Psychological medicine*, *33*(2), 351-356.

Topp, C. W., Østergaard, S. D., Søndergaard, S., & Bech, P. (2015). The WHO-5 Well-Being Index: a systematic review of the literature. *Psychotherapy and psychosomatics*, *84*(3), 167-176.

World Health Organization. (1998). Wellbeing measures in primary health care/the DEPCARE project: report on a WHO meeting, Stockholm, Sweden 12-13 February 1998. In *Wellbeing measures*

in primary health care/the DEPCARE project: report on a WHO meeting, Stockholm, Sweden 12-13 February 1998.

FIGURES

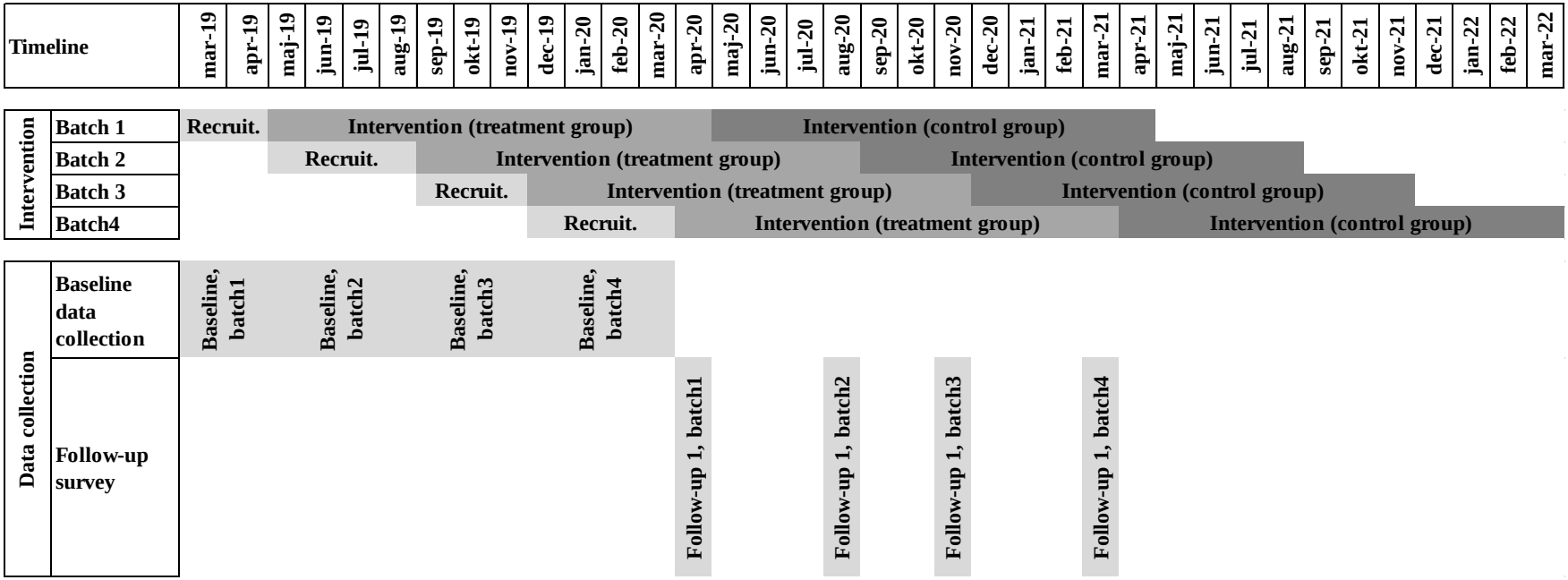


Fig. 1: Project timeline. For both financial and logistical reasons, the intervention was rolled out progressively over a period of 13 months. The first batch consisted of eligible individuals currently on the waiting list with an anticipated waiting time over a year. Subsequent batches included eligible individuals who had signed up for TUBA's intervention within a 4-month period on the basis of "first-come, first-enrolled in the study." The first batch included 150 individuals, batches 2 and 3 70 individuals each, and batch 4 68 individuals.. Once eligible individuals agreed to participate in the study, they were invited to register by completing our baseline questionnaire. For each batch, follow-up data was collected one year after treatment respondents began therapy and just before those in the control group begin theirs.

TABLES

Table 1: Sample description

Variables	# Obs. in total	Control group Mean	Sd	Diff. T. vs. C. Coeff.	S.e.	
Panel A: Background characteristics						
Female	341	0.778	0.417	-0.068	0.048	
Age	342	29.494	3.238	0.025	0.318	
Is in a relationship	342	0.651	0.478	0.021	0.051	
Has children	342	0.244	0.431	-0.015	0.046	
Occupation						
In paid work	342	0.541	0.500	-0.063	0.054	
In education	342	0.273	0.447	-0.029	0.047	
Other occupation	342	0.186	0.390	0.091	0.046	**
Panel B: Information on the problem for which help is sought						
Grew up in problematic family	342	0.983	0.131	0.011	0.012	
Alcohol abuse	342	0.977	0.151	0.011	0.014	
Biological father	342	0.709	0.455	0.023	0.049	
Biological mother	342	0.448	0.499	-0.058	0.053	
Other	342	0.302	0.461	-0.060	0.048	
Drug abuse	342	0.424	0.496	-0.092	0.052	*
Biological father	342	0.215	0.412	-0.085	0.041	**
Biological mother	342	0.145	0.353	-0.018	0.038	
Other	342	0.215	0.412	-0.048	0.043	
Age at which they became aware of the problem	299	11.311	4.822	-0.387	0.596	
Experienced violence in childhood	338	0.782	0.414	0.041	0.043	
Ever attempted suicide	339	0.152	0.360	-0.010	0.039	
Panel C: Past psychological help and medication						
Psychological help						
Now	341	0.157	0.365	-0.039	0.038	
In the past 12 months	342	0.448	0.499	0.010	0.054	
Ever	341	0.807	0.396	-0.047	0.045	
Medication						
Now	334	0.151	0.359	-0.012	0.039	
In the past 12 months	336	0.250	0.434	-0.023	0.047	
Ever	342	0.413	0.494	-0.013	0.054	
Panel D: Psychological health (total scores)						
Wellbeing (WHO5, score out of 100)	342	46.93	17.074	2.734	2.011	
Depression (MDI, score out of 50)	342	23.744	10.176	-1.040	1.189	
Psychological distress (CORE, score out of 40)	342	17.221	6.005	-0.829	0.705	
Post-Traumatic Stress Disorder (ITQ, score out of 24)	342	11.361	5.899	-0.695	0.620	

Notes: In this table, we provide a description of individuals included in our control group (C) and compare their characteristics with those of the individuals in the treatment group (T) using data collected at baseline. In order to do so, equation (1) is estimated for each variable displayed in the left column. Robust standard errors were computed. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.

Table 2: Participation in TUBA's intervention

	In the past 12 months			Now
	Received help from TUBA	**Number of sessions**	**Session length (in min.)**	**Still receives help from TUBA**
Panel A: With strata dummies but without additional covariates				
Treatment	0.907***	13.411***	54.668***	0.672***
	(0.023)	(0.512)	(1.333)	(0.038)
P-value	0.000	0.000	0.000	0.000
Panel B: With strata dummies and covariates (selected by double-lasso)				
Treatment	0.905***	13.395***	54.583***	0.672***
	(0.023)	(0.497)	(1.318)	(0.037)
P-value	0.000	0.000	0.000	0.000
Sample size	339	290	297	339
Control mean	0.0760	0.205	2.994	0.0234

Notes: In this table, we describe the average treatment effect of offering individuals to enroll in the program offered by TUBA on respondents' level of exposure to the intervention at follow-up (12 months after the randomization). In order to do so, we estimate equations (1) (*Panel A*) and (2) (*Panel B*) for each of the measure of exposure displayed in the top row of the table. We compute robust standard errors. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.

Table 3: Other form of psychological help received

		Help received					Number of sessions				
	# Obs. in	Control group		Diff. T. vs. C.			Control group		Diff. T. vs. C.		
Variables	total	Mean	Sd	Coeff.	S.e.		Mean	Sd	Coeff.	S.e.	p.v.
Any organization other than TUBA	337	0.485	0.501	-0.163	0.051	***	6.297	13.693	-3.815	1.278	***
Partner, relatives and/or friends	337	0.213	0.411	-0.062	0.041		1.779	5.704	-0.977	0.593	*
ACA/Al-Anon/Alateen	337	0.012	0.108	-0.006	0.010		0.053	0.620	0.255	0.305	
School or social	337	0.006	0.077	0.006	0.010		0.237	3.077	-0.225	0.247	
Municipalities	337	0.006	0.077	0.000	0.008		0.030	0.385	0.152	0.172	
GP	337	0.107	0.309	-0.009	0.033		0.335	1.264	-0.098	0.124	
Psychologist	337	0.219	0.415	-0.128	0.039	***	1.640	4.468	-1.254	0.381	***
Psychiatrist	337	0.047	0.213	-0.013	0.022		0.298	1.708	-0.190	0.145	
Psychotherapist	337	0.071	0.258	-0.060	0.022	***	0.568	3.446	-0.527	0.275	*
Psychiatric daycare	337	0.012	0.108	0.005	0.013		0.592	5.532	-0.545	0.444	
Psychiatric residential facility	337	0.006	0.077	-0.006	0.006		0.018	0.231	-0.019	0.019	
Alcohol treatment (own abuse)	337	0.006	0.077	0.000	0.008		0.030	0.385	-0.030	0.029	
Alcohol treatment (relative)	337	0.012	0.108	-0.006	0.010		0.059	0.633	-0.043	0.054	
Other	337	0.095	0.294	-0.060	0.027	**	1.054	4.643	-0.541	0.585	

Notes: In this table, we describe the average treatment effect of offering individuals to enroll in the program offered by TUBA on respondents' level of exposure to other types of psychological help at follow-up (12 months after the randomization). In order to do so, In order to do so, we estimate equation (2) for each of the measure of exposure displayed in the left column of the table. We compute robust standard errors. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.

Table 4: Impact on psychological health

	Psychological health (standardized scores)			
	WHO-5	MDI	CORE	ITQ
Panel A: With strata dummies but without additional covariates				
Treatment	0.608***	-0.449***	-0.452***	-0.402***
	(0.106)	(0.113)	(0.108)	(0.110)
P-value (unadj.)	0.000	0.000	0.000	0.000
P-value (adj.)				
Panel B: With strata dummies and covariates (selected by double-lasso)				
Treatment	0.526***	-0.363***	-0.353***	-0.292***
	(0.092)	(0.089)	(0.086)	(0.090)
P-value (unadj.)	0.000	0.000	0.000	0.001
P-value (adj.)				
Sample size	337	337	337	337
Control mean (non-standardized scores)	46.13	23.74	17.22	11.36

Notes: In this table, we describe the average treatment effect of offering individuals to enroll in the program offered by TUBA on respondents' psychological health at follow-up (12 months after the randomization). In order to do so, we estimate equations (1) (*Panel A*) and (2) (*Panel B*) for each of the measure of exposure displayed in the top row of the table. Note that positive WHO-5 coefficients and negative MDI, CORE, and ITQ coefficients indicate improvements in psychological health. We use standardized mental health outcomes in the regressions. We compute robust standard errors. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.

Table 5: Impact on psychological health diagnoses

Variables	# Obs. in total	Sample description: Control group Mean	Sd	Impact: Diff. T. vs. C. Coeff.	S.e.	
Panel A: Wellbeing (WHO-5)						
No wellbeing problem	342	0.337	0.474	0.190	0.048	***
Poor wellbeing	342	0.663	0.474	-0.190	0.048	***
Panel B: Depression (MDI)						
No depression	342	0.355	0.480	0.219	0.045	***
Mild depression	342	0.151	0.359	-0.103	0.035	***
Moderate depression	342	0.215	0.412	-0.072	0.034	**
Severe depression	342	0.279	0.450	-0.051	0.041	
Panel C: Psychological distress (CORE)						
No psychological distress	342	0.035	0.184	0.033	0.031	
Low level of problems	342	0.081	0.274	0.119	0.044	***
Mild psychological distress	342	0.256	0.438	0.029	0.045	
Moderate distress	342	0.360	0.482	-0.104	0.044	**
Moderately severe	342	0.192	0.395	-0.103	0.038	***
Severe psychological distress	342	0.076	0.265	0.021	0.021	
Panel D: Post-Traumatic Stress Disorder (ITQ)						
No PTSD problem	342	0.593	0.493	0.127	0.045	***
PTSD	342	0.407	0.493	-0.127	0.045	***

Notes: In this table, we describe the average treatment effect of offering individuals to enroll in the program offered by TUBA on respondents' psychological health at follow-up (12 months after the randomization). In order to do so, equation (2) is estimated for each variable displayed in the left column. We use a double-lasso procedure to select covariates. We compute robust standard errors. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.

Table 6: Impact on psychological health by subgroups

	Participation		Psychological health			
	# Sessions	Still receives help	WHO-5	MDI	ITQ	CORE
Panel A: Age						
G1: Below 30	13.475 ***	.665 ***	.358 ***	-.257 **	-.325 ***	-.325 ***
	(.709)	(.051)	(.129)	(.121)	(.118)	(.118)
G2: Above 30	13.307 ***	.681 ***	.734 ***	-.487 ***	-.379 ***	-.38 ***
	(.701)	(.055)	(.132)	(.134)	(.136)	(.129)
Testing G1=G2						
Chi2-stat	0.028	0.047	4.025	1.602	0.088	0.099
P-value	0.867	0.829	0.045	0.206	0.767	0.754
Sample size	290	339	337	337	337	337
Control mean (G1)	0.245	0.011	0.159	-0.057	-0.110	-0.017
Control mean (G2)	0.156	0.039	-0.199	0.072	0.138	0.021
Panel B: Gender						
G1: Women	13.743 ***	.716 ***	.534 ***	-.343 ***	-.281 ***	-.304 ***
	(.604)	(.042)	(.103)	(.1)	(.1)	(.097)
G2: Men	12.728 ***	.567 ***	.528 ***	-.426 **	-.526 ***	-.485 ***
	(.911)	(.076)	(.205)	(.188)	(.192)	(.186)
Testing G1=G2						
Chi2-stat	0.842	2.862	0.001	0.155	1.271	0.746
P-value	0.359	0.091	0.981	0.693	0.260	0.388
Sample size	289	338	336	336	336	336
Control mean (G1)	0.212	0.023	0.033	-0.019	-0.045	-0.027
Control mean (G2)	0.132	0.026	-0.121	0.079	0.149	0.105
Panel C: Mental health at baseline						
G1: Bottom 50%	14.362 ***	.781 ***	.609 ***	-.364 ***	-.397 ***	-.383 ***
	(.706)	(.046)	(.13)	(.124)	(.126)	(.121)
G2: Top 50%	12.841 ***	.573 ***	.426 ***	-.371 ***	-.304 **	-.331 ***
	(.717)	(.057)	(.134)	(.128)	(.127)	(.125)
Testing G1=G2						
Chi2-stat	2.106	7.826	0.925	0.002	0.263	0.088
P-value	0.147	0.005	0.336	0.967	0.608	0.767
Sample size	290	339	337	337	337	337
Control mean (G1)	0.313	0.021	-0.345	0.266	0.207	0.238
Control mean (G2)	0.067	0.027	0.443	-0.341	-0.265	-0.306

Notes: In this table, we describe the average treatment effect of the intervention on respondents' mental health for different subgroups of our sample. In order to do so, we estimate equation (3) for each pair of subgroups and each dependent variable displayed in top row of the table. Robust standard errors are computed. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.

Table 7: Impact on other outcomes

Variables	# Obs. in total	*Sample description* Control group Mean	Sd	*Impact* Diff. T. vs. C. Coeff.	S.e.	
Panel A: Alcohol problem						
Has an alcohol problem now	341	0.058	0.2347	-0.0243	0.02	
CAGE-C test (pb w/ alcohol)	342	0.233	0.4237	-0.0958	0.036	***
Ever felt you should cut down on your drinking	342	0.238	0.4273	-0.0416	0.037	
Ever been annoyed by anyone critizing your drinking	342	0.081	0.2742	-0.0446	0.025	*
Ever felt bad or guilty about your drinking	342	0.238	0.4273	-0.0888	0.037	**
Ever felt like drinking when waking up	342	0	0	0	0	
Average number of drinking age per week	342	0.948	1.1761	-0.0379	0.095	
Drink outside of meal time	342	0.169	0.3755	-0.0186	0.038	
Talk to a doctor about your drinking	342	0.145	0.3535	-0.0465	0.033	
Panel B: Main source of income						
Paid work	342	0.663	0.4741	-0.0077	0.043	
SU	342	0.221	0.4161	0.0169	0.034	
Benefits	337	0.174	0.3806	-0.0041	0.036	
Panel C: Family						
Is in a relationship	342	0.698	0.4606	-0.0152	0.039	
Has children	342	0.267	0.4439	0.0575	0.023	**
Panel D: Anxieties related to Covid						
Not at all	342	0.117	0.3223	0.1106	0.04	***
A little	342	0.497	0.5015	-0.0755	0.054	
Moderately	342	0.24	0.4282	-0.0318	0.045	
Quite a bit	342	0.099	0.3001	0.0075	0.032	
Extremely	342	0.047	0.2118	-0.0109	0.018	

Notes: We provide a description of individuals included in our control group and compare their characteristics with those in individuals in the treatment group using data collected at baseline. In order to do so, equation (2) is estimated for each variable displayed in the left column. We use a double-lasso procedure to select covariates. We compute robust standard errors. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.

SUPPLEMENTARY MATERIALS

S.1 Correspondence between the different mental health measures

	Wellbeing (WHO-5)		Depression (MDI)				Psychological distress (CORE)						PTSD (ITQ)	
	No problem	Poor wellbeing	No problem	Mild	Mod.	Severe	No problem	Low	Mild	Mod.	Mod. severe	Severe	PTSD	No PTSD
N	358	358	358	358	358	358	358	358	358	358	358	358	358	358
Mean	0.38	0.62	0.41	0.13	0.16	0.30	0.05	0.11	0.23	0.32	0.21	0.07	0.38	0.62
Wellbeing (WHO-5)														
No wellbeing problem			0.71	0.30	0.22	0.04	0.94	0.80	0.56	0.31	0.05	0.00	0.25	0.46
Poor wellbeing			0.29	0.70	0.78	0.96	0.06	0.20	0.44	0.69	0.95	1.00	0.75	0.54
Depression (MDI)														
No depression	0.77	0.19					1.00	0.85	0.67	0.31	0.04	0.00	0.23	0.52
Mild depression	0.10	0.15					0.00	0.10	0.19	0.19	0.07	0.00	0.12	0.14
Moderate depression	0.10	0.20					0.00	0.03	0.12	0.26	0.20	0.08	0.17	0.16
Severe depression	0.03	0.46					0.00	0.03	0.02	0.24	0.69	0.92	0.49	0.18
Psychological distress (CORE)														
No psychological distress	0.13	0.00	0.12	0.00	0.00	0.00							0.01	0.08
Low level of distress	0.23	0.04	0.23	0.09	0.02	0.01							0.03	0.16
Mild distress	0.35	0.17	0.38	0.34	0.17	0.02							0.11	0.31
Moderate distress	0.26	0.36	0.25	0.47	0.52	0.26							0.37	0.30
Moderately severe distress	0.03	0.32	0.02	0.11	0.26	0.49							0.35	0.12
Severe distress	0.00	0.11	0.00	0.00	0.03	0.22							0.14	0.03
Post-Traumatic Stress Disorder (ITQ)														
PTSD	0.25	0.46	0.21	0.34	0.40	0.63	0.06	0.10	0.18	0.43	0.64	0.76		
No PTSD	0.75	0.54	0.79	0.66	0.60	0.37	0.94	0.90	0.82	0.57	0.36	0.24		

Notes: In this table, we show the extent to which each measure of mental health captures a different dimension than the others. For each tool (WHO-5, MDI, CORE, and ITQ), we use the baseline data and indicate the share of respondents falling into each of the different categories that can be constructed from the total score and validated thresholds (e.g., no, mild, moderate or severe depression for the DMI tool). Then, for each category we indicate the share of respondents falling in each of the other categories.

S.2 Attrition

Variables	# Obs. in total	Control group Mean	Sd	Diff. T-C Coeff.	S.e.
Panel A: Attrition rates for the entire sample					
Overall attrition	358	0.039	0.194	0.011	0.022
Panel B: Attrition rates for each batch taken separately					
Batch 1 attrition	150	0.053	0.226	0.000	0.037
Batch 2 attrition	70	0.029	0.169	0.000	0.040
Batch 3 attrition	70	0.057	0.236	0.000	0.056
Batch 4 attrition	68	0.000	0.000	0.060	0.042

Notes: In this table, we provide a description of the non-response rate to the follow-up survey for individuals in the control group (C) and compare it to that of individuals in the treatment group (T). We make this comparison first for the sample as a whole *(Panel A*) and then for each batch separately (*Panel B*). In each case, equation (1) was estimated. Robust standard errors were computed. *, **, *** denote significance at the 10, 5 and 1 percent levels respectively.