This is the accepted version of the following article: Deacon, R.M., Mammen, K., Holmes, J., Dunlop, A., Bruno, R., Mills, L., Graham, R. and Lintzeris, N. (2020), Assessing the validity of the Australian Treatment Outcomes Profile for telephone administration in drug health treatment populations. Drug Alcohol Rev. doi:10.1111/dar.13088 which has been published in final form at www.onlinelibrary.wiley.com

Title: Assessing the validity of the Australian Treatment Outcomes Profile (ATOP) for telephone administration in drug health treatment populations

Running title: Assessing the telephone validity of the ATOP

Authors:
Rachel M. Deacon[1,2,3], PhD, Research Associate
Kristie Mammen[2,3], BPsych (Hons), Program Manager
Jennifer Holmes[3,8], MSc, Senior Program Manager Data and Informatics
Adrian Dunlop[3,5,6,7], MBBS, PhD, Area Director and Senior Staff Specialist
Raimondo Bruno[4,5], PhD, Associate Professor, Psychology
Llewellyn Mills[1,2,3], PhD, Research Fellow
Robert Graham[3,9,10], MBBS, Addiction Medicine Specialist, Conjoint Senior Lecturer
Nicholas Lintzeris[1,2,3], BMedSci MBBS PhD FAChAM, Director, Drug & Alcohol Services; Conjoint Professor

Affiliations:
[1]Sydney School of Medicine (Central Clinical School), Faculty of Medicine and Health, The University of Sydney, NSW, Australia
[2]Drug and Alcohol Services, South Eastern Sydney Local Health District, NSW, Australia
[3]NSW Drug and Alcohol Clinical Research and Improvement Network (DACRIN), NSW, Australia
[4]School of Psychological Sciences, University of Tasmania, Tasmania, Australia
[5]National Drug and Alcohol Research Centre, UNSW Sydney, NSW, Australia
[6]Drug and Alcohol Services, Hunter New England Local Health District, NSW, Australia
[7]Hunter Medical Research Institute, The University of Newcastle, NSW, Australia
[8]Alcohol and other Drugs, Centre for Population Health, Ministry of Health, NSW, Australia
[9]Drug Health, Western Sydney Local Health District, NSW, Australia
[10]School of Medicine, Western Sydney University, NSW, Australia

Corresponding author:
Rachel Deacon, The Langton Centre, 591 South Dowling St, Surry Hills, NSW 2010 Australia. Ph (02) 9332 8777

Conflict of Interest: No competing interests to declare

This is the accepted version of the following article: Deacon, R.M., Mammen, K., Holmes, J., Dunlop, A., Bruno, R., Mills, L., Graham, R. and Lintzeris, N. (2020), Assessing the validity of the Australian Treatment Outcomes Profile for telephone administration in drug health treatment populations. Drug Alcohol Rev. doi:10.1111/dar.13088 which has been published in final form at www.onlinelibrary.wiley.com

Abstract:

Introduction and Aims:

The ATOP is a brief clinical tool measuring recent substance use, health and wellbeing among clients attending AoD treatment services. It has previously been assessed for concurrent validity and inter-rater reliability. In this study we examine whether it is suitable for administration over the telephone.

Design and Methods:
We recruited a sample of 107 AoD clients across public sector specialist AoD treatment services in NSW, Australia between 2016 and 2018. Participants had a mean age of 47 years and 46% were female. Participants completed a face-to-face ATOP, and a phone ATOP with a researcher within five days. Comparisons between the two administration modes were undertaken using Spearman's rank correlation coefficient for continuous or ordinal variables, and Cohen's Kappa for nominal variables.

Results:
Among 107 participants, 59% were attending for alcohol treatment and 41% for opioid treatment. Most ATOP items (76%) reached above 0.7 (good) or 0.9 (excellent) agreement between face to face and telephone use.

Discussions and Conclusions:
Our findings suggest that the ATOP is a suitable instrument for telephone monitoring of recent substance use, health and social functioning among AoD clients. Its validation for remote use over the telephone will support staff to monitor clients' risks and outcomes – of particular relevance in response to the COVID-19 pandemic in which services are increasingly relying on telework approaches to client monitoring.

Keywords: telemedicine; outcome assessment, health care; psychometrics; substance-related disorders

Introduction

The Australian Treatment Outcomes Profile (ATOP) is a validated instrument used for structured brief clinical assessments, risk screening, care planning and patient reported outcome measurement in Alcohol and other Drug (AoD) treatment settings. It examines the domains of substance use, health and wellbeing over a 4-week period [1]. The ATOP was adapted for Australian conditions from the Treatment Outcomes Profile (TOP) developed in the UK [2] and has been implemented in services across Australia. It has demonstrated acceptable inter-rater reliability and concurrent validity against gold-standard instruments among clients attending treatment for alcohol [3], opioid [1, 3] and cannabis use disorders [4] as well as older AoD clients [5]. An advantage of the TOP and ATOP compared to lengthier instruments is their brevity and acceptability for use in clinical settings as part of routine care.

The ATOP was originally developed to be administered in face-to-face interactions during a clinical or research appointment. However, there is a demand for tools that can be reliably used over the telephone, particularly during the current COVID-19 pandemic, in rural settings [6, 7] and for follow-up after discharge. Use of the telephone can be superior to other remote technologies like online self-complete questionnaires, which require internet access and a certain level of literacy not always present in AoD populations [8]. Thus there is a need for the ATOP to be formally assessed for telephone completion.

Other instruments commonly used in AoD treatment populations have been validated for use over the telephone. The alcohol time-line follow-back scale on which the substance use section of the TOP and ATOP is based has been validated for use by telephone and computer [9] against face-to-face, and by telephone against online administration [10]. The SF-36, used for assessing functional physical and mental health [11-17] and the WHOQoL-BREF [18], used to assess patient-reported health related quality of life have also been validated for use via telephone compared to self or clinician completion.

Aim: to assess the concurrent validity of the ATOP for telephone administration against face-to-face administration in an AoD treatment population.

Methods

Participants were enrolled in a larger study (the Clinical Outcomes and Quality Indicators (COQI) study [3]) which aimed to assess the concurrent validity, inter-rater reliability and test-retest reliability of the ATOP in treatment-seeking populations for alcohol or opioid use disorders. They were recruited (written informed consent) from AoD services across South Eastern Sydney, Western Sydney and Hunter New England Local Health Districts between October 2016 and February 2018. Participants attended research and clinical visits at baseline and 4 weeks follow up which included completion of an ATOP.

The proposed gold standard approach for validating different modes of administration of an instrument are outlined in Evans et al [19] and references therein [20-22]. Key methods include: that the same interviewer should be used each time; that the time between applications should be short

enough to limit changes in the participants' behaviour or state, but long enough to allow for participants to 'forget' their previous responses; and the order of application of each mode should be randomised. For the current study, participants attending the follow up interviews were asked to complete another ATOP over the telephone with the same trained researcher, between one and five days before or after the research visit. These limits ensured participants had time to forget their previous responses, but not for their health and wellbeing to change significantly. However, due to operational constraints, the order of ATOP completion (telephone/face to face or face to face/telephone) was unable to be randomised as the research face-to-face interview was tied to completion of a clinical appointment due to the requirements of the larger study. At the start of the phone call, researchers checked participants were on their own and unlikely to be interrupted, to ensure participant confidentiality.

Participants were reimbursed for the research interview and telephone ATOP with a $50 shopping voucher. Ethical approval was provided by the South Eastern Sydney Local Health District Human Research Ethics Committee (16/112 (HREC/16/POWH/222).

Eligibility criteria:

A new or existing client of a participating service, in treatment for either alcohol or opioids (principal drug of concern), and no recent ATOP completion (past 28 days). Clients with severe mental health or cognitive problems that impaired their ability to provide informed consent or participate in interviews were excluded, as were clients considered very likely to be lost to follow-up (e.g. pending incarceration).

Measures

Basic demographic data (age, gender) were collected upon enrolment to the larger study.

The ATOP domains are: days of substance use (alcohol, cannabis, methamphetamines, benzodiazepines, heroin, other opioids (excluding prescribed opioid treatment medications such as methadone or buprenorphine) and cocaine, plus typical amount of alcohol used on a use day); days of paid work and education; any acute housing issue or risk of eviction; caring for children; experience of violence (received and inflicted); arrest; and self-rated psychological and physical health, and quality of life on 0 (poor)-10 (good) scales. All items refer to the previous 4 weeks.

Data management and analysis:

ATOP and demographic data were collected and managed using Research Electronic Data Capture (REDCap) online electronic data capture software [23, 24] or on paper forms entered into the database post-interview. Data were exported from REDCap to SPSS statistics package version 25.0 [25] for analysis.

Domain agreement was assessed using Pearson's correlation coefficient for continuous and ordinal variables, and Cohen's κ for nominal variables. Cutoffs used were inadequate (<0.5), moderate (0.5-0.7), good (0.71-0.9) and excellent agreement (>0.9) [26]. Bootstrap procedures using 10 000 samples were used to estimate 95% confidence intervals (CIs). Order of administration effects were tested by creating method difference scores for key variables and testing for order differences using one-way analysis of variance (ANOVA).

Results

The larger COQI study had 278 participants, of whom 236 completed follow up 1 and 165 completed the telephone ATOP. 107 telephone ATOPs were completed within 1-5 days of the face to face ATOP and were used in the current analysis (Figure 1). Participants did not have a completed telephone ATOP because they either did not have a phone, or were not contactable after several contact attempts.

Table 1 describes demographics, interview order, and the mean time difference between completion of each ATOP for the 107 participants. Participants had a mean age of 47 years (SD 10 years) and just over half (54%) were male. Our sample was somewhat older and had more women than clients entering public AoD services in NSW in general [27], where 54% were aged 20-39, and 66% were male. Eighty-seven (81%) of participants completed the face to face ATOP first, and the mean time difference between completion was 2 days.

<insert Figure 1 here>
<insert Table 1 here>

Table 2 presents concurrent validity of the ATOP items for telephone administration compared to face to face use. Nearly all items reached good (0.71-0.9 on the Pearson's correlation coefficient or Cohen's κ) or excellent (>0.9) agreement. Exceptions were use of the other opioids (any use and days used) where only moderate agreement was reached, and injection with use of other equipment was unable to be accurately assessed due to very low use in this sample. Cocaine use, risk of eviction and arrest were also low in this sample, and could not be properly assessed for agreement.

Assessment for differences between administration order is presented in Supplementary table 1 for key continuous variables – days of substance use for the most frequently endorsed categories (alcohol, cannabis, other opioids and benzodiazepines), and psychological health, physical health and quality of life. No bias was exhibited towards the telephone or face to face ATOP being performed first.

<insert Table 2 here>

Discussion

When administered via the telephone, nearly all ATOP measures reached moderate or excellent validity compared to face to face administration. This shows the ATOP can be confidently used in situations where face to face contact is impractical.

Low rates of reporting cocaine use, sharing of injection equipment, risk of eviction and arrest in this sample meant these items could not be validated. There was lower agreement on use of 'other opioids' than for other substances. We have previously noted lower reliability of the other opioids item [3], which may relate to the study being conducted whilst over-the-counter codeine products were available in Australia [28], and some participants may not have considered these relevant. This confusion is now less likely as all codeine-containing products became prescription-only in Australia in early 2018. Clients may also have not reported other prescribed opioid use (excluding opioid

treatment medications) despite researchers being trained to query for such use. The operational definitions for some items (including opioid use), an ATOP Manual, and training of interviewers are being reviewed to optimise use of the ATOP.

Differences in demographics between our sample and that of clients entering treatment in NSW generally may reflect our sample consisting of continuing AoD clients as well as newly-entered clients.

Limitations: The order of ATOP completion (telephone/face to face or face to face/telephone) was not randomised for reasons discussed in the methods section. However, we were able to achieve completion of the second ATOPs within a short timeframe, and each ATOP pair was completed by the same researcher. Operationally, we were unable to utilise clinicians to complete ATOPs which would have been closer to a real-world clinical use; we have used trained researchers as proxies for clinicians. We have previously found there to be good to excellent agreement between clinician and researcher-completed ATOPs [3]. Finally, telephone validity was only assessed among people attending for alcohol or opioid treatment and for those with reliable phone contact; validity among those attending for other substances and with unreliable communications has not been determined.

Conclusion

The ATOP can be reliably used over the telephone in clinical situations where face to face contact is not possible. We note that some items with low incidence could not be properly assessed in this study; however their clinical utility may still be valid. Our findings suggest that the ATOP is a valid tool for remote use over the telephone in AoD settings. This finding means it can be used to support staff to monitor clients' outcomes and risks, a finding of particular relevance in response to the COVID-19 pandemic in which services are increasingly relying on telehealth approaches to patient monitoring.

Acknowledgements

We thank clients and study participants at the recruitment sites: AoD services in the Hunter New England, Western Sydney and South Eastern Sydney Local Health Districts. We also thank other research staff: Elaine Murray, Susan Hazelwood, Janet Dunbabin, Woroud Alzaher, Consuelo Rivas, Dr Nghi Phung, Jennifer (Mary) Luksza and Dr Thao Lam.

This work was funded as part of the Clinical Outcomes and Quality Improvement Framework study funded by the Centre for Population Health at the NSW Ministry of Health.

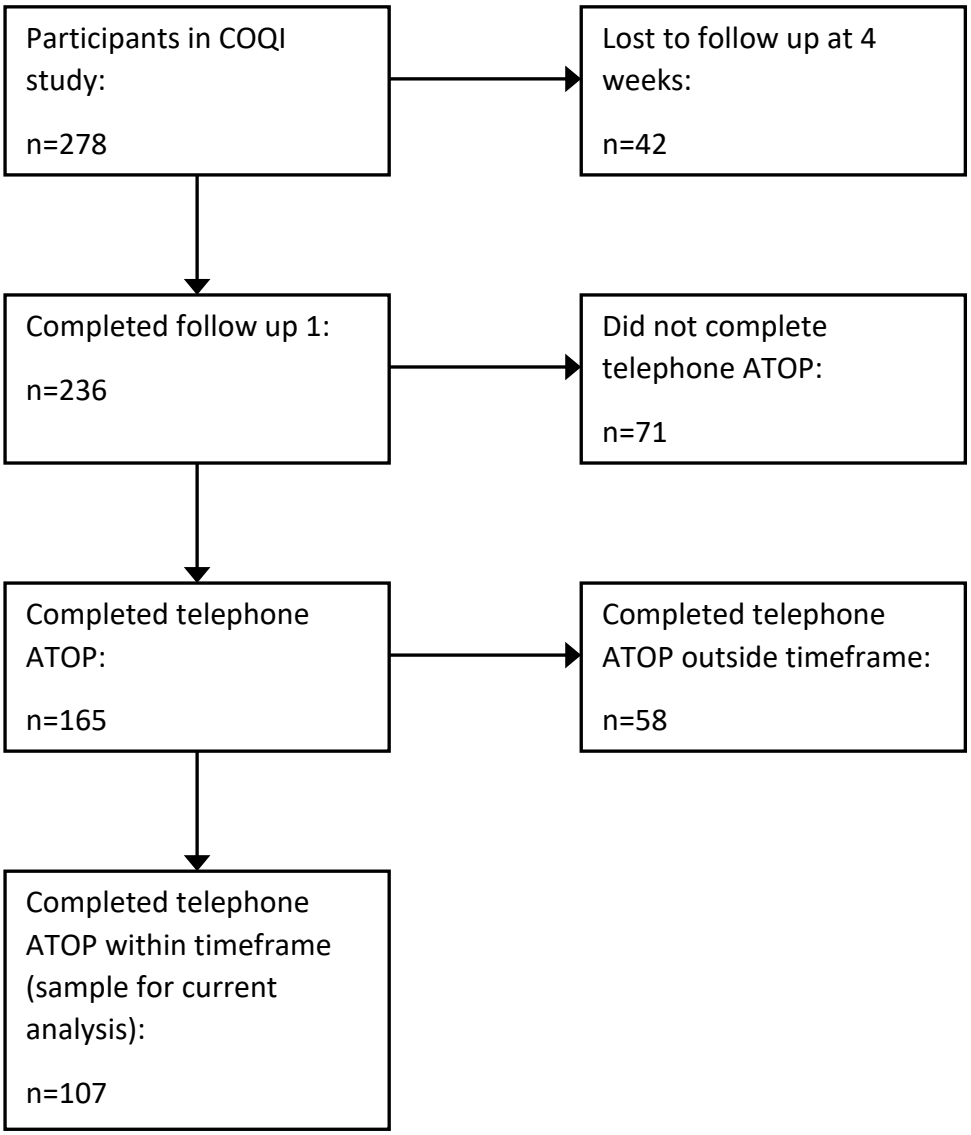
References:

1. Ryan A, Holmes J, Hunt V, Dunlop A, Mammen K, Holland R, et al. Validation and implementation of the Australian Treatment Outcomes Profile in specialist drug and alcohol settings. Drug Alcohol Rev. 2014;33(1):33-42. doi.org/10.1111/dar.12083.
2. Marsden J, Farrell M, Bradbury C, Dale-Perera A, Eastwood B, Roxburgh M, et al. Development of the treatment outcomes profile. Addiction. 2008;103(9):1450-60. doi.org/10.1111/j.1360-0443.2008.02284.x.
3. Deacon RM, Mammen K, Bruno R, Mills L, Dunlop A, Holmes J, et al. Assessing the concurrent validity, inter-rater reliability and test-retest reliability of the Australian Treatment Outcomes Profile (ATOP) in alcohol and opioid treatment populations. Under review.
4. Mills L, Lintzeris N, Bruno R, Montebello M, Dunlop A, Deacon RM, et al. Validation of the Australian Treatment Outcomes Profile for use in clients with cannabis dependence. Drug Alcohol Rev. 2020; early view. doi.org/10.1111/dar.13050.
5. Lintzeris N, Monds LA, Rivas G, Leung S, Withall A, Draper B. The Australian Treatment Outcomes Profile instrument as a clinical tool for older alcohol and other drug clients: A validation study. Drug Alcohol Rev. 2016;35(6):673-7. doi.org/10.1111/dar.12393.
6. NSW Health. NSW Health Telehealth framework and implementation strategy. Sydney: NSW Ministry of Health; 2016.
7. NSW Health. eHealth Strategy for NSW Health 2016-2026. Sydney: NSW Ministry of Health; 2016.
8. Wang M, Chen RC, Usinger DS, Reeve BB. Evaluating measurement invariance across assessment modes of phone interview and computer self-administered survey for the PROMIS measures in a population-based cohort of localized prostate cancer survivors. Qual Life Res. 2017;26(11):2973-85. doi.org/10.1007/s11136-017-1640-3.
9. Sobell LC, Brown J, Leo GI, Sobell MB. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. Drug Alcohol Depend. 1996;42(1):49-54. doi.org/10.1016/0376-8716(96)01263-X.
10. Maisto SA, Conigliaro JC, Gordon AJ, McGinnis KA, Justice AC. An experimental study of the agreement of self-administration and telephone administration of the timeline followback interview. J Stud Alcohol Drugs. 2008;69(3). doi.org/10.15288/jsad.2008.69.468.
11. Perkins JJ, Sanson-Fisher RW. An examination of self- and telephone-administered modes of administration for the Australian SF-36. J Clin Epidemiol. 1998;51(11):969-73. doi.org/10.1016/S0895-4356(98)00088-2.
12. McHorney CA, Kosinski M, Ware JE. Comparisons of the costs and quality of norms for the SF-36 Health Survey collected by mail versus telephone interview: Results from a national survey. Med Care. 1994;32(6):551-67. doi.org/10.1097/00005650-199406000-00002.
13. Watson E, Firman D, Baade P, Ring I, editors. Telephone administration of the SF-36 in the Queensland general population, and a comparison of telephone versus mail administration. Proceedings of the Health Outcomes and Quality of Life Measurement Conference; 1995; Canberra ACT, Australia: Australian Health Outcomes Clearing House.
14. Watson EK, Firman DW, Baade PD, Ring I. Telephone administration of the SF-36 health survey: validation studies and population norms for adults in Queensland. Aust N Z J Public Health. 1996;20(1326-0200 (Print)):359-63. doi.org/10.1111/j.1467-842x.1996.tb01046.x.
15. García M, Rohlfs I, Vila J, Sala J, Pena A, Masiá R, et al. Comparison between telephone and self-administration of Short Form Health Survey Questionnaire (SF-36). Gac Sanit. 2005;19(6):433-9. doi.org/10.1016/S0213-9111(05)71393-5.
16. Conwell Y, Simning A, Driffill N, Xia Y, Tu X, Messing SP, et al. Validation of telephone-based behavioral assessments in aging services clients. Int Psychogeriatr. 2018;30(1):95-102. doi.org/10.1017/S1041610217001752.
17. Ware J. SF-36 physical and mental health summary scales: A user`s manual. Boston: The Health Institute, New England Medical Center; 1994.

18. Lin S-Y, Kerse N, McLean C, Moyes S. Validation of quality of life and functional measures for older people for telephone administration. J Prim Health Care. 2010;2(1):35-42. doi.org/10.1071/HC10035.
19. Evans M, Kessler D, Lewis G, Peters TJ, Sharp D. Assessing mental health in primary care research using standardized scales: can it be carried out over the telephone? Psychol Med. 2004;34(1):157-62. doi.org/10.1017/s0033291703008055.
20. Sobin MC, Weissman BM, Goldstein DR, Adams DP, Wickramaratne DP, Warner DV, et al. Diagnostic interviewing for family studies: comparing telephone and face-to-face methods for the diagnosis of lifetime psychiatric disorders. Psychiatr Genet. 1993;3(4):227-34. doi.org/10.1097/00041444-199324000-00005.
21. Wells KB, Burnam MA, Leake B, Robins LN. Agreement between face-to-face and telephone-administered versions of the depression section of the NIMH Diagnostic Interview Schedule. J Psychiatr Res. 1988;22(3):207-20. doi.org/10.1016/0022-3956(88)90006-4.
22. Robins LN. Epidemiology: Reflections on testing the validity of psychiatric interviews. Arch Gen Psychiatry. 1985;42(9):918-24. doi.org/10.1001/archpsyc.1985.01790320090013.
23. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform. 2009;42(2):377-81. doi.org/10.1016/j.jbi.2008.08.010.
24. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform. 2019;95:103208. doi.org/10.1016/j.jbi.2019.103208.
25. IBM Corp. IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY: IBM Corp; 2016.
26. Nunnally JC. Psychometric theory. 2nd ed. New York: McGraw-Hill; 1978.
27. AIHW. Alcohol and other drug treatment services in Australia 2017–18: key findings Canberra: Australian Institute of Health and Welfare; 2019 [Available from: https://www.aihw.gov.au/reports/alcohol-other-drug-treatment-services/aodts-2017-18-key-findings/contents/summary].
28. Roberts DM, Nielsen S. Changes for codeine. Aust Prescr. 2018;41:2-3. doi.org/10.18773/austprescr.2018.006.

This is the accepted version of the following article: Deacon, R.M., Mammen, K., Holmes, J., Dunlop, A., Bruno, R., Mills, L., Graham, R. and Lintzeris, N. (2020), Assessing the validity of the Australian Treatment Outcomes Profile for telephone administration in drug health treatment populations. Drug Alcohol Rev. doi:10.1111/dar.13088 which has been published in final form at www.onlinelibrary.wiley.com

Figure 1: Flowchart showing number of participants recruited to the main Clinical Outcomes and Quality Indicators (COQI) study, numbers lost to follow-up and the number of participants included in the analysis of concurrent validity of the ATOP for telephone administration against face-to-face administration



This is the accepted version of the following article: Deacon, R.M., Mammen, K., Holmes, J., Dunlop, A., Bruno, R., Mills, L., Graham, R. and Lintzeris, N. (2020), Assessing the validity of the Australian Treatment Outcomes Profile for telephone administration in drug health treatment populations. Drug Alcohol Rev. doi:10.1111/dar.13088 which has been published in final form at www.onlinelibrary.wiley.com

Table 1: Demographics, principal drug of concern and order of ATOP completion (face to face or telephone first) for the sample

N	**107**
Age, mean (SD)	47 (10)
Gender, n (%)	
Female	49 (46)
Male	58 (54)
Principal drug of concern (%)	
Alcohol	63 (59)
Opioids including heroin	44 (41)
Order of ATOP completion (%)	
Face to face first	87 (81)
Telephone first	20 (19)
Mean (SD) absolute time difference between completion, days	2.1 (1.2)

This is the accepted version of the following article: Deacon, R.M., Mammen, K., Holmes, J., Dunlop, A., Bruno, R., Mills, L., Graham, R. and Lintzeris, N. (2020), Assessing the validity of the Australian Treatment Outcomes Profile for telephone administration in drug health treatment populations. Drug Alcohol Rev. doi:10.1111/dar.13088 which has been published in final form at www.onlinelibrary.wiley.com

Table 2: Procedural validity of the ATOP items for telephone administration compared to face to face administration by the same researcher. All items refer to behaviour in the previous 4 weeks.

ATOP item	Face to face ATOP n=107	Telephone ATOP n=107	Test value[a] (95% CI; p)
Alcohol			
Used, n (%)	42 (39)	38 (35)	0.88 (0.78, 0.96; <0.01)
Days used, m (SD)	13 (10)	12 (10)	0.97 (0.93, 0.99; <0.01)
Units/day, m (SD)	11 (10)	11 (9)	0.86 (0.77, 0.96, <0.01)
Cannabis			
Used, n (%)	24 (22)	22 (21)	0.94 (0.85, 1.00; <0.01)
Days used, m (SD)	13 (10)	13 (11)	0.94 (0.86, 0.99; <0.01)
Amphetamines			
Used, n (%)	6 (6)	6 (6)	1.00 (1.00, 1.00; <0.01)
Days used, m (SD)	5 (4)	4 (3)	0.87 (0.02)[b]
Benzodiazepines			
Used, n (%)	24 (22)	21 (20)	0.92 (0.80, 1.00; <0.01)
Days used, m (SD)	12 (11)	11 (11)	0.96 (0.88, 1.00; <0.01)
Heroin			
Used, n (%)	7 (6)	5 (5)	0.82 (0.48, 1.00; <0.01)
Days used, m (SD)	7 (5)	6 (6)	0.94 (<0.01)[b]
Other opioids			
Used, n (%)	17 (16)	14 (13)	0.66 (0.42, 0.84; <0.01)
Days used, m (SD)	10 (12)	9 (12)	0.55 (-0.01, 0.96; <0.01)
Cocaine			
Used, n (%)	1 (1)	1 (1)	1.00 (1.00, 1.00; <0.01)
Days used, m (SD)	4 (-)	3 (-)	-
Tobacco			
Daily use, n (%)	75 (70)	75 (70)	0.82 (0.69, 0.93; <0.01)
Injected drugs			
Injected, n (%)	11 (10)	10 (19)	0.95 (0.81, 1.00; <0.01)
Days inject, m (SD)	7 (6)	6 (5)	0.94 (0.69, 0.99; <0.01)
Injected with used equipment			
n (%)	0 (0)	1 (1)	0.00 (0,0)[b]
Employment			
Any, n (%)	31 (29)	30 (28)	0.93 (0.84, 1.00; <0.01)
Days, m (SD)	15 (7)	14 (7)	0.89 (0.79, 0.89; <0.01)
Training			
Any, n (%)	10 (9)	8 (7)	0.88 (0.65, 1.00; <0.01)
Days, m (SD)	7 (6)	6 (7)	0.92 (0.60, 1.00; <0.01)
Homelessness			
n (%)	6 (6)	5 (5)	0.90 (0.65, 1.00; <0.01)
At risk of eviction			
n (%)	1 (1)	1 (1)	1.00 (1.00, 1.00; <0.01)
Caring for children			
<5 years old, n (%)	14 (13)	13 (12)	0.96 (0.85, 1.00; <0.01)
5-15 years old, n (%)	20 (19)	20 (19)	1.00 (1.00, 1.00; <0.01)
Arrested			
n (%)	1 (1)	1 (1)	1.00 (1.00, 1.00; <0.01)
Violence to you			
n (%)	6 (6)	7 (6)	0.75 (0.38, 1.00; <0.01)
Violence to others			
n (%)	5 (5)	5 (5)	1.00 (1.00, 1.00; <0.01)
Psychological health			
m (SD)	6.0 (2.0)	5.9 (2.0)	0.83 (0.73, 0.91; <0.01)
Physical health			
m (SD)	5.9 (2.0)	6.0 (1.9)	0.88 (0.81, 0.93; <0.01)
Quality of life			
m (SD)	6.4 (2.2)	6.4 (2.1)	0.87 (0.80, 0.92; <0.01)

[a] Pearson's correlation coefficient for continuous and ordinal items; Cohen's κ for nominal items

[b] no bootstrapping performed on these items as cell sizes were too small

n: number; m: mean; SD: standard deviation; CI: confidence interval

This is the accepted version of the following article: Deacon, R.M., Mammen, K., Holmes, J., Dunlop, A., Bruno, R., Mills, L., Graham, R. and Lintzeris, N. (2020), Assessing the validity of the Australian Treatment Outcomes Profile for telephone administration in drug health treatment populations. Drug Alcohol Rev. doi:10.1111/dar.13088 which has been published in final form at www.onlinelibrary.wiley.com

Table S1: Test results for differences between telephone and face to face administration for key continuous variables (substance use categories with the highest use) and the self-rated scores. Mean differences between the modes of administration for the two groups (telephone ATOP first and face to face ATOP first) were calculated and ANOVA test applied to assess any overall difference.

	Mean differences between modes of administration (95% CI)		
	Telephone first (n=20)	**Face to face first (n=87)**	**One-way ANOVA F(df), p**
Alcohol use days	-0.25 (-0.52, 0.05)	-0.28 (-0.70, 0.07)	0.004 (1,105), 0.951
Cannabis use days	-0.40 (-1.58, 0.29)	-0.05 (-0.37, 0.29)	0.689 (1,105), 0.405
BZD use days	-0.20 (-0.67, 0.00)	-0.25, (-0.67, 0.01)	0.018 (1,105), 0.894
Other opioid use days	-0.15 (-0.50, 0.00)	-0.29 (-0.67, 0.01)	0.013 (1,105), 0.909
Psychological health	-0.35 (-1.00, 0.27)	-0.10 (-0.33, 0.12)	0.738 (1,105), 0.392
Physical health	0.20 (-0.12, 0.61)	0.01 (-0.19, 0.22)	0.631 (1,105), 0.429
Quality of life	0.05 (-0.42, 0.53)	0.02 (-0.20, 0.25)	0.010 (1,105), 0.921