

## **A Measurement Invariance Analysis of Selected Opioid Overdose Knowledge Scale (OOKS)**

### **Items among Bystanders and First Responders**

James A. Swartz<sup>1\*</sup>¶, Qiao Lin<sup>2¶</sup>, and Yerim Kim<sup>1¶</sup>

<sup>1</sup>Jane Addams College of Social Work, University of Illinois Chicago,  
Chicago, Illinois, United States of America

<sup>2</sup>College of Education, University of Illinois Chicago,  
Chicago, Illinois, United States of America

\* Corresponding author  
E-mail: [jaswartz@uic.edu](mailto:jaswartz@uic.edu)

¶ These authors contributed equally to this work.

## Abstract

The Opioid Overdose Knowledge Scale (OOKS) is widely used as an adjunct to opioid education and naloxone distribution (OEND) for assessing pre- and post-training knowledge. However, the extent to which the OOKS performs comparably for bystander and first responder groups has not been well determined. We used exploratory structural equation modeling (ESEM) to assess the measurement invariance of an OOKS item subset when used as an OEND training pre-test. We used secondary analysis of pre-test data collected from 446 first responders and 1,349 bystanders (N=1,795) attending OEND trainings conducted by two county public health departments. Twenty-four items were selected by practitioner/trainer consensus from the original 45-item OOKS instrument with an additional 2 removed owing to low response variation. We used exploratory factor analysis (EFA) followed by ESEM to identify a factor structure, which we assessed for configural, metric, and scalar measurement invariance by participant group using the 22 dichotomous items (correct/incorrect) as factor indicators. EFA identified a 3-factor model consisting of items dealing with: basic overdose risk information, signs of an overdose, and rescue procedures/advanced overdose risk information. Model fit by ESEM estimation versus confirmatory factor analysis showed the ESEM model afforded a better fit. Measurement invariance analyses indicated the 3-factor model fit the data across all levels of invariance per standard fit statistic metrics. The reduced set of 22 OOKS items appears to offer comparable measurement of pre-training knowledge on opioid overdose risks, signs of an overdose, and rescue procedures for both bystanders and first responders.

**Keywords:** exploratory structural equation modeling, ESEM, opioid education and naloxone distribution, OEND, Opioid Overdose Knowledge Scale, OOKS, measurement invariance

## Introduction

Although overshadowed the past two years by the rapid, global emergence of COVID-19 as a major public health threat, the decades-long opioid epidemic continues unabated, worsening in 2020. For the 12-month period ending in May 2020, 81,243 OORF were reported to the CDC, representing the largest 12-month increase in fatalities since 2015 (1). At least part of the recent trend reversal in opioid overdose-related fatalities (OORF) is directly related to COVID-19 (2). The pandemic increased stress, isolation and loneliness, as well as housing instability, and made accessing drug treatment more difficult (3). Given the enduring nature of the opioid epidemic, however, it is likely that even after the COVID-19 pandemic has run its course, opioid misuse and OORF will remain substantial public health concerns.

Prominent among CDC recommendations for reducing OORF is to “expand the provision and use of naloxone and overdose prevention education” with an emphasis on raising awareness about the “critical need for bystanders to have naloxone on hand and use it during an overdose” (1). Expanding community naloxone access has been one of three US Department of Health and Human Services priority areas for addressing the opioid crisis (4).

Providing take-home naloxone (THN), a competitive opioid antagonist that can rapidly reverse the life-threatening effects of an opioid-related overdose, to non-medically trained “bystanders” (e.g., a heterogeneous group that can include opioid users, their social networks and family members, and social service staff) has become an important component of expanding naloxone access (5-7). THN has become more widespread over the past 5 years, abetted by the passage of “Good Samaritan” laws; increased availability through pharmacies, hospitals, and emergency departments; Medicaid expansion under the Affordable Care Act; and

the availability of a nasal spray, obviating the need for administration by injection (8-10).

Despite increased availability, there continues to be a need for further expansion of community-based overdose education and naloxone distribution (OEND) for bystanders as well as for non-medical first responders such as the police.

Though studies of whether THN is a moral hazard that increases risky opioid use (i.e., risk compensation) have yielded mixed results, the preponderance of the evidence suggests that, on balance, THN reduces OORF (4, 11, 12). Moreover, while it is possible the effectiveness of providing THN is enhanced when combined with education and training on recognizing the signs of an overdose and how to correctly administer naloxone, current evidence that providing education and training provides additional benefit has not been well established (4, 12-17).

As the continuation of THN programs that incorporate education and training is likely and because there is no single standardized set of instructional materials or manualized training protocol, it is important to have validated instruments for assessing whether any specific training effectively increases basic knowledge on recognizing the signs of an opioid overdose and how to correctly administer naloxone for the intended audience, whether bystanders or first responders. Having valid information on whether education and training effectively increase knowledge and skills are necessary pre-requisites to determining if OEND further contributes to reductions in OORF beyond naloxone distribution alone and for whom and for determining which training content and materials are most effective.

The Opioid Overdose Knowledge Scale (OOKS) was the first validated instrument of which we are aware to be used in conjunction with OEND (18) although other instruments have

become available more recently (19). Consisting of 45 items in its original version, the OOKS continues to be frequently cited in studies that evaluate THN/OEND trainings (20-22). The OOKS has been translated into several languages for use in European countries and a short, 10-item form (i.e., the Brief Opioid Overdose Knowledge Scale [BOOK]) has been developed as well as a version adapted for use to assess knowledge specific to prescription opioids (23-25).

Despite its popularity, there remain questions about whether the OOKS is optimally constructed for use with bystanders as well as first responders. In the original study, the OOKS was administered to both bystanders (i.e., small samples of friends and family members of heroin users; N=42) as well as healthcare professionals (N=56). The BOOK was developed based on larger samples of illicit opioid users and patients prescribed an opioid for chronic pain treatment [N=848; Dunn, Barrett (23)]. In both studies, most participants would be considered bystanders with the status of the trained healthcare professionals unclear. Police, who are very often the first responders when overdoses occur, were not involved in the development of the OOKS or the BOOK. Considering the continuing need to assess the effectiveness of different OEND training strategies for groups of participants with varying knowledge and professional training, this study sought to evaluate the performance of OOKS items when used as a training pre-test for first responders as well as bystanders. Specifically, we conducted a multigroup analysis of OOKS item equivalence by assessing levels of measurement invariance (i.e., configural, metric, and scalar). We also wanted to re-examine the OOKS item factor structure using the relatively new analytic tool, exploratory structural equation modeling (ESEM), which combines elements of exploratory factor analysis and confirmatory factor analysis (26, 27). We describe the potential advantages of ESEM in the analyses section below.

## Material and Methods

This study was a secondary analysis of de-identified pre-test data collected as part of an evaluation of a project to reduce OORF through increasing THN availability and conducting OEND trainings in Illinois communities with high OORF rates. The University of Illinois Chicago IRB determined the study did not constitute human subjects research and granted an exemption.

### Study setting and sample

The Substance Use and Mental Health Services Administration (SAMSHA) provided funding for a five-year project (2016–2021) to reduce OORF among users of any type of opioid including street drugs such as heroin or fentanyl. Group trainings on naloxone administration and recognizing an overdose were scheduled as needed at each of the six participating sites based on outreach efforts in participating communities. As part of the project evaluation, two sites agreed to administer pre- and post-test quizzes to assess knowledge gain and assess training effectiveness. A 24-item version of the OOKS was used and required about 10 minutes to administer. One site trained both first responders (N=498) and bystanders (N=506) whereas the second site exclusively trained bystanders (N=1,137). We eliminated data from participants with more than 3 missing responses or with a response pattern indicating rote responding (N=346, 16.2%). This yielded a final analytic sample of 1,795 participants composed of 446 (24.8%) first responders and 1,349 (75.2%) bystanders.

First responders were composed predominantly of police whereas composition of the bystander groups varied considerably. Aggregate data provided by the site that conducted bystander trainings indicate the following types of bystanders participated: active drug

users/syringe exchange program clients; family members and friends of drug users; staff working in substance use treatment programs; staff in human services agencies such as shelters and transitional living programs; and staff working in public institutions such as schools, restaurants, and libraries. In total, the two participating sites conducted 109 trainings between March 10, 2018 and January 20, 2020. The mean number of participants per training was 36.1 (sd=25.3), with the size of bystander trainings larger and more variable (mean=38.7, sd=27.1) compared with first responder trainings (mean=24.9, sd=9.8;  $t_{(df=2,029)}=10.0$ ,  $p < .001$ ).

We did not collect individual demographic information to preserve training participant anonymity. However, aggregated data indicate the following demographic composition of the study sample: most participants were male (80.1%); white (82.1%); and between the ages of 25 to 44 (57.2%) with smaller proportions reporting they were 18 to 24 years old (19.4%) or 45 to 64 years old (10.6%). About 8.3% were African American/Black with 12.7% indicating Latinx ethnicity.

## Measures

Because training content and delivery varied by trainer, setting, and participant composition, we restricted our analyses to the pre-tests to eliminate variance owing to training-related factors that could affect post-test responses. At project outset, a panel composed of site directors and training practitioners reviewed the original OOKS for use in the evaluation of trainings. Because of the limited time for trainings, 50 to 60 minutes at most, the review panel shortened the 45-item OOKS by removing 21 items they believed were redundant or less informative to the project training goals. After preliminary analyses of the collected training data, we removed two additional items – one should call an ambulance when managing an

overdose as well as stay with the person until an ambulance arrives - that were answered correctly by more than 98% of all participants and consequently lacked variability, causing convergence problems during the measurement invariance analyses. The resulting 22-items for the modified version of the OOKS are shown in Table 1 with the original OOKS section and item number provided parenthetically.

The pre-test was administered as a self-report paper questionnaire. For each item, participants were asked to indicate true, false, or unsure/unknown. The self-reported data were then entered into REDCap (28), downloaded and scored as correct or incorrect according to the answer key provided with the original OOKS (18). All unsure/don't know responses were coded as incorrect, effectively converting the items from trichotomous to dichotomous.

### **Analyses**

We used Stata version 17.0 for data screening and generating bivariate statistics (29). We first ran descriptive analyses to assess percentage correct by item as well as number of total correct responses, disaggregated into three groups: bystanders trained at site A; first responders trained at site A; and bystanders trained at site B. For each item, a bivariate logistic regression was used to compare the odds of a correct response by group. We also calculated a mean score for each group based on the total correct and compared these using a one-way ANOVA.

To assess the 22 OOKS items for measurement invariance, we first ran an exploratory factor analysis (EFA) to determine the number of factors for the pre-test instrument. Although a factor structure has been identified for the original 45-item OOKS, we wanted to re-assess this given we were using an item subset. We examined factor structure in several ways: using principal factor analysis with goemin rotation and robust weighted least squares in Mplus version 8.7

(30) to generate models with 2 to 6 factors and comparing the CFI, TLI, RMSEA, and SRMR fit statistics for each model. We also conducted parallel analysis in the R software program version 4.1.3 (31) using the `fa.parallel` function in the `psych` package for R (32) assessing polychoric correlations, given the factor indicators were binary (correct/incorrect). Based on analyses of the relative accuracy of different metrics for most accurately determining the number of factors in an EFA under varying circumstances (e.g., sample size, underlying number of factors, factor correlations and loading), we gave greater weight to the RMSEA statistic (33) in selecting a final model. Our goal was to find the model with the minimum number of factors that still provided adequate fit to the data.

We then used exploratory structural equation modeling (ESEM) to test levels of measurement invariance for the selected factor model. ESEM combines EFA, and CFA/SEM and is less restrictive than CFA because factor cross-loadings are not constrained to zero. This potentially permits estimation of better-fitting and more realistic models whereby items can be associated with multiple factors (26, 34). We directly assessed this by estimating two models based on CFA – one which did and did not allow for selected residual covariances – with two corresponding models based on ESEM. We used conventional fit statistics and thresholds to evaluate model fit (35): root mean square error of approximation (RMSEA)  $\leq 0.5$ ; comparative fit (CFI) and Tucker–Lewis indices (TLI)  $\geq 0.90$ ; and standardized root mean square residual (SRMR)  $\leq 0.8$ .

We assessed the OOKS pre-test for measurement invariance for first responders and bystanders across three increasingly restrictive measurement levels using the best fitting model derived from the CFA-ESEM comparative analyses: configural, metric or “weak invariance”, and

scalar (36). Briefly, configural invariance is indicated when factors are composed of the same items across groups, but the item factor loadings and intercepts/thresholds are allowed to vary across groups; with metric invariance, an equality constraint is applied to the factor loadings of each item; and finally, with scalar invariance, an additional equality constraint is applied to item intercepts/thresholds. For each successive model, we again used a conventional set of fit statistics and thresholds to evaluate model fit.

## Results

### Descriptive statistics

Table 1 shows the percentage of total correct responses by item as well as by participant group. In general, the percentages of correct answers were similar among the two bystander groups with only two of the more difficult items – whether naloxone can be used to reverse an overdose for *any* drug and how long the effects of naloxone last – showing significantly different proportions of correct responses. There were many more significant differences between the combined bystander groups and the first responders with first responders having a higher proportion of correct responses to 18 of the 22 OOKS items. Across all participants, the three items that were answered correctly least often were the length of time naloxone lasts (21.4% correct) and having blood shot eyes (32.1% correct) or displaying agitated behavior (38.0% correct) as indicators of an opioid overdose.

**Table 1. Percentage Correct Responses to Selected OOKS Items by OEND Training Group**

Selected OOKS Pre-test Items	Site A - First Responders	Site A - Bystanders	Site B - Bystanders	Total	Sig	Sig
	(N = 446)	(N = 412)	(N = 937)	(N = 1,795)	(FR - BYS)	(BYS_A - BYS_B)
<b>Which of the following factors increase the risk of a heroin (opioid) overdose?</b>	Correct %	Correct %	Correct %	Correct %		
Switching from smoking to injecting heroin (A2)	82.7	77.5	77.9	79.0 %	NS	NS
Using heroin with other substances, such as alcohol or sleeping pills (A3)	96.6	95.6	97.6	96.9	NS	NS
Increase in heroin purity (A4)	95.1	94.4	95.1	94.9	NS	NS
Using heroin again soon after release from prison (A8)	85.0	77.0	75.0	78.0	***	NS
Using heroin again after a detoxification treatment (A9)	90.4	85.0	84.0	85.8	**	NS
<b>Which of the following are indicators of an opioid overdose?</b>						
Having blood-shot eyes (B1)	39.9	27.4	30.4	32.1	***	NS
Slow or shallow breathing (B2)	95.1	88.1	84.2	87.8	***	NS
Lips, hands or feet turning blue (B3)	89.0	78.9	79.8	81.9	***	NS
Loss of consciousness (B4)	97.3	94.0	91.1	93.3	***	NS
Deep snoring (B7)	66.8	55.2	49.6	55.2	***	NS
Very small pupils (B8)	82.3	73.6	69.9	73.8	***	NS
Agitated behaviour (B9)	47.3	38.7	33.2	38.0	***	NS
Rapid heartbeat (B10)	46.0	33.4	29.6	34.5	***	NS
<b>Which of the following should be done when managing a heroin (opioid) overdose?</b>						
Give stimulants (e.g. cocaine or black coffee) (C5)	56.1	50.4	52.6	53.0	NS	NS
Place the person in the recovery position (on their side with mouth clear) (C6)	92.8	80.2	79.6	83.0	***	NS
Put the person in bed to sleep it off (C11)	69.5	54.2	55.6	58.7	***	NS
<b>What is naloxone used for?</b>						
To reverse the effects of any overdose (D4)	63.9	46.5	38.2	46.5	***	**
<b>How can naloxone be administered?</b>						
Into mouth or swallowed orally (E4)	79.4	63.0	45.8	58.1	***	***
<b>How long do the effects of naloxone last?</b>						
2 to 6 hours (H3)	27.3	15.7	21.1	21.4	***	NS
<b>Please indicate which of the following statements are correct</b>						
If the first dose of naloxone has no effect a second dose can be given (I1)	72.2	52.3	49.8	56.0	***	NS
Someone can overdose again even after having received naloxone (I3)	79.6	72.6	70.2	73.1	***	NS
Naloxone can provoke withdrawal symptoms (I6)	50.2	43.8	40.6	43.7	**	NS

**Note.** All figures shown are percentages of correct responses for each item by participant group and site. The labels in parentheses by each item reference the original OOKS instrument item section and number. All significance tests are based on Pearson chi-square tests with 2 degrees of freedom. Only results significant at  $p < .01$  are reported. The first set of significance tests compare the combined set of bystanders with first responders. The second set compares the percent of correct responses for bystanders at site A with bystanders at site B.

\*\* =  $p < .01$ , \*\*\* =  $p < .001$ , NS = Non-significant

On average, first responders answered 16 of the 22 items correct (mean=16.1, SD=3.5) whereas bystanders at site A averaged just under 14 correct responses (mean=13.9, SD=3.9) as did bystanders at site B (mean=13.5, SD=4.2). These differences were statistically significant ( $F_{(2,1792)}=63.82, p < .001$ ) with post-hoc analyses indicating that the first responders scored significantly higher than either bystander group but that the small difference among the bystander groups was non-significant.

### **Exploratory factor analyses**

Results for the exploratory analyses for the 2 to 6 factor models ruled out only the 2-factor model as not providing adequate fit (RMSEA=.053 [95% CI=.050 - .056]; CFI=.905; TLI=.883; SRMR=.077). Models with 3 to 6 factors fit the data well per these same statistics with each showing improvement in model fit over the preceding model as determined by Chi-square tests comparing each model with the model having k-1 factors. These comparisons supported the 6-factor model as providing the best fit as did inspection of the scree plot obtained following parallel analysis.

Despite these results, and for several reasons, we selected the 3- (RMSEA=.044 [95% CI=.041 - .048]; CFI=.941; TLI=.919; SRMR=.067) and 4-factor models (RMSEA=.044 [95% CI=.034 - .041]; CFI=.962; TLI=.941; SRMR=.057) for further analysis. First, as noted by Finch (2020) in his simulation study, when the fit statistics were inconsistent with the underlying factor structure of the simulated models, the statistics tended to favor over-factored results. Parsimony was also an important consideration. One of the factors in the 4-factor model had only two items with loadings greater than .50 with the 5 and 6-factor models having a similar factor structure. In subsequent preliminary measurement invariance analyses, the 4-factor model failed to

converge at the metric invariance step suggesting it could be overly complex. Given that both the 3- and 4-factor model fit statistics indicated both fit the data well and the item to factor structure appeared to be more robust in the 3-factor model, which had more than 2 items with loadings  $> .50$  on every factor, and the pattern of factor loadings for the items made sense substantively, we selected the 3-factor model for further measurement invariance testing.

### **ESEM model**

The final ESEM factor structure for the 3-factor model identified in the EFA step is shown in Table 2, which displays factor loadings and significance levels of the selected OOKS items. Based on the significant factor loadings, we assessed the factor 1 (**overdose risks**) as determined by correct identification of opioid overdose risk items drawn from section A of the OOKS. Items corresponding to recognition of overdose signs and drawn from section B of the OOKS were most strongly associated with factor 2 (**overdose signs**). Factor 3 (**rescue/advanced knowledge**) was composed of knowledge related to what to do to rescue a person from an opioid overdose as well as some items corresponding to more advanced knowledge (as reflected by more difficult items on the pre-test) such as whether blood shot eyes or agitated behavior indicate an opioid overdose or how long the effects of naloxone last. Items loading on factor 3 were drawn from multiple OOKS sections B – I.

**Table 2. - ESEM Three-Factor Model Structure for Selected OOKS Pre-test Items**

Which of the following factors increase the risk of a heroin (opioid) overdose?	Factor 1 Overdose Risks		Factor 2 Overdose Signs		Factor 3 Rescue/Advanced Knowledge	
		Sig		Sig		Sig
Switching from smoking to injecting heroin (A2)	0.609	***	0.193	**	-0.023	
Using heroin with other substances, such as alcohol or sleeping pills (A3)	0.555	***	0.191		-0.007	
Increase in heroin purity (A4)	0.624	***	0.118		0.099	
Using heroin again soon after release from prison (A8)	0.490	***	0.264	***	0.074	
Using heroin again after a detoxification treatment (A9)	0.516	***	0.146	*	0.172	***
<b>Which of the following are indicators of an opioid overdose?</b>						
Having blood-shot eyes (B1)	-0.398	***	0.016		0.673	***
Slow or shallow breathing (B2)	-0.040		0.815	***	0.031	
Lips, hands or feet turning blue (B3)	0.017		0.767	***	-0.015	
Loss of consciousness (B4)	0.235	***	0.544	***	0.034	
Deep snoring (B7)	0.208	***	0.510	***	0.121	**
Very small pupils (B8)	0.143	**	0.423	***	-0.036	
Agitated behaviour (B9)	-0.288	***	0.061		0.603	***
Rapid heartbeat (B10)	-0.198	***	0.166	**	0.593	***
<b>Which of the following should be done when managing a heroin (opioid) overdose?</b>						
Give stimulants (e.g. cocaine or black coffee) (C5)	-0.211	***	-0.016		0.471	***
Place the person in the recovery position (on their side with mouth clear) (C6)	0.286	***	0.210	**	0.129	*
Put the person in bed to sleep it off (C11)	-0.213	***	0.143	*	0.391	***
<b>What is naloxone used for?</b>						
To reverse the effects of any overdose (D4)	0.053		0.030		0.654	***
<b>How can naloxone be administered?</b>						
Into mouth or swallowed orally (E4)	0.038		0.138	*	0.607	***
<b>How long do the effects of naloxone last?</b>						
2 to 6 hours (H3)	0.095		-0.104		0.691	***
<b>Please indicate which of the following statements are correct</b>						
If the first dose of naloxone has no effect a second dose can be given (I1)	0.159	***	0.062		0.626	***
Someone can overdose again even after having received naloxone (I3)	0.413	***	-0.158	**	0.526	***
Naloxone can provoke withdrawal symptoms (I6)	0.320	***	-0.167	**	0.588	***

Note. All figures shown are standardized factor loadings on each item based on exploratory structural equation modeling with target rotation for a 3-factor model. Model parameters were estimated using weighted least square with means and variance adjusted estimator. The labels in parentheses by each item reference the original OOKS instrument section and item number. Shaded items indicate the highest factor loading for that item.

\* =  $p < .05$ ; \*\* =  $p < .01$ ; \*\*\* =  $p < .001$

### **CFA and ESEM model estimation comparison**

To assess whether allowing estimation of factor cross-loadings in ESEM provided better model fit than constraining the cross-loadings to zero as in CFA, we estimated both CFA and ESEM models with and without allowing 3 selected correlated residual terms that modification indices suggested would improve model fit: OOKS items B10 (rapid heartbeat) and B9 (agitated behavior) as overdose indicators; items C11 (put person in bed to “sleep it off”) and C5 (place person in recovery position); and items I6 (provoke withdrawal symptoms) and I3 (someone can overdose again after receiving naloxone). The results of these analyses are shown in Table 3 and provide support for ESEM estimation as resulting in a better fitting model (RMSEA=.032 [95% CI=.028 - .035]; CFI=.971; TLI=.959; SRMR=.049) compared with the best fitting CFA model (RMSEA=.043 [95% CI=.040 - .046]; CFI=.934; TLI=.924; SRMR=.072).

### **Measurement invariance assessment**

Our final analyses tested whether the factor structure for the best fitting ESEM model (shown in Table 2) fit the data equally well for bystanders and first responders or whether a different factor structure would better fit the data for each group. The results of these measurement invariance analyses are also shown in Table 3. We began by estimating the configural model, which has the same overall factor structure but allows the item factor loadings and thresholds to vary across groups. This model fit the data reasonably well (RMSEA=.041 [95% CI=.038 - .045]; CFI=.935; TLI=.911; SRMR=.071). The metric model showed improved fit (RMSEA=.024 [95% CI=.020 - .028]; CFI=.974; TLI=.969; SRMR=.072) relative to the configural invariance model. The scalar invariance model fit statistics showed only slightly poorer but still very good model fit compared with the metric invariance model. Moreover, per thresholds recommended

by Finch (e.g., change in CFI  $\geq$  .01, change in SRMR  $>$  .03, change in RMSEA  $\geq$  .015) for comparing fit differences between models where differences greater than the thresholds indicate poorer fit, differences between the scalar and metric invariance models were well below the recommended thresholds. We therefore concluded that the OOKS items met the criteria for establishing scalar invariance for first responders and bystanders.

## Discussion

We believe the findings from this study have implications for assessing participants attending OEND trainings as well as for anticipating areas of training that might require more emphasis, especially among bystander trainees. From a measurement standpoint, we found that a 3-factor model adequately represented the variances and covariances among the selected OOKS items. Although our goal was not to develop a shortened version of the OOKS given the existence of the 12-item BOOK, the study that developed the BOOK also found a corresponding though not entirely overlapping 3-factor structure: opioid knowledge, opioid overdose knowledge, and opioid overdose response knowledge (23). Our analyses focused instead on determining if the selected set of OOKS items met criteria for measurement invariance among first responders and bystanders. The results supported use of this subset of OOKS items for assessing pre-training knowledge among both groups of participants. The results also supported using the items for comparative purposes given the 3-factor model evidenced scalar invariance. This indicates that comparisons of the mean scores for these two groups of OEND participants are valid. Further work would need to establish fully whether these findings are applicable to the BOOK and the OOKS parent instrument, but our findings suggest the OOKS and an item subset have broad applicability.

**Table 3. Exploratory Models, Parameter Constraints, and Fit Statistics for Increasingly Restrictive Measurement Invariance Models**

	Model Parameters			Fit Statistics				
	Factor Loadings	Item Thresholds	Residual Covariances	Chi-2 (df)	TLI	CFI	RMSEA (95% CI)	SRMR
<b>Exploratory Models (No Grouping)</b>								
CFA - Uncorrelated Residuals	Constrained	NA	None	1598.967 (206)	0.845	0.861	.061 (.059 - .064)	0.087
CFA - Correlated Residuals	Constrained	NA	Included	822.203 (203)	0.924	0.934	.043 (.040 - .046)	0.072
ESEM - Uncorrelated Residuals	Targeted	NA	None	762.781 (168)	0.919	0.941	.044 (.041 - .048)	0.062
ESEM - Correlated Residuals	Targeted	NA	Included	459.627 (165)	0.959	0.971	.032 (.028 - .035)	0.049
<b>Multi-group Measurement Invariance Models</b>								
ESEM - Configural Invariance (Correlated Residuals)	Free	Free	Included	602.752 (330)	0.952	0.966	.030 (.026 - .034)	0.061
ESEM - Metric Invariance (Correlated Residuals)	Fixed	Free	Included	593.864 (387)	0.969	0.974	.024 (.020 - .028)	0.072
ESEM - Scalar Invariance (Correlated Residuals)	Fixed	Fixed	Included	654.755 (406)	0.950	0.969	.026 (.022 - .030)	0.073

Note. All measurement invariance modeling was conducted using Exploratory Structural Equation Modeling (ESEM) with targeted rotation and weighted least squares mean and variance adjusted robust estimation using Mplus (Muthén, & Muthén, 2021). 'Free' means the model parameter was allowed to vary across groups, whereas fixed means the parameter was constrained to be equal across groups. Models where residual covariances were included allowed estimation of the residual covariances between three sets of OOKS items as identified in the text. Otherwise, residual covariances were fixed to zero. In the CFA models, cross-factor loadings were constrained to be zero whereas in the ESEM exploratory and measurement invariance models, cross-factor loadings were estimated to be as close to zero as possible.

CFI = Comparative Fit Index; TLI = Tucker Lewis Index; RMSEA = Root Mean Square Error of Approximation; and SRMR = Standardized Root Mean Square Residual.

From a practical standpoint, the results indicated that first responders are likely to have greater knowledge of the risks for an opioid-related overdose, signs of an overdose, and what to do to reverse an overdose prior to training compared with first responders. That this finding held across different groups of bystanders trained in two different counties, with considerable heterogeneity within each bystander group, gives us some though not complete confidence in the generalizability of the findings.

Whereas trainings for both groups should address all necessary knowledge relevant to recognizing and reducing opioid-related overdoses and fatalities, trainings for first responders might focus more on details such as exactly how long naloxone can be expected to last and which symptoms do not specifically indicate an opioid-related overdose (e.g., having blood-shot eyes, agitated behavior, and a rapid heartbeat). These were the issues generating the highest proportion of incorrect responses among first responders. Conversely, bystanders tended to begin trainings without knowledge in these same areas but also lacked information more generally such as whether naloxone can be used to reverse overdoses for any drug. If possible, longer OEND trainings might be needed to address the greater knowledge deficit among bystanders relative to first responders.

### **Limitations**

Our determination to seek a model with the smallest number of factors to represent the OOKS data means there is a possibility that our model is “under-factored”, and that the data might be better represented by a model with more factors. However, we assessed models with more factors and did not find they provided substantially better fit, had issues with model estimation during the measurement invariance testing, and the factors included had only several items

with loadings greater than .50. The models with more factors did not seem to improve either fit or substantive interpretation. In addition, although the collaborating sites carefully deliberated on which items to remove to avoid redundancy, it is possible that better performing items were removed or the original OOKS item set could have demonstrated better or worse measurement invariance. For this reason, we recommend a replication of this study on the full item set.

Generalizability of study findings remains a potential issue. Although we had a large and diverse sample, participants were drawn from two counties in a single midwestern state in the United States and might not be representative of OEND training participants in other parts of the country or in other countries or with other backgrounds. Moreover, the bystander group was heterogeneous and composed of people who use drugs, family members, professionals working in treatment settings, and so on. Future research should examine whether this shortened version of the OOKs works equally well for these different groups of bystanders.

Last, we also do not know how applicable the findings are with respect to post-training measurement. Clearly, an important goal of administering pre- and post-tests at trainings, is to assess training effectiveness through knowledge gain. It would therefore seem important as a next step, to assess measurement invariance across occasions rather than types of participants. This would help determine if comparing mean pre- and post-training factor scores is statistically valid (37). We also believe because of the high number of items with significant cross-loadings in the final ESEM model, a bi-factor model with a single factor representing general knowledge across items and a 3-factor representation of specific knowledge as revealed in this study is worth further exploration(38).

## Conclusions

This study supports use of a subset of items of the OOKS and inferentially the full OOKS and shorter version BOOK, as adjuncts to assess pre-training knowledge among broadly constituted groups of OEND participants. We found that the OOKS items demonstrated scalar measurement invariance supporting the use of scale score comparisons of opioid overdose related knowledge among bystanders and first responders as valid. Although not a focus of the study, we also found that bystanders tend to have larger knowledge gaps in key areas related to recognizing and reversing an opioid overdose, particularly with respect to details on naloxone use and duration of effect. These areas could be more emphasized in future trainings to address these knowledge gaps.

## Acknowledgements

We would like to thank Lauren Cox for her assistance in assembling the manuscript; Sharuti Madan for her help with obtaining and reviewing the background literature; and Kelly Horn for entering all of the study data. We also want to thank the support of the project site directors and staff that afforded access to and collected the training data for this study: Mila Tsigalis, Kathleen Burke, Rabia Mukhtar, and Scott Kaufman.

## References

1. CDC Health Alert Network. Increase in fatal drug overdoses across the United States driven by synthetic opioids before and during the COVID-19 pandemic: Centers for Disease Control and Prevention; December 17, 2020 [CDCHAN-00438:[Available from: <https://emergency.cdc.gov/han/2020/han00438.asp>].
2. Collins AB, Ndoye CD, Arene-Morley D, Marshall BDL. Addressing co-occurring public health emergencies: The importance of naloxone distribution in the era of COVID-19. *Int J Drug Policy*. 2020;83:102872.
3. Volkow ND. Collision of the COVID-19 and Addiction Epidemics. *Ann Intern Med*. 2020.
4. Kerensky T, Walley AY. Opioid overdose prevention and naloxone rescue kits: what we know and what we don't know. *Addict Sci Clin Pract*. 2017;12(1):4.
5. McDonald R, Campbell ND, Strang J. Twenty years of take-home naloxone for the prevention of overdose deaths from heroin and other opioids-Conception and maturation. *Drug Alcohol Depend*. 2017;178:176-87.
6. Ellis MS, Kasper ZA, Cicero TJ. Destigmatizing Naloxone: Associations of Perceived Availability on Opioid Use Patterns. *Ann Emerg Med*. 2021;77(1):134-5.
7. Bazazi AR, Zaller ND, Fu JJ, Rich JD. Preventing opiate overdose deaths: examining objections to take-home naloxone. *Journal of health care for the poor and underserved*. 2010;21(4):1108-13.
8. Lambdin BH, Davis CS, Wheeler E, Tueller S, Kral AH. Naloxone laws facilitate the establishment of overdose education and naloxone distribution programs in the United States. *Drug Alcohol Depend*. 2018;188:370-6.

9. Frank RG, Humpheys K, Pollack HA. Does naloxone availability increase opioid abuse?

The case for skepticism. Health Affairs Blog. March 19, 2018.

10. Kim HK, Connors NJ, Mazer-Amirshahi ME. The role of take-home naloxone in the epidemic of opioid overdose involving illicitly manufactured fentanyl and its analogs. Expert Opin Drug Saf. 2019;18(6):465-75.

11. Tas B, Humphreys K, McDonald R, Strang J. Should we worry that take-home naloxone availability may increase opioid use? Addiction (Abingdon, England). 2019;114(10):1723-5.

12. McDonald R, Strang J. Are take-home naloxone programmes effective? Systematic review utilizing application of the Bradford Hill criteria. Addiction (Abingdon, England). 2016;111(7):1177-87.

13. Seal KH, Thawley R, Gee L, Bamberger J, Kral AH, Ciccarone D, et al. Naloxone distribution and cardiopulmonary resuscitation training for injection drug users to prevent heroin overdose death: a pilot intervention study. J Urban Health. 2005;82(2):303-11.

14. Clark AK, Wilder CM, Winstanley EL. A systematic review of community opioid overdose prevention and naloxone distribution programs. J Addict Med. 2014;8(3):153-63.

15. Doe-Simkins M, Quinn E, Xuan Z, Sorensen-Alawad A, Hackman H, Ozonoff A, et al. Overdose rescues by trained and untrained participants and change in opioid use among substance-using participants in overdose education and naloxone distribution programs; a retrospective cohort study. BMC Public Health. 2014;14.

16. Walley AY, Doe-Simkins M, Quinn E, Pierce C, Xuan Z, Ozonoff A. Opioid overdose prevention with intranasal naloxone among people who take methadone. J Subst Abuse Treat. 2013;44(2):241-7.

17. Williams AV, Marsden J, Strang J. Training family members to manage heroin overdose and administer naloxone: randomized trial of effects on knowledge and attitudes. *Addiction* (Abingdon, England). 2014;109(2):250-9.
18. Williams AV, Strang J, Marsden J. Development of Opioid Overdose Knowledge (OOKS) and Attitudes (OOAS) Scales for take-home naloxone training evaluation. *Drug Alcohol Depend.* 2013;132(1-2):383-6.
19. Franklin Edwards G, 3rd, Mierisch C, Mutcheson B, Horn K, Henrickson Parker S. A review of performance assessment tools for rescuer response in opioid overdose simulations and training programs. *Prev Med Rep.* 2020;20:101232.
20. Giordano NA, Whitney CE, Axson SA, Cassidy K, Rosado E, Hoyt-Brennan AM. A pilot study to compare virtual reality to hybrid simulation for opioid-related overdose and naloxone training. *Nurse Educ Today.* 2020;88:104365.
21. Kwon M, Moody AE, Thigpen J, Gauld A. Implementation of an Opioid Overdose and Naloxone Distribution Training in a Pharmacist Laboratory Course. *American Journal of Pharmaceutical Education.* 2020;84(2):7179.
22. Stover AN, Grogg K, Patel J, Thornton D, Dwibedi N. Opioid Overdose Knowledge among College Students in a High Overdose Death State. *J Hum Behav Soc Environ.* 2019;29(7):887-96.
23. Dunn KE, Barrett FS, Yopez-Laubach C, Meyer AC, Hruska BJ, Sigmon SC, et al. Brief Opioid Overdose Knowledge (BOOK): A Questionnaire to Assess Overdose Knowledge in Individuals Who Use Illicit or Prescribed Opioids. *J Addict Med.* 2016;10(5):314-23.

24. Shoup JA, Mueller SR, Binswanger IA, Williams AV, Strang J, Glanz JM. Modifying and evaluating the Opioid Overdose Knowledge Scale for prescription opioids: A pilot study of the Rx-OOKS. *Pain Med.* 2020;21(10):2244-52.
25. Laconi S, F. P-AR, Stöver H, C. P, D. J, Meroueh F, et al. Validation of the Opioid Overdose Knowledge (OOKS) and Attitudes (OOAS) Scales in French, Spanish, and German languages, among a sample of opioid users European Addiction Research. 2021;3:1-6.
26. Asparouhov T, Muthen B, Muthen L. Exploratory structural equation modeling. *Structural Equation Modeling.* 2009;16:397-498.
27. Booth T, Hughes DJ. Exploratory structural equation modeling of personality data. *Assessment.* 2014;21(3):260-71.
28. Harris PA, Taylor R, Thielke R, Payne J, Gonzales N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Information.* 2011;42(2):377-81.
29. StataCorp. *Stata 17.1 for Mac.* College Station, TX: StataCorp; 2021.
30. Muthén B, Muthén LK. *Mplus. 8.7 ed.* Los Angeles: Muthén & Muthén; 2021.
31. R Development Core Team. *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing; 2022.
32. Revelle W. *psych: Procedures for psychological, psychometric, and personality research.* 2.0.12 ed. Northwestern University 2020.
33. Finch WH. Using fit statistic differences to determine the optimal number of factors to retain in an exploratory factor analysis. *Educ Psychol Meas.* 2020;80(2):217-41.

34. Marsh HW, Morin AJ, Parker PD, Kaur G. Exploratory structural equation modeling: an integration of the best features of exploratory and confirmatory factor analysis. *Annu Rev Clin Psychol.* 2014;10:85-110.
35. West SG, Taylor AB, Wu W. Model fit and model selection in structural equation modeling. In: Hoyle RH, editor. *Handbook of structural equation modeling.* New York, NY: Guilford Press; 2012. p. 209-31.
36. Svetina D, Rutkowski L, Rutkowski D. Multiple-group invariance with categorical outcomes using updated guidelines: An illustration using Mplus and the lavaan/semTools packages. *Structural Equation Modeling: A Multidisciplinary Journal.* 2019;27(1):111-30.
37. Putnick DL, Bornstein MH. Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Dev Rev.* 2016;41:71-90.
38. Reise SP, Moore TM, Haviland MG. Bifactor models and rotations: exploring the extent to which multidimensional data yield univocal scale scores. *J Pers Assess.* 2010;92(6):544-59.