



The author(s) shown below used Federal funding provided by the U.S. Department of Justice to prepare the following resource:

Document Title: Mystery Shop Programs to Reduce Underage Alcohol Sales

Author(s): Brad S. Krevor, Ph.D., Joel Grube, Ph.D., William DeJong, Ph.D.

Document Number: 251090

Date Received: August 2017

Award Number: 2012-AH-FX-4004

This resource has not been published by the U.S. Department of Justice. This resource is being made publically available through the Office of Justice Programs' National Criminal Justice Reference Service.

Opinions or points of view expressed are those of the author(s) and do not necessarily reflect the official position or policies of the U.S. Department of Justice.

Mystery Shop Programs to Reduce Underage Alcohol Sales

Brad S. Krevor, Ph.D, Responsible Retailing Forum
Joel Grube, Ph.D, Prevention Research Center/PIRE
William DeJong, Ph.D, Boston University

A. Abstract

This project examined a new community-level intervention strategy to achieve compliance with alcohol sales-to-minors laws: Mystery Shops. Mystery Shops are attempted purchases of age-restricted products by young, legal-age inspectors for the purpose of providing feedback to licensees on actual staff age-verification conduct. Unlike law enforcement inspections, licensees and staff face no legal penalties for failing to check the ID of someone young enough to trigger an ID-check but not under the legal purchase age. Mystery shops have shown to be effective with large national chains that implemented programs under agreements with state attorneys general. Researchers working with the Responsible Retailing Forum hypothesized that Mystery Shops will similarly improve staff performance for independently owned and operated licensees. To test this, we partnered with the Oregon Liquor Control Commission and the Texas Alcoholic Beverage Control Commission to conduct 13 monthly Mystery Shops in 24 communities, 12 in each state: in 8 communities, licensees received on-the-spot feedback and follow-up reports; in 8 communities, licensees received reports on aggregate community-level performance and Responsible Retailing best practices; and 8 other communities served as no-intervention controls. If proven to be effective, Mystery Shops will provide communities with a low-cost means for assisting licensees to reduce underage sales without taxing the limited resources of law enforcement.

Table of Contents

A. Executive Summary	p. 4
B. Overview	p. 7
C. Study Design	p. 10
D. Outcomes	p. 12
E. Conclusions	p. 14
F. Limitations	p. 15
G. References	p. 16

Acknowledgements

This project was supported by Grant #2012-AH-FX-4004 awarded by the Office of Juvenile Justice and Delinquency Prevention, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice.

The authors thank the Oregon Liquor Control Commission and the Texas Alcoholic Beverage Commission for their assistance in this project.

A. Executive Summary

Background

Underage alcohol use is associated with a range of negative outcomes, including increased risk of traffic crashes (Yi, Williams, & Dufour, 2001) and highway fatalities (NHTSA, 2002), accidental injuries and assaults (Hingson, 2002), vandalism and property damage (Wechsler et al., 1995), high-risk sexually activity (Strunin & Hingson, 1992; Tapert et al., 2001), damage to the developing brain (Spear, 2002), and future alcohol dependence (Grant et al., 2006; Agrawal et al., 2009). Reducing adolescent access to commercial alcohol sources has been recommended by the Institute of Medicine (IOM) report, *Preventing Underage Drinking: A Collective Responsibility* (IOM, 2003), the Surgeon General's *Call to Action to Prevent Underage Drinking* (US HHS, 2007), and the Federal Trade Commission's *Report to Congress: Alcohol Marketing and Advertising* (FTC, 2003) as a key strategy for reducing the societal and personal costs of adolescent alcohol use.

The principal approach for reducing underage access to commercial sources of alcohol has been law enforcement compliance checks, in which underage decoys attempt to purchase an alcoholic beverage while being observed by a law enforcement officer who cites the seller and/or licensee if an unlawful sale is made. Although they can be effective, some limitations of compliance checks can be noted:

- The effects of enforcement upon retailer behavior are relatively short-lived and often modest (Wagenaar, Toomey, & Erickson (2005);
- Enforcement compliance checks are labor-intensive and expensive and thus cannot be implemented with sufficient frequency to maintain high levels of compliance with underage sales laws; and
- Compliance checks alone do not show licensees what steps to take to avert non-compliance in the future.

That is, enforcement is necessary, but insufficient. Identifying other approaches to achieve compliance with sales-to-minors laws is therefore a priority.

Mystery Shop programs, which use young, but legal-age customers who attempt to purchase an age-restricted product for the purpose of observing staff ID-checking conduct, are a possible supplement to enforcement compliance checks. These programs provide immediate feedback to clerks about their performance and reports to owners/managers about staff performance. They have the advantage of being less labor-intensive and less expensive than enforcement compliance checks. They can be undertaken by retailer groups, community organizations, or public health agencies without the involvement of enforcement agencies. As a result, they can be implemented with greater frequency than enforcement compliance checks and with a broader coverage of outlets. Mystery Shop programs also provide an opportunity for retailer education on responsible sales and service practices and can motivate retailers to take steps to improve staff performance relating to underage sales.

In the present study, we addressed the following questions: 1) Will individualized Mystery Shop reports on staff performance sent to independently owned and operated licensees improve adherence to age-verification protocols? 2) Will broadly disseminated, non-individualized reports, using aggregated community-level Mystery Shop data, also improve adherence to age-verification protocols? If effective, programs using aggregated, community-level data would further reduce costs and improve program reach.

Study Design

The study was a randomized community trial that assessed the impact of Mystery Shops for off-premise alcohol establishments on ID-checking for age-verification. Twenty-four communities in Oregon and Texas were matched on baseline rates of ID checking for alcohol sales to young-appearing mystery shoppers and then randomly assigned to one of three study arms:

Direct Feedback Reports (8 communities). Licensees received direct feedback on staff ID-checking behavior and RR resources (Arm 1);

Community-Level Reports (8 communities). Licensees received periodic reports on aggregate results of ID-checking behavior and RR resources (Arm 2);

Control/No Reports (8 communities). Licensees in these communities received no feedback or community-level reports (Arm 3).

Within each community 11 or more outlets were randomly selected for inclusion in the study. Three monthly baseline mystery shops were conducted during which no feedback was provided, after which the communities were matched and randomly assigned to condition. The baseline was followed by 10 monthly Mystery Shop inspections. For outlets in communities assigned to the direct feedback condition, these inspections were reported to staff in real time and through follow-up letters sent to managers, which also described various responsible retailing strategies to promote staff ID-checking. For communities in the community-level report condition, these inspections were not reported to management or staff. Instead, we sent quarterly reports to every off-premise licensee in the community, whether inspected or not, that summarize the aggregated results for the community's Mystery Shop inspections during the previous 3 months. The responsible retailing resources to promote staff ID-checking were included with the reports.

Results

ID-checking rates for the communities in the Direct Feedback condition increased from 77.8% to 83.3% from baseline to the intervention period, and for communities in the Community Feedback condition they increased from 76.0% to 80.9%. For communities in the Control condition, they decreased from 77.7% to 76.8%. Generalized Linear Mixed Model logistic regressions indicated that the Direct Feedback Condition had a marginally significant ($p < .08$) effect when compared with the Control condition. Specifically, providing Direct Feedback was related to a 53% increase in the odds of checking IDs. Although not statistically significant ($p < .19$), the Community Feedback Condition was associated with a 35% increase in the odds of checking IDs. In addition, a large difference in ID-checking rates was observed between Oregon and Texas ($p < .002$); the odds of a clerk in Oregon checking IDs was 244% greater than that for a clerk in Texas. Notably, the ID-checking rates in Oregon were very high during the baseline period ($\approx 86\%-90\%$), leaving little room for improvement as a result of the interventions. In fact, ID-checking rates declined in Oregon between baseline and the intervention period; although not a significant difference, these declines were somewhat smaller in the Direct and Community-level feedback conditions compared with the Control condition. Because of the large difference in ID-checking rates between Oregon and Texas, we conducted separate analyses to explore whether the interventions were differentially effective in the two states. Consistent with the main analyses, the interventions were each related to increases in the relative odds of ID-checking, although the effect was not statistically significant for either state alone, perhaps due to the smaller sample size for each state.

Conclusion

The primary analysis indicated that the Direct Feedback intervention was associated with a modest and marginally significant increase in ID-checking. This effect was smaller than has been observed in previous studies using similar mystery shop feedback interventions. A number of factors may have contributed to this situation. Notably, the ID-checking rates in Oregon were very high even before the intervention, leaving little room for improvement. Rates of ID-checking in Texas were much lower, and the absolute percentage improvement observed there was substantially greater than that observed in Oregon. Unfortunately, the study was not powered to detect intervention effects at the individual state level. It is also unclear whether the Community Feedback intervention is effective in increasing rates of ID-checking. Looking at the percentage data, the Community Feedback intervention had an equivalent effect to the Direct Feedback intervention in Texas (i.e., $\approx 10\%$ increase in ID-checking). In Oregon, the Community Feedback condition appeared to be somewhat less effective. Given the lack of statistical significance,

however, this pattern is difficult to interpret. With full-scale implementation, providing community-level feedback based on monthly random samples of retailers would be far less costly than providing direct feedback to every retailer. For that reason, and because of the suggestive results seen in this investigation, additional research should be conducted to explore the viability of this option.

Limitations

Some limitations of the study should be noted. Most importantly, very high ID-checking rates were observed in Oregon during the baseline period. This likely created a ceiling, making it very difficult to detect intervention effects. Unfortunately, the study was not sufficiently powered in terms of numbers of communities to detect intervention effects within the two states separately. Future studies should include more communities across more states to more definitively address whether community-level feedback can be an effective approach to increasing retailers' compliance with underage sales laws.

An additional limitation relates to the frequency of the Community-level reports. Whereas the Direct reports were delivered monthly, the Community reports were delivered only quarterly. Thus, frequency of report was confounded with type of report. As a result, any difference between the two types of reports could be due to differences in frequency or format. It is unknown whether more frequent Community-level reports may have been more effective and would have produced similar effects as observed for Direct reports. These are issues that should be considered in future research.

B. Overview of the Problem

The harm due to alcohol use by adolescents has been brought to national attention by the Institute of Medicine (IOM) report, *Preventing Underage Drinking: A Collective Responsibility* (IOM, 2003) and by the Surgeon General's *Call to Action to Prevent Underage Drinking* (US HHS, 2007). Underage alcohol use is associated with increased risk of traffic crashes (Yi, Williams, & Dufour, 2001) and highway fatalities (NHTSA, 2002), accidental injuries and assaults (Hingson, 2002), vandalism and property damage (Wechsler et al., 1995), high-risk sexually activity (Strunin & Hingson, 1992; Tapert et al., 2001), damage to the developing brain (Spear, 2002) and future alcohol dependence (Grant et al., 2006; Agrawal et al., 2009). The primary sources of alcohol beverages for minors are acquaintances and friends, both under and over age 21 (Paschall et al., 2007; Dent et al., 2005; Harrison et al., 2000), plus parents and other adults (Century Council, 2002). Nonetheless, 15% to 30% of minors (Century Council, 2002; Dent, Grube, & Biglan, 2005; Paschall et al., 2007) report that they purchase alcohol beverages from retailers directly, and many of these youth become the source of alcohol for their peers. Research has confirmed minors' ability to purchase alcohol beverages (Preusser & Williams, 1992; Forster et al., 1994; Durkin, Wolfe, & Phillips; 1996; Wolfson et al., 1996; Freisthler et al., 2003; Paschall et al., 2007; Toomey et al., 2008), with or without a fake ID. Consequently, reducing adolescent access to commercial alcohol sources has been recommended by the IOM Report, the Surgeon General's report, and the Federal Trade Commission's *Report to Congress: Alcohol Marketing and Advertising* (FTC, 2003) as a key strategy for reducing the societal and personal costs of adolescent alcohol use.

The principal approach for reducing underage access to commercial sources of alcohol has been law enforcement compliance checks in which underage decoys attempt to purchase an alcoholic beverage while being observed by a law enforcement officer who will cite the seller and/or licensee if an unlawful sale is made. Compliance checks have been demonstrably successful in reducing the ability of adolescents to purchase alcohol (Wagenaar & Wolfson, 1995; Grube, 1997; Preusser, Williams, & Weinstein, 1994; Wagenaar et al., 2000). The CDC's *Community Guide on Preventive Practices* recommends compliance checks to enforce sales-to-minors laws (CDC, 2006). A report developed for the Center for Substance Abuse Prevention (CSAP) as a federal guidance document, *Best Practices for Responsible Retailing* (CSAP, 2003), characterized law enforcement as the keystone strategy for preventing sales to minors: absent law enforcement of sales-to-minors laws, most retailers will not commit time and resources to developing employee policies and point-of-sales protocols, hiring and training committed staff, and providing management supervision, all of which additively contribute to compliance with alcohol sales laws and practices that are collectively termed "Responsible Retailing" (RR).

It is important to note, however, that law enforcement compliance checks are limited in three important respects: 1) The effects of enforcement upon retailer behavior are short-lived and modest. Wagenaar, Toomey, & Erickson (2005) found that underage sales refusals improved 17% immediately following a compliance check, but within 2 months fell to only 8.5% at on-premise establishments, and within 3 months fell to 0% at off-premise retailers. Extremely high staff turnover—e.g., in convenience stores, estimates of turnover are 200-300% a year (Trayte, 2006)—further limits the deterrent effect of compliance checks on future staff conduct. 2) Enforcement is labor-intensive and expensive. Typically, a single officer or a two-person team of officers accompany an underage decoy who attempts to make an alcohol purchase. Costs for compliance checks can vary greatly: clearly, a protocol that calls for two officers rather than one

increases costs, and higher violation rates require substantially more time for issuing and processing citations. For example, in Green Bay, WI, compliance checks of eight licensees in April, 2012, each involving two-person teams of police officers, required 16 man-hours, inclusive of planning and processing citations, which incurred \$640 in labor costs, or \$80 per inspection, based upon a normal duty rate of \$40/hour, which would have risen to \$110 per inspection if based upon the overtime rate of \$60/hour (Bongle, 2012). The intensive manpower needed for compliance checks means that alcohol law enforcement must always compete with other public safety concerns, many of which may be seen as more “serious” than enforcing sales-to-minors laws. 3) Compliance checks do not show licensees what steps to take to avert non-compliance in the future. For most violations of law, the specific steps by which an individual or business can avert a repeat infraction are self-evident. Being aware of the age 21 law, retailers already have explicit policies for verifying age and refusing sales to underage customers, plus explicit penalties for failing to do so (Krevor, Lieberman, & Gerlach, 2002). There is, however, a significant gap between policies that mandate ID checks for customers under a designated “trigger age” (typically 30 years), and actual clerk/server conduct at the point of sale. After conducting a study of clerks in North Carolina who had sold alcohol to undercover decoys, former Alcohol Law Enforcement Director William Chandler, Ph.D., concluded that a sale to minors is, in most cases, a “crime of inattention” (Chandler & Vasu, 2005). As Chandler explains, sales to minors often occur because the clerk is multi-tasking (e.g., monitoring other customers, overseeing deliveries, processing lottery tickets on a separate register) or simply loses focus, due to the highly repetitive behavior involved in ringing up alcohol sales (In this respect, underage sales are similar to industrial accidents.) At best, penalties for sales to minors will only temporarily reduce this susceptibility to inattention (Wagenaar et al., 2005)

This, then, is the underlying problem of relying upon enforcement alone to prevent unlawful sales to minors: compliance checks remind licensees that they are subject to occasional inspection, and that establishments whose staff sell to minors will face legal consequences; but because compliance checks are labor-intensive and costly, most law enforcement agencies cannot conduct them anywhere near the frequency needed to sustain high levels of compliance. Moreover, compliance checks do not show licensees the responsible retailing measures they can take to prevent future underage sales. *In short, enforcement is necessary but insufficient.* As a result, identifying other approaches to achieve compliance with sale-to-minors laws becomes a priority.

Mystery Shop programs – in which young, but legal-age customers attempt to purchase an age-restricted product for the purpose of observing staff conduct – can fulfill this need. Originally, Mystery Shops were used as a measurement tool to assess the impact of an RR practice or technology. Mystery Shops have been used to examine the impact of training programs on age-verification conduct, to determine the frequency of use and effectiveness of electronic ID readers (Krevor et al., 2003), and to measure the effectiveness of Assurances of Voluntary Compliance (AVCs) executed by state attorneys general and large national chains (Krevor, Lieberman, & Gerlach, 2002). One of the AVC signatory chains, ExxonMobil, posted Mystery Shop results on a management website for all store managers to view and required re-training for the staff of any gas station/convenience store that failed a Mystery Shop. It became apparent to the AVC signatory chains that Mystery Shops were not just a measurement tool: because the feedback itself could impact behavior, Mystery Shops were also an intervention.

Mystery Shops have been shown to be an effective tool for large chains. Our analysis of

Mystery Shops conducted by the BARS Program for national and regional chains between 2005-2012 indicates that grocery stores, convenience store/petroleum marketers, and “big box” chains all achieved correct ID-checking performance rates in excess of 91% after participating in a Mystery Shop program, while liquor stores chains achieved rates in excess of 95% (Fournier, 2012). As shown in Table 1, pass rates for the national and regional chain outlets participating in the BARS Program improved steadily over time, increasing from 85% to just under 90% by the 10th month of participation. The improvements in ID-checking are statistically significant, but it should also be noted that the baseline rates are quite high, which reflects the numerous RR practices that these chains had adopted before instituting monthly MS inspections. Since independently owned and operated alcohol licensees seldom employ formal clerk/server training programs or sophisticated point-of-sales technologies for age-verification (RRForum, 2009), we anticipate much lower baseline rates for these licensees and correspondingly greater improvements in ID-checking as a result of a MS inspection program.

Table 1. Pass Rates for National and Regional Chains by Number of Months in the BARS Program

Month	Pass Rate	Count	Month	Pass Rate	Count
Month 1	85.07%	19,066	Month 8	89.76%	8,113
Month 2	87.84%	10,297	Month 9	89.49%	8,837
Month 3	88.09%	9,925	Month 10	89.90%	7,868
Month 4	89.05%	9,985	Month 11	90.61%	8,224
Month 5	88.95%	8,981	Month 12	90.56%	7,208
Month 6	89.33%	8,928	Month 13	90.50%	8,135
Month 7	89.70%	9,420			

In the Mystery Shop protocol used in this study, an unfamiliar customer, male or female, age 21-24, enters an alcohol retail outlet and attempts to purchase an alcoholic beverage. If the clerk/server requests and examines an ID, the Mystery Shopper presents the clerk/server with a Green Card that explains that an ID-check had been correctly performed and complements the clerk/server. If the clerk/server fails to request an ID, or is willing to sell without an ID being presented, the Mystery Shopper presents the clerk/server and the manager on duty with a Red Card that indicates a failed ID-check and outlines the possible penalties if this had instead been a law enforcement compliance check. In addition, a written follow-up report summarizing the Mystery Shop outcome is mailed to the store owner or manager, along with RR materials for the licensee’s use. Because Mystery Shoppers are of legal age, the failure to check IDs may be a violation of store policy but not a violation of state or local law: neither the clerk/server nor the licensee is exposed to criminal or administrative penalties. Nevertheless, results of inspections

are confidential and are shared only with the licensee. Thus, Mystery Shops are a quality improvement tool that serve as a proxy for law enforcement compliance checks by reinforcing the need for consistent age-verification and reminding licensees and their staff of the consequences of illegal sales to minors.

To date, Mystery Shops have been employed primarily by retail chain stores and multi-location independents, and the data presented above shows that direct feedback on staff ID-checking conduct improves future staff conduct. What had not been determined is whether community-level aggregate reports on performance will also improve retailer staff conduct. Accordingly, a key objective of this project was to determine if Mystery Shop programs with off-premise alcohol retailers can be employed as a *community* intervention to increase ID-checking for age-verification and thereby reduce underage alcohol sales to minors. We addressed the following questions: 1) Will individualized Mystery Shop reports on staff performance sent to independently owned and operated licensees improve adherence to age-verification protocols, compared with no-intervention controls? 2) Will broadly disseminated, non-individualized reports, using community-level Mystery Shop data, also improve adherence to age-verification protocols?

C. Study Design

This study was a randomized community trial that assessed the impact of Mystery Shops for off-premise alcohol establishments on ID-checking for age-verification. Twenty-four communities in Oregon and Texas were assigned to one of three study arms:

- Direct Feedback Reports (8 communities). Licensees receive direct feedback on staff ID-checking behavior and RR resources
- Community-Level Reports (8 communities). Licensees receive periodic reports on aggregate results of ID-checking behavior and RR resources.
- Control/No Reports (8 communities). Licensees in these communities received no feedback or community-level reports.

Selection of communities. With the assistance of the Oregon Liquor Control Commission and the Texas Alcoholic Beverage Control Commission, we identified 12 communities in each state with 15 or more off-premises alcohol sales licensees. No communities were contiguous.

Random Assignment of Communities to Treatment Conditions. We conducted 3 baseline Mystery Shop inspections (protocols are described below) at 11 or more randomly selected licensees in each of the 24 communities, and then calculated each community's overall baseline pass rate. Based on these rates, we stratified the communities into three groups—high, medium, and low compliance.

Next, we randomly assigned communities within each group to the three treatment conditions: *Study Arm 1: Direct Feedback Reports; Study Arm 2: Community-Level Reports; and Study Arm 3: Control/No Reports*. Note that we inspected 11 to 13 licensees per site to allow for possible attrition over the course of the study. Because of random assignment within these baseline sales strata, we expected the outlets in the different conditions to show very similar sales rates before the intervention period. The research design is summarized in Table 2.

To launch the study, RRF and OLCC or TABC mailed an announcement to all licensees in the two intervention arms describing the study and indicating that they might receive periodic Mystery Shops—for educational, and not enforcement purposes—sometime in the future. The participating licensees in Study Arm 1 (Direct Feedback Reports) and Study Arm 2 (Community-Level Reports)

were not aware that the Mystery Shops had been initiated until month 4, after completion of the three-month baseline period.

Study Arm 1: Direct Feedback Reports (8 communities, 99 licensees). After the 3 baseline inspections, we conducted 10 monthly Mystery Shop inspections (months 4-13) at each of the selected off-premise establishments. These inspections were reported to staff in real time with Green or Red Cards and through follow-up letters sent to managers, which also describe various RR strategies to promote staff ID-checking.

Table 2: Mystery Shop Inspections and Reports for Study Arms 1, 2, and 3

Experimental Condition	1	2	3	4	5	6	7	8	9	10	11	12	13
1) Direct Feedback Reports				F	F	F	F	F	F	F	F	F	F
2) Community-Level Reports				Q			Q			Q			Q
3) Control/No Reports													
Mystery Shop Inspections	B	B	B	I	I	I	I	I	I	I	I	I	I

F = Monthly feedback (Green or Red Cards, follow-up letter to managers); Q = Quarterly community-level report for previous three months of inspections; B= Baseline Mystery Shop inspection; I = Mystery Shop inspection.

Study Arm 2: Community-Level Reports (8 communities, 102 licensees). After the 3 baseline inspections, we conducted 10 monthly Mystery Shop inspections (months 4-10) at each of the selected off-premise establishments. These inspections were not reported to management or staff. Instead, we sent *quarterly reports* to every off-premise licensee in the community, whether inspected or not, that summarized the aggregated results for the community’s Mystery Shop inspections during the previous 3 months. Additional RR resources to promote staff ID-checking were included.

3) Study Arm 3: Control/No Reports (8 communities, 97 licensees). After the 3 baseline inspections, we conducted 10 monthly Mystery Shop inspections (months 4-13) at each of the selected off-premise establishments. Staff and managers in these communities received no feedback or community-level reports.

Mystery Shoppers. The Mystery Shoppers were young, legal-age males and females, age 21-24, recruited and trained by the BARS Program, a Colorado-based vendor that has conducted inspections for several of the national AVC signatory chains, and which the research team had employed in past research projects. Mystery Shoppers were instructed to make no attempt to alter their apparent age.

Mystery Shop Inspection Protocols. Consistent with the BARS protocol, the Mystery Shop inspections were scheduled for varying times of day and days of the week.

One protocol was used for the unreported Mystery Shops, including the 3 baseline inspections for Study Arm 1, and all 13 inspections in Study Arms 2 and 3. The Mystery Shopper (MS)

entered the retail establishment and attempted to purchase an alcoholic beverage. If the clerk requested an ID, the MS stated that he/she brought money (clearly visible in his/her hand), but not a wallet. If the clerk refused to complete the transaction without an ID, the MS replied that he/she needed to get his/her wallet and then left the store. If the clerk was willing to complete the transaction without an ID, the MS would purchase the alcoholic beverage and leave the store.

In Study Arm 1, for inspections 4 - 13, the Mystery Shopper entered the retail establishment and similarly attempted to purchase an alcoholic beverage. If the clerk requested an ID, the MS stated that he/she brought money (clearly visible in his/her hand), but not a wallet. If the clerk refused to complete the transaction without an ID, the MS presented the clerk with a Green Card that explained that an ID-check had been correctly performed and complemented the clerk. If the clerk failed to request an ID or was prepared to make a sale without an ID, the MS presented the clerk and the manager on duty with a Red Card that indicated a failed ID check and outlined the possible penalties if this had instead been a compliance check by law enforcement.

Multiple MS were used to avoid having the same shopper visit the same store multiple times. For all inspections, if the clerk on duty had waited on the MS in a prior inspection, the MS was instructed to leave the establishment and another inspection was conducted at a later time.

D. Outcomes

Analysis strategy. The initial analyses consisted of descriptive statistics and data plots of the mystery shop results over time for the *Direct Feedback*, *Community Feedback*, and *No Feedback Control* groups. The primary outcome analyses were a series of Generalized Linear Mixed Model logistic regressions of the mystery shop pass rates (0= failed to check ID; 1 = checked ID). These analyses took into account the nesting of observations (mystery shops) within outlets and outlets within cities. State was used as a covariate in order to control for differences between Oregon and Texas in ID-checking rates. The intervention groups were coded as two dummy variables that individually represented the Direct Feedback and Community Feedback conditions, with the Control condition as the contrast group. A variable representing the 3 baseline mystery shops

versus the 10 intervention period mystery shops (0 = Baseline; 1 = Intervention) was included, as were the interactions between each intervention condition variable and baseline vs. intervention period. These latter effects are the primary outcomes of interest because they represent differential change in ID-checking between

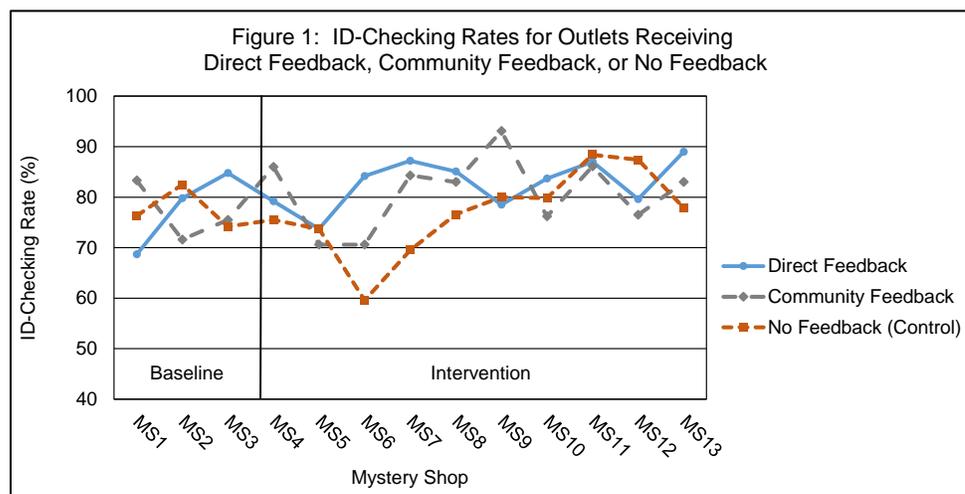


Table 3
Mean ID-Checking Rates During Baseline and Intervention Periods for Direct Feedback, Community Feedback, and No Feedback Control Groups

Condition	Baseline	Intervention Period	Change in ID-checking
Direct Feedback	77.8%	83.3%	+ 5.5%
Community Feedback	76.8%	80.9%	+ 4.1%
No Feedback (Control)	77.7%	76.8%	-0.9%

the treatment and control conditions from baseline to the intervention period. Follow-up analyses investigated the effects of the combined treatments (i.e., any intervention) on ID-checking rates. Separate analyses were also conducted by state.

Descriptive analyses. Figure 1 depicts the pass rates for the two intervention groups and control group over the baseline and intervention mystery shop periods. Overall, the three groups showed very similar rates of ID-checking at baseline, averaging 77.8% for the Direct Feedback condition, 76.8% for the Community Feedback condition, and 77.7% for the No Feedback Control condition. ID-checking rates improved during the intervention period for the two intervention groups, averaging 83.3% (+5.5%) for the Direct Feedback group and 80.9% (+4.1%) for the Community Feedback group, compared with 76.8% (-0.9%) for the Control group (Table 3).

Primary outcome analysis. The effects of greatest interest are the interactions of Time (baseline vs. intervention period) with the Direct Feedback and Community Feedback conditions. These interactions represent the differential change in ID-checking between each group and the control condition from

before to after the interventions were implemented. The logistic regression analysis (Table 4) indicated that the Direct Feedback Condition (OR = 1.53, $p < .08$) had a

marginally significant effect when compared with the Control condition. That is, providing Direct Feedback was related to a 53% increase in the odds of checking IDs. Although not statistically significant, the Community Feedback Condition (OR = 1.35, $p < .19$) was associated with a 35% increase in the odds of checking IDs. In addition, a large difference in ID-checking rates was observed between Oregon and Texas (OR = 2.44, $p < .002$); the odds of a clerk in Oregon checking IDs was 244% greater than that for a clerk in Texas.

Any intervention vs. no intervention. We conducted an additional analysis to explore whether providing

any feedback (either Direct or Community) improved ID-checking rates (Table 5). The results of this analysis paralleled

those for the separate interventions. Overall, providing any feedback was associated with a 43% increase in the odds of ID-checking (OR = 1.43). This effect was marginally statistically significant ($p < .08$). Again, there was a large difference between Oregon and Texas in ID-checking rates.

Table 4
Results of Generalized Linear Mixed Model Logistic Regression Testing Effects of Direct and Community Feedback on ID-Checking

Predictor	Regression Coefficient	Standard Error	t	p	Odds Ratio	95% CI Odds Ratio
State (Oregon)	.893	.243	3.68	< .002	2.44	1.47, 4.05
Direct Feedback	.024	.346	.07	.95	1.02	.51, 2.07
Community Feedback	-.076	.343	-.22	.83	.93	.46, 1.86
Time (Intervention vs. Baseline)	-.043	.166	-.26	.80	.96	.69, 1.33
Direct Feedback x Time	.424	.238	1.78	.08	1.53	.96, 2.44
Community Feedback x Time	.301	.231	1.30	.19	1.35	.86, 2.13

Table 5
Results of Generalized Linear Mixed Model Logistic Regression Testing Effects of Any Feedback on ID-Checking

Predictor	Regression Coefficient	Standard Error	t	p	Odds Ratio	95% CI Odds Ratio
State (Oregon)	.889	.239	3.73	< .001	2.43	1.48, 4.03
Any Feedback	-.025	.295	-.08	.93	.975	.54, 1.77
Time (Intervention vs. Baseline)	-.043	.166	-.26	.80	.96	.69, 1.33
Any Feedback x Time	.358	.203	1.77	.08	1.43	.96, 2.13

Analyses by state. Because of the large difference in ID-checking rates between Oregon and Texas, we conducted separate analyses to explore whether the interventions were differentially effective in the two states. Table 6 shows the results of these analyses. Consistent with the previous analyses, the interventions were each related to increases in the odds of ID-checking, although the effect was not statistically significant for either state alone.

Table 6
Results of Generalized Linear Mixed Model Logistic Regression Testing Effects of Direct and Community Feedback on ID-checking by State

Predictor	Regression Coefficient	Standard Error	t	p	Odds Ratio	95% CI Odds Ratio
Oregon						
Direct Feedback	.076	.601	.13	.90	1.08	.31, 3.80
Community Feedback	-.239	.584	-.41	.69	.79	.23, 2.70
Time (Intervention vs. Baseline)	-.532	.294	-1.81	.07	.59	.33, 1.05
Direct Feedback x Time	.553	.432	1.28	.20	1.74	.75, 4.06
Community Feedback x Time	.355	.398	.89	.37	1.43	.65, 3.12
Texas						
Direct Feedback	-.045	.441	-.10	.92	.96	.40, 2.27
Community Feedback	.028	.441	.06	.95	1.03	.43, 2.44
Time (Intervention vs. Baseline)	.235	.207	1.14	.26	1.27	.84, 1.90
Direct Feedback x Time	.314	.293	1.07	.28	1.37	.77, 2.43
Community Feedback x Time	.305	.292	1.05	.30	1.36	.77, 2.40

Table 7
Mean ID-checking Rates During Baseline and Intervention Periods for Direct Feedback, Community Feedback, and No Feedback Control Groups by State

Condition	Baseline	Intervention Period	Change in ID-checking
Oregon			
Direct Feedback	89.8%	88.5%	-1.3%
Community Feedback	86.3%	84.1%	-2.2%
No Feedback (Control)	88.9%	82.9%	-6.0%
Texas			
Direct Feedback	66.0%	77.0%	10.0%
Community Feedback	67.3%	77.4%	10.1%
No Feedback (Control)	66.7%	71.8%	5.1%

Table 7 shows the mean ID-checking rates for Oregon and Texas by condition for the baseline and intervention periods. There are several things to note. First, consistent with the results of the regression analyses, ID-checking rates were much higher in Oregon than in Texas both at baseline and after implementation of the interventions. Notably, the ID-checking rates in Oregon were very high during the baseline period ($\approx 86\%$ - 90%), leaving little room for improvement as a result of the interventions. Second, ID-checking rates overall fell slightly in Oregon over the course of the project, whereas they increased in Texas. Third, the decreases in Oregon were greater in the outlets in the Control group than for the outlets in the intervention groups. That is, although the interventions did not increase ID-checking rates in Oregon, they may have prevented or lessened deterioration of the rates. Finally, although ID-checking rates increased among all three groups of outlets in Texas, these increases were somewhat greater among outlets in the two intervention conditions.

E. Conclusion

The primary analysis indicates that the Direct Feedback intervention was associated with a modest ($OR = 1.53$) and marginally significant ($p < .08$) increase in ID-checking. This effect is somewhat smaller than has been observed in previous studies using similar mystery shop feedback interventions. A number of factors may have contributed to this situation. Notably, the ID-checking rates in Oregon were very high even before the intervention, leaving little room for improvement. Rates of ID-checking in Texas were much lower, and the absolute percentage improvement observed there was substantially greater than that observed in Oregon. Unfortunately, the study was not powered to detect intervention effects at the individual state

level. It is also unclear whether the Community Feedback intervention is effective in increasing rates of ID-checking. Looking at the percentage data, the Community Feedback intervention had an equivalent effect to the Direct Feedback intervention in Texas (i.e., $\approx 10\%$ increase in ID-checking). In Oregon, the Community Feedback condition appeared to be somewhat less effective in moderating the decrease in ID-checking than was the Direct Feedback condition (-2.2% vs. -1.3%). Given the lack of statistical significance, however, this pattern is difficult to interpret. With full-scale implementation, providing community-level feedback based on monthly random samples of retailers would be far less costly than providing direct feedback to every retailer. For that reason, and because of the suggestive results seen in this investigation, additional research should be conducted to explore the viability of this option.

F. LIMITATIONS

Some limitations of the study should be noted. Most importantly, very high ID-checking rates were observed in Oregon during the baseline period. This likely created a ceiling, making it very difficult to detect intervention effects. Unfortunately, the study was not sufficiently powered in terms of numbers of communities to detect intervention effects within the two states separately. Future studies should include more communities across more states to more definitively address whether community-level feedback can be an effective approach to increasing retailers' compliance with underage sales laws. An additional limitation relates to the frequency of the Community-level reports. Whereas the Direct reports were delivered monthly, the Community reports were delivered only quarterly. Thus, frequency of report was confounded with type of report. As a result, any difference between the two types of reports could be due to differences in frequency or format. It is unknown whether more frequent Community-level reports may have been more effective and would have produced similar effects as observed for Direct reports. These are issues that should be considered in future research.

G. References

- Agrawal A, Sartor CE, Lynskey MT, Grant JD, Pergadia ML, Gruzca R, Bucholz KK, Nelson EC, Madden PA, Martin NG, Heath AC. (2009). Evidence for an interaction between age at first drink and genetic influences on DSM-IV alcohol dependence symptoms. *Alcohol Clin Exp Res*. Dec;33(12):2047-56. Epub 2009 Sep 17.
- Bongle, Capt. William (2012). Phone conversation May 14, 2012.
- Centers for Disease Control and Prevention (2006). *Guide to Community Preventive Services. Preventing excessive alcohol consumption: enhanced enforcement of laws prohibiting sales to minors*. Retrieved from: www.thecommunityguide.org/alcohol/lawsprohibitingsales.html.
- Center for Substance Abuse Prevention (2002). Best Practices for Responsible Retailing. U.S. Department for Health and Human Services, Substance Abuse and Mental Health Services Administration, Centers for Substance Abuse Prevention, Conference Edition Draft, March, 2002. Available at <http://www.rforum.org/reports>.
- Century Council. (2002). National Research on Underage Alcohol Access and Parental Discussions, Washington, DC.
- Chandler WC, Vasu ML (2005). Reducing the Sale of Alcohol to Minors by Targeting Sellers and Licensed Outlets: The Case of North Carolina. *Law Enforcement Executive Forum* 5 (3), 75 - 92.
- Dent C, Grube J W, Biglan A. (2005). Community Level Alcohol Availability and Enforcement of Possession Laws as Predictors of Youth Drinking. *Preventive Medicine*, 40, 355-362.
- Durkin KF, Wolfe TW, Phillips, DW III. (1996). College students' use of fraudulent identification to obtain alcohol: an exploratory analysis. *J of Alcohol and Drug Education* 41(2):92-104.
- Flewelling, R., Grube, J.W., Paschall, M.J., Biglan, A., Kraft, A., Black C., Hanley, S., Ringwalt, C., Wiesen, C., Ruscoe, J. (in press). Reducing youth access to alcohol: Findings from a community-based randomized trial. *American Journal of Community Psychology*. PMC Journal-In Process.
- Freisthler B, Gruenewald PJ, Treno AJ, Lee J. (2003). Evaluating alcohol access and the alcohol environment in neighborhood areas. *Alcoholism: Clinical and Experimental Research*, 27, 477-84.
- Forster JL, McGovern PG, Wagenaar AC, Wolfson M, Perry CL, Anstine PS. (1994). The ability of young people to purchase alcohol without age identification in northeastern Minnesota, USA. *Addiction* 89: 699-705.
- Fournier S. (2012). Mystery shop performance data from national retail chains. Presented at 10th RRFForum national conference, April 12, 2012, Boston, MA.
- Grant JD, Scherrer JF, Lynskey MT, Lyons MJ, Eisen SA, Tsuang MT, True WR, Bucholz KK. (2006). Adolescent alcohol use is a risk factor for adult alcohol and drug dependence: evidence from a twin design. *Psychol Med*. Jan; 36(1):109-18. Epub 2005 Sep 29.
- Grube J W. (1997). Preventing sales of alcohol to minors: Results from a community trial. *Addiction*, 92 (suppl. 2), S251-S260.
- Harrison PA, Fulkerson JA, Park E. (2000). Relative importance of social versus commercial sources in youth access to tobacco, alcohol, and other drugs. *Preventive Medicine*, 31, 39-48.
- Institute of Medicine and National Research Council. (2003). *Reducing Underage Drinking: A Collective Responsibility*. Committee on Developing a Strategy to Reduce and Prevent Underage Drinking, Richard J. Bonnie and Mary Ellen O'Connell, Editors. Board on Children,

- Youth and Families, Division of Behavioral and Social Sciences and Education. Washington, D.C.: The National Academies Press.
- Krevor BS, Ponicki WR, Grube JW, DeJong W. The Impact of Mystery Shopper Reports on Age Verification for Tobacco Purchases. *Journal of Health Communication*, 16(8): 820-830, September 2011.
- Krevor BS, Lieberman A, Gerlach K. (2002). The application of consumer protection authority in preventing tobacco sales to minors. *Tobacco Control*, 11,109-111.
- Krevor BS, Capitman JA, Cannon J, Oblak L, Ruwe M (2003). Preventing Illegal Tobacco and Alcohol Sales to Minors Through Electronic Age Verification Devices: A Field Effectiveness Study. *J Public Health Policy* 24, 251-69.
- Moore, R.S., Roberts, J., McGaffigan, R., Calac, D., Grube, J.W., Gilder, D.A., & Ehlers, C.L. (In Press). Implementing a reward and reminder underage drinking prevention program in convenience stores near Southern California American Indian reservations. *American Journal of Drug and Alcohol Abuse*.
- National Highway Traffic Safety Administration (2002). Youth fatal crash and alcohol facts 2000. DOT HS 809 406. Washington, D.C. U.S. Department of Transportation.
- Paschall MJ, Grube JW, Black C, Flewelling RL, Ringwalt CL, Biglan A. (2007). Alcohol outlet characteristics and alcohol sales to youth: results of alcohol purchase surveys in 45 Oregon communities, *Prevention Science*, 8, 153-159.
- Preusser DF, Williams AF. (1992). Sales of alcohol to underage purchasers in three New York counties and Washington, D.C. *J Public Health Policy* 13:306-17.
- Raudenbush S, Bryk A, Cheong YF, Congdon R. (2004). HLM 6: Hierarchical Linear and Nonlinear Modeling. Lincolnwood, IL: Scientific Software International.
- Responsible Retailing Forum. (2009). Recommended Practices for Off-Premise Alcohol Retailers. Available at <http://www.rforum.org/reports>.
- Responsible Retailing Forum. (2008). Report on the Responsible Retailing Systems Project. Available at <http://www.rforum.org/reports>.
- Spear L. (2002). Adolescent brain and the college drinker: Biological basis of propensity to use and misuse alcohol. *Journal of Studies on Alcohol (Suppl. 14):71–81*.
- Strunin L, Hingson R. (1992). Alcohol, drugs, and adolescent sexual behavior. *International Journal of the Addiction*. 27(2), 129-146.
- Tapert SF, Aarons GA, Sedlar G, Brown SA. (2001). Adolescent substance use and sexual risk taking behavior. *Journal of Adolescent Health*, 28, 181-189.
- Toomey TL, Komro KA, Oakes JM, Lenk KM. (2008). Propensity for illegal alcohol sales to underage youth in Chicago. *Journal of Community Health*, 33,134-138.
- Trayte R. The Case for Responsible Retailing (2005). Downloadable at http://fcpr.fsu.edu/retail/RR%20Forum%20Slides/2006_RRF_presentations/The_Case_for_Responsible_Retailing_2006.pdf
- U.S. Department of Health and Human Services. (2007). The Surgeon General's Call to Action to Prevent and Reduce Underage Drinking. US Department of Health and Human Services, Office of the Surgeon General.
- Wagenaar AC, Toomey TL, Erickson DJ. (2005). Preventing youth access to alcohol: outcomes from a multi-community time-series trial. *Addiction*, 100, 335-345.
- Wagenaar AC, Wolfson M. (1995). Deterring sales and provision of alcohol to minors: A study of enforcement in 295 counties in four states. *Public Health Reports*, 110, 419-427.

- Wagenaar AC, Murray DM, Gehan JP, Wolfson M, Forster JL, Toomey TL, Perry CL, Jones-Webb R. (2000). Communities mobilizing for change on alcohol: Outcomes from a randomized community trial. *Journal of Studies on Alcohol*, 61, 85-94.
- Wechsler H, Moeykens B, Davenport A, Castillo S, Hansen J. (1995). The adverse impact of heavy episodic drinkers on other college students. *Journal of Studies on Alcohol* 56(6):628-634.
- Wolfson M, Toomey TL, Forster JL, Murray DM, Wagenaar AC. (1996). Alcohol outlet policies and practices concerning sales to underage people. *Addiction* 91: 670-74.
- Yi H, Williams GD, Dufour MC. (2001). Surveillance report #56: Trends in alcohol-related fatal traffic crashes, United States, 1977-99. Bethesda, MD: National Institute on Alcohol Abuse and Alcoholism.