The Influence of Misperceptions about Social Norms on Substance Use among School-age Adolescents

Amialchuk, Aliaksandr and Ajilore, Gbenga and Egan, Kevin

University of Toledo, Center for American Progress, University of Toledo

12 February 2019

The Influence of Misperceptions about Social Norms on Substance Use among School-age Adolescents

Author:	Aliaksandr Amialchuk
Co-author:	Olugbenga Ajilore (Gbenga.Ajilore@utoledo.edu)
Co-author:	Kevin Egan (Kevin.Egan@utoledo.edu)
Affiliation:	University of Toledo
Address:	2801 Bancroft St. Department of Economics Toledo, OH 43606-3390
Phone:	(419) 530-4754
E-mail:	aliaksandr.amialchuk@utoledo.com

Declarations of interest: none

JEL Classification: C23, I12, J13
Keywords: Substance use, Misperception, Social Norms, Adolescents, Add Health

1. Introduction

Despite a steady decline in the prevalence of illicit drugs other than marijuana among American adolescents in recent years this decline has stalled in 2017 and the prevalence of substance use remains alarmingly high. A 2017 survey reported that among American 12th-graders, 33.2% consumed alcohol, 9.7% smoked cigarettes and 22.9% used marijuana in the past 30 days [Johnston, Miech, O'Malley, et al. 2018]. Consumption of these drugs bears high social and economic costs due to more accidents, premature mortality, productivity, health, and behavioral issues [NRC and IOM 2004]. Alcohol-related auto crashes alone are the leading cause of death among adolescents between the ages of 17 and 20 and each year more than 400,000 students ages 18 to 24 have unprotected sex due to drinking with about a quarter of those students reporting being too intoxicated to know if they consented to having sex [DHHS 2013].

Most people initiate alcohol, smoking and marijuana use early, usually as teens [DHHS 2012; Jensen and Lleras-Muney 2012; D'Amico and McCarthy 2006], putting them at greater risk of developing dependence and abuse compared to those who initiate substance use during adulthood [Strashny 2014]. While public health policy is focused on prevention and reduction in substance use among adolescents, achieving this goal remains a considerable challenge [Clark and Loheac 2007; Johnston, et al. 2018]. Peer influence is an important driver of individual substance use and it may be responsible for the limited success of the public health interventions [Rees and Wallace 2014].

Peer effect, i.e., the direct influence of the behavior of peers on an individual's own behavior, has been identified as an important factor for alcohol consumption, smoking and marijuana use [Lundborg 2006; Clark and Loheac 2007; Nakajima 2007; Ajilore, Amialchuk and Egan 2016]. In addition, individual behavior has been found to be driven by something more

proximal – the individual's (mis)perception of their peers' behavior [Prentice and Miller 1993; Borsari and Carey 2003; Perkins 2014]. Intuitively, as long as group norms influence behavior and as long as "situations or circumstances perceived as real are real in their consequences" [Thomas and Thomas 1928], one can expect the influence of the norm to occur via individual's perception of it regardless of how accurate the perception is [Perkins 2014]. However, if they discover that their behavior is different from the normative behavior, they may experience discomfort and try to resolve the discrepancy by adjusting their behavior to be more in line with the norm [Prentice and Miller 1993; Borsari and Carey 2003; Cooter, et al. 2008]. In addition, if it is the (mis)perceptions, and not the actual peer norms *per se*, which is driving the individual's behavior, then the estimates of the traditional peer effects that rely on arguably exogenous measures of actual peers' norms can be viewed as biased versions of the effect of the more proximal determinants of behavior - the (mis)perceptions; and the significance of the peer effects should decrease once the (mis)perception is added as a regressor.

An important question is who are your peers? Close friends are the most proximal and influential peer group [Prentice and Miller 1993; Halliday and Kwak 2012]. Most of the previous research used "typical/average student" as a reference group which lacks psychological reality for the respondents since they have in mind a group of real people, their close friends [Prentice and Miller 1993], which are more predictive of individual behavior [Perkins 2014]. In this paper we use a large nationally representative sample of U.S middle and high school adolescents and include information on the actual peers (nominated friends) who were directly asked about their substance use. Our data is from the school social networks of friends in the National Longitudinal Survey of Adolescent to Adult Health (Add Health). From it, we estimate the effect

of misperception about the rates of friends' substance use on own use by adolescents in grades 7-12.

We utilize "any use" as the measure of perceived and actual substance use. This helps alleviate possible social desirability bias due to stigma and recall biases in the self-reports of substance use, which are less likely to be a concern when focusing on the extensive rather than the intensive margin of substance use [Johnson 2014]. For example, for the intensive margin, memory impairment that was found in heavy and frequent drug users was implicated in the inaccurate and inconsistent reporting of substance use in surveys. Also, given the illicit and socially undesirable nature of substance use behaviors, they tend to be stigmatized and the perceived stigma was found to be associated with underreporting of such behaviors [for the review of literature, see Johnson 2014].

For the extensive margin decision, we include *all three* of the high use substances by adolescents- alcohol, smoking, and marijuana-[Johnston, et al. 2018] which are also the most amenable to peer influence [Clark and Loheac 2007]. Despite the fact that substance use rises sharply during teen years and is heavily influenced by school peers, little research has utilized data on school-aged adolescents [Juvonen, Martino, Ellickson et al. 2007; D'Amico and McCarthy 2006] with the previous research almost exclusively focusing on college students [Borsari and Carey 2003; Perkins 2014].

The richness of information and the panel design of the Add Health data allow us to address endogeneity concerns and also account for relevant confounding factors by using lagged and fixed effects estimation controlling for school and neighborhood environment, family background characteristics, peer selection and actual peer norms. Most of the previous studies only included perception of peer substance use and did not control for the actual peer norm

[Perkins 2014]; however, both actual and perceived behaviors of peers are independently correlated with individual's behavior [Perkins 2014; Juvonen et al. 2007]. The previous empirical studies, which included both actual and perceived norms, were limited to small and non-representative samples focusing mainly on college students, and only contained correlational analysis. Our study attempts to identify the independent effects of actual and perceived norms in a large nationally representative sample.

Finally, our study is the first to consider economic equilibrium framework of conformity to social norms developed by Cooter et al. [2008].[1] Our empirical analysis, which is directly based on this framework, employs a continuous misperception measure that allows precise estimation of the effects for the full range of biases, including underestimation to overestimation of the norm.

We find that overestimation of friend's substance use-alcohol, cigarettes, or marijuana-significantly increases adolescent's own use approximately one year later; and the estimated effect is robust across specifications including individual-level fixed effects regression. The effect size is larger for males than for females (although not always statistically significantly). Uniquely, we also find the effect to be strong among those who underestimate the norm, which suggests a possible rebound/boomerang effect which we show has important policy implications.

2. Empirical method.

2.1. Measures

2.1.1. Data Source

The data for this study are drawn from the first two waves of the Add Health data set; an on-going nation-wide survey of adolescents in 132 schools who were in grades 7 to 12 in 1994. A nationally representative subset of the initial sample was interviewed in their homes in 1994

[1] Due to space constraints, we provide an online appendix that summarizes fully Cooter et al.'s rational choice cost-benefit framework.

(Wave 1, 20,745 respondents) with follow-up surveys in 1996 (Wave 2, about 15,000 respondents), in 2001-2002, in 2008, and in 2016-2018. A unique feature of Add Health is that the first two waves contain information on individuals' nominations of their close friends (up to five male and up to five female friends, best friends nominated first). Because almost all of these friends were in the same school, peer substance use measures were constructed from their direct responses.

2.1.2. Outcome Variables

We created three outcome indicator variables that measure the respondent's alcohol, smoking, and marijuana use in Wave 2.[2] The alcohol variable indicates having drank alcohol more than once a month and is based on the question "During the past 12 months, on how many days did you drink alcohol?" with possible responses 1) "every day or almost every day", 2) "3 to 5 days a week", 3) "1 or 2 days a week", 4) "2 or 3 days a month", 5) "once a month or less (3-12 times in the past 12 months)", 6) "1 or 2 days in the past 12 months" and 7) "never". The alcohol variable was coded as 1 if the respondent chose response 1, 2, 3 or 4, and zero otherwise.[3] The smoking indicator variable is based on the question "During the past 30 days, on how many days did you smoke cigarettes?" with possible responses ranging between 0 and 30 days; it was coded as 1 for the responses ranging between 1 and 30. The marijuana use indicator is based on the question "During the past 30 days, how many times have you used marijuana?" with possible responses ranging between 0 to 900 times; it was coded as 1 for any nonzero response.

[2] Computer-assisted personal interviewing of Add Health was used to ensure confidentiality of responses and reduce reporting bias.

[3] The substance use variables were chosen to correspond to the questions about perceived substance use of friends, which ask how many friends use a given substance at least once a month. We have conducted robustness check by including response category (5) for alcohol question and the results remained unchanged. We note that there remains a slight inconsistency between the alcohol questions asking about actual drinking (more than once a month) and perceived friends' drinking (at least once a month).

2.1.3. Misperception measures

We created three misperception variables (scores) for each substance in Wave 1 as a difference between the perceived and the actual proportions of friend users. In Wave 1, the respondents were asked: "Of your 3 best friends, how many drink alcohol at least once a month?", "Of your 3 best friends, how many smoke at least 1 cigarette a day?" and "Of your 3 best friends, how many use marijuana at least once a month?" The perceived proportion of friend users was computed by dividing the perceived number of friend users (out of three best friends) by number three or the actual number of friendship nominations if it was less than three, with the possible values of 0, 1/3, 1/2, 2/3 and 1. The actual proportion of friend users was determined from the direct responses of nominated friends and was computed as the average across non-missing responses of the three best male and three best female friends of the indicator variables for alcohol, smoking, and marijuana during Wave 1, resulting in the possible values of 0; 1/N, N=2,3,4,5,6; 2/N, N=3,4,5,6; 3/N, N=4,5,6; 4/N, N=5,6; 5/6, and 1.[4] The misperception score is thus equal to the difference between the fraction of friends perceived to use the substance and the actual fraction of friend-users, $Fraction_{perceived\ peer\ use} - Fraction_{actual\ peer\ use}$. The resulting misperception score ranges between -1 and 1, where negative values indicate under-estimation and positive numbers indicate over-estimation of the proportion of friend users, and the misperception score of zero indicates correct estimation (no perceptual bias). Compared to a binary indicator of over-estimation used in some studies, this misperception score has the advantage of capturing finer variation in misperception and allows for greater measurement precision.

2.1.4. Socio-demographic controls

[4] Following an anonymous referee's suggestion, we re-estimated the models after randomly assigning nomination position to friends who were not named best male or female friend. The results did not differ. The results of this robustness check are available from the authors upon request.

We include controls for actual substance use among peers (descriptive norm) by using the average of substance use among best friends. We also have other controls that were measured in wave 1: grade level indicators for grades 7-12 (grade 7 being the omitted category), age, gender, race (white, black, other race), Hispanic ethnicity, Picture Vocabulary Test (PVT) score, log of pretax family income, indicator of whether the adolescent lived with both biological parents, whether mother or father has a college degree, variable to indicate attendance of religious services once a month or more, indicator for the presence of older siblings; and separate indicators for whether alcohol, cigarettes, or illegal drugs are easily available at home. We also control for age of the adolescent when they first moved to their current location and whether the parents chose their residence because of the school district. These indicators help to account for endogeneity of school choice or residential location [Clark and Loheac 2007].

2.1.5. Sample

The sample for this study is drawn from wave 2 (1996) respondents (N=14,738). We link the data on the nominated friends to each respondent. The average number of nominated friends per individual was 2.54 and the data was only available for the nominated friends who are in the same school as the respondent, which is approximately 85% of all friendship nominations. After restricting to individuals who were 20 years old or younger in wave 2 (18 years old or younger in wave 1), who nominated at least one friend, had non-missing observations on all the variables and had non-missing values on wave 2 sample weight, our sample reduced to 4,557 respondents.[5] The sample used for fixed effects regressions is further restricted to non-missing observations on misperception variables in wave 2.

[5] Add Health utilizes a multistage clustered sample design with observations having unequal probability of selection and requires the use of sampling weights in order to make the estimates nationally representative (Chen 2014).

Moreover, note that we are using nominated friendships from both waves 1 and 2 to estimate the fixed effects model. The respondents were asked to nominate up to 5 male and up to 5 female friends in both waves. In this way our fixed effects model uses measures of perception and actual behavior of friends nominated in waves 1 and 2. Thus, our fixed effect measures refer to *current* peers and to the extent that peer groups change from wave 1 to wave 2, our individual-level fixed effects regression reflects both the change in old peer's behavior (and perception of it) and the change in peer group composition.[6] Table 1 has summary statistics for the sample and all variables, which we will discuss further in the results section.

2.2. *Empirical models*

We estimate a linear probability model for substance use by individual *i* in school *s* observed in wave *t*:

$$Y_{ist} = \alpha + \beta_1 M_{ist-1} + \beta_2 F_{ist-1} + \beta_3 X_{ist-1} + \gamma_s + \varepsilon_{ist} \qquad \text{(Equation 1)},$$

where Y_{ist} is the substance use indicator from Wave 2, M_{ist-1} is the substance use misperception score measured in Wave 1, X_{ist-1} is the vector of individual demographic and family characteristics measured in Wave 1, and ε_{ist} is the error term. We control for unobserved environmental influences by including school fixed-effects, γ_s. For example, adolescents who are in neighborhoods with high density of alcohol and tobacco outlets may have both elevated misperception of peer drinking rates and higher level of own use of those substances. We also control for the average actual substance use among close friends, F_{ist-1}, in order to allow it to have an independent effect on individual substance use, and because perceived and actual peer substance use may be correlated.

[6] In the sample, 53% of the respondents changed at least one friend and 19% of the respondents changed all of their friends between the waves.

The strategy of using lagged values for the misperception score helps to identify the effect on current behavior if the substance use behavior of adolescents is likely to be out of steady state and not highly correlated over time [Clark and Loheac 2007; Manski 2000]. Substance use behavior during adolescence is likely to be out of steady state due to changing preferences and surroundings, which is reflected in the biggest number of initiations and highest growth in consumption during teen years [DHHS 2012; Jensen and Lleras-Muney 2012; Strashny 2014; D'Amico and McCarthy 2006]. This is supported by the relatively low estimates of the year-on-year correlation coefficients between the first two waves in our sample: 0.53 for smoking, 0.40 for drinking, and 0.42 for marijuana use.[7]

The theoretical model outlined in the online appendix implies that the effect of the misperception score on individual substance use, β_1, is positive regardless of the sign of misperception: misperception score is positive in case of overestimation of the norm, and higher score will lead to more substance use; the misperception score is negative in case of underestimation of the norm, and a lower score (indicating more misperception of the norm) will lead to lower substance use (father away from the norm). We also estimate individual-level fixed effects regression in first differences between waves 1 and 2:

$$\Delta Y_{is} = \beta_1 \Delta M_{ist} + \beta_2 \Delta F_{ist} + \beta_3 t + \varepsilon_{ist} \qquad \text{(Equation 2).}$$

In contrast with Equation 1, this regression controls for all of the individual-level differences that are related to substance use and its misperception, but don't change over time. It also exploits a different source of variation: instead of looking at how past misperception

[7] The identification of the observed friends' substance use (peer effect) was accomplished in a similar way to the identification of misperception and follows Clark and Loheac (2007) - by relying on lagged peer measure (to overcome the reflection problem in the standard social interactions equation); and by using individual-level covariates (including indicators for whether the families concerned are recent movers, and whether the adolescent's parents chose the neighborhood for the school), school fixed effects and individual fixed effects in order to control for the contextual and correlated effects (Manski 2000).

influences present behavior, it looks at how change in misperception over time affects the simultaneous change in individual behavior. If all of the individual-level factors correlated with misperception stay constant, and in the absence of the effect of individual behavior on his/her misperception, this strategy will identify the causal effect of normative misperception on individual substance use. It should also be noted that the fixed effects model better controls for the time-invariant individual heterogeneity, which, depending on the correlation with individual-level unobserved confounding factors, could otherwise bias the estimates up or down. For example, if the respondent and her friends are all members of a tight-knit network of regular marijuana smokers, their substance use would be higher while misperceptions about friends' use would be lower, resulting in a downward bias in the lagged model that fails to adjust for interconnectedness of the individual friendship network. In addition, the coefficients in the fixed effects model may be upward biased due to reverse causality whereby individuals who use substances more have elevated misperception of friends' use. Finally, the fact that the lagged model relies on friends from wave 1, while the fixed effects model relies on current friends may also drive differences in the coefficients. The lagged model coefficients are likely to be higher in magnitude because there is more room for the individual's behavior to change (at least one year, between wave 1 and wave 2) in response to misperceptions formed in wave 1, while there is less room for the behavior to adjust when it is measured in the same wave as the misperception.

3. Results

3.1. *Main results*

According to Table 1, an average adolescent in Wave 1 of our sample overestimated the proportion of friends who drink alcohol at least once a month by 23.4%, overestimated the proportion of friends who smoke at least 1 cigarette a day by 3.7%, and overestimated the

proportion of friends who use marijuana at least once a month by 8.2%. Between Waves 1 and 2, the prevalence of drinking at least once a month increased from 15.3% to 18.1%, the prevalence of smoking increased from 24% to 31.9% and the prevalence of marijuana use increased from 12.3% to 14.6% (statistically significantly for all three substances). The actual average substance use among close friends nominated by the respondent is slightly higher than the respondents' own use for each of the substances in wave 1 (statistically significantly for all three substances), possibly because the respondents tend to nominate popular friends who are also heavier users than the average adolescent. Table 2 shows that, compared to females, males have a higher prevalence of use of each substance.

Tables 3, 4 and 5 contain results of estimating equations 1 (lagged model) and 2 (fixed effects model) for the full sample and after stratifying by gender. [8] The estimated effect of misperception is positive and statistically significant in the lagged model, indicating that 0.1 unit increase in the misperception score leads to a 2.68 percentage point increase in the probability of the respondent's drinking. In the fixed effects model, the magnitude of the effect of drinking misperception is lower, suggesting that a 0.1 unit increase in the misperception score leads to a 1.39 percentage point increase in the probability of the respondent's drinking. Interestingly, the effect of misperception about friends' drinking is smaller in magnitude than the effect of the actual prevalence of friends' drinking, as indicated by the coefficient on average friends' drinking (descriptive norm). Panels B and C suggest that the effect of drinking misperception is significant for both genders, and the magnitude of the effect is higher (although not statistically significantly) for males.

[8] We have also estimated other lagged model specifications, gradually adding covariates by starting with only misperception score and average friends' substance use, then adding demographic and socio-economic controls, and then adding school fixed effects. These specifications produced similar coefficients on the misperception score for all three substances (results are available upon request).

The pattern of the results for smoking (Table 4) is similar to that for drinking (Table 3), except the magnitude of the effect of misperception is higher for smoking. In the lagged model, a 0.1 unit increase in the misperception score leads to a 3.13 percentage point increase in the probability of the respondent's smoking. The actual proportion of friends who smoke has a significantly greater effect on the likelihood of respondent's smoking than perceived friends' smoking in the lagged model.

Table 5 suggests that the effect of misperception on marijuana use is approximately the same in magnitude as in the case of smoking. In the lagged model, a 0.1 unit increase in the misperception score leads to a 3.16 percentage point increase in the probability of the respondent's use of marijuana. Just like in the case of alcohol and smoking, the actual descriptive norm has more influence than misperception of the norm. In addition, the size of the effects is larger for males.[9,10]

3.2. Estimates in the case of overestimation and underestimation of the norm

In order to investigate whether the effect of misperception differs between those who overestimate and underestimate the group norm, we re-estimated the models for those with a misperception score greater than or equal to 0 (overestimation and correct estimation) and for those with a misperception score less than or equal to 0 (underestimation and correct estimation) in wave 1. Table 6 shows that the effect of misperception in the overestimation range of scores is positive in all cases and are higher among males compared to females. The effects of

[9] In the case of the fixed effects model, the coefficient on actual friends' substance use was statistically significantly different from the coefficient on misperception only in the case of marijuana.

[10] Because perceptions are formed through social learning and learning may be more important at younger ages, we re-estimated the models separately for the younger (12-15) and older (16-18) respondents following an anonymous referee's suggestion. Only the fixed effects model showed a consistently stronger effect of misperception at the younger ages. The results are available upon request.

misperception are strongest for marijuana use, followed by smoking and then drinking. Table 6 also shows that among those who underestimated friends' substance use, the effect of misperception is positive and significant in all cases and generally larger than for those who overestimated substance use.

4. Discussion

We first referenced how perceptual bias influences substance use in an economic cost-benefit framework and then used Add Health data to test the prediction that higher overestimation of the group norm leads to higher individual's substance use. We used information from social networks of middle and high school adolescents to estimate the effect of misperception about the rates of friends' substance use on their own use of the three most prevalent substances among adolescents: alcohol, smoking and marijuana. This analysis helps to bring social comparison and psychological biases into the realm of economic analysis and adds to the scarce research on school-aged adolescents' perceptual biases and their influence on substance use behavior during the period when most initiation is taking place. Unlike any of the existing studies, we use a large and nationally representative sample of U.S middle and high school adolescents and include information on peers who were directly asked about their substance use. We also focus on the most proximal and salient peer group – nominated friends from the same school. The richness and panel design of Add Health allow us to overcome several endogeneity problems. We are able to use a time lag as well as control for school fixed effects and individual fixed effects which may remove much of the omitted variable bias. We conduct our analysis by gender and also separately for those who underestimated and overestimated the group norm.

This study contributes several new findings about the influence of substance use misperceptions concerning peer norms. Consistent with the prediction of the cost-benefit framework, we find that for each of the three substances, alcohol, smoking and marijuana, there is strong evidence that higher misperception of the group norm (relative to the actual group norm) significantly increases adolescent's own use. Furthermore, the effect of misperception of the norm is comparable, but slightly smaller than (~75-90% of) the magnitude of the effect of actual peer behavior. The significance and magnitude of the effect of misperception is robust across specifications including school-level fixed effects regression. Fixed effects regressions, which hold all individual-level factors constant, also suggest a significant positive effect with the magnitudes around 50-65% of the magnitudes in the lagged models, with the smaller effect sizes possibly reflecting better control for reverse causality or the fact that individual behavior did not change much in response to the misperception because both were measured in the same wave. The coefficient in the lagged models does not significantly change in magnitude after the additional covariates or school fixed effects are included (these results are not shown but available upon request). This suggests that a shared school environment does not explain much of the correlation between normative misperception and substance use. The effect of misperception is the strongest for marijuana use (illegal in all states at the time), followed by cigarette smoking and then alcohol consumption. One possible explanation for this finding is an incomplete information failure, where overall smoking and marijuana use is lower than alcohol leading to individuals to have less general knowledge of use and relying more on peer group information. Misperception of the group norm affects substance use more for males than for females in the case of marijuana. This is consistent with the previous finding that females' substance use is less influenced by group norms than males' [Clark and Loheac 2007; Cooter et

al. 2008], except now we find this to also be true for *perceived* group norms. In addition, a positive and surprisingly significant effect of higher perception on substance use among those who underestimate the group norm suggests the possibility of a boomerang (or rebound) effect.

Although several interventions have been found to decrease substance among school and college students, the rate of alcohol consumption and other drug use remains high in the U.S. Our findings suggest an opportunity to reduce substance initiation and subsequent use at the national level by conveying to middle and high school students the message about the true levels of substance use among their peers. Furthermore, our estimates suggest that interventions to change normative perceptions would be more effective when it comes to smoking and marijuana use and also when targeting males.

Our findings also suggest that policy makers should exercise caution and take into account the possibility of an undesirable boomerang (or rebound) effect, where students start using substances more after they realize that others use them more than they thought. For example, a school campaign targeting alcohol consumption might motivate students who previously consumed less alcohol than the norm to consume more now. This is because descriptive norms only provide a standard from which people do not want to deviate, and individuals are likely to strive to adhere to this standard regardless of whether they are currently above or below the norm. For example, given our estimate that a 0.1 unit increase in the misperception score increases the probability of smoking by 4.16 percentage points among those who underestimate and only by 2.47 percentage points among those who overestimate their friends' smoking, a policy that informs all students about the correct percentage of students who smoke would likely result in a higher smoking prevalence in school given that about the same

percentage of the students (about 20%) underestimate and overestimate peer smoking rate in wave 1.

The boomerang effect has been implicated for the lack of universal success of descriptive normative messages in changing behavior in multiple field contexts, especially where prevalence of underestimation of reference group's norm was more likely; and it was proposed that supplementing descriptive messages with injunctive messages could solve the problem [Schultz, Nolan, Cialdini et al. 2007]. Injunctive norms (also known as "prescriptive norms") represent perceived moral rules and reflect the individual's perception of what is commonly approved or disapproved by others. In fact, adding an injunctive message that the desired behavior is approved by others was found to be an important factor in substance use by itself [Pedersen, Osilla, Miles et al. 2017] and was shown to reduce the boomerang effect in the context of energy conservation [Schultz, Nolan, Cialdini et al. 2007]. An example of an injunctive message in the context of energy use is when a household received a happy face (☺) on his/her energy report after it had consumed less than the average amount of energy for the neighborhood, indicating the social desirability for the household to *remain better* than average and thus mitigate the boomerang effect [Schultz, Nolan, Cialdini et al. 2007]. Unfortunately, no information about other students' views or friends' approval of substance use is available in Add Health data.

Overall our analysis is extensive, investigating three different substances, all with various econometric specifications, and all variations supporting our conclusions. Thus we believe future research in this area is warranted, particularly determining whether accounting for injunctive norms reduces the influence of underestimating peer norms and ameliorates the boomerang effect in the context of substance use.

References:

Ajilore O, Amialchuk A, Egan K. (2016). Alcohol consumption by youth: Peers, parents, or prices? *Econ Hum Biol*. 23:76-83.

Borsari B, Carey KB. (2003). Descriptive and injunctive norms in college drinking: A meta-analytic integration. *Journal of Studies on Alcohol*. 64:331-341.

Chen P. 2014. Appropriate Analysis in Add Health Correcting for Design Effects & Selecting Weights. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill

Clark A, Loheac Y. (2007). It wasn't me, it was them! Social influence in risky behavior by adolescents, *Journal of Health Economics*, 26(4):763-784.

Cooter R.D., Feldman M., Feldman Y. (2008). The Misperception of Norms: The Psychology of Bias and the Economics of Equilibrium, *Review of Law & Economics*, 4(3):889-911.

D'Amico E.J., McCarthy D.M. (2006). Escalation and initiation of younger adolescents' substance use: The impact of perceived peer use. *Journal of Adolescent Health*, 39(4):481-487.

DHHS (2012). U.S. Department of Health and Human Services. Preventing Tobacco Use Among Youth and Young Adults: A Report of the Surgeon General. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2012 [accessed 2018 Jun 15]

DHHS (2013). US Department of Health and Human Services. Report to congress on the prevention and reduction of underage drinking. Retrieved from https://store.samhsa.gov/shin/content/PEP13-RTCUAD/PEP13-RTCUAD.pdf

Halliday T.J., Kwak S. (2012). What is a peer? The role of network definitions in estimation of endogenous peer effects. *Applied Economics* 44(3):289-302.

Jensen R., Lleras-Muney A. (2012). Does staying in school (and not working) prevent teen smoking and drinking?, *Journal of Health Economics*, 31(4):644-657.

Johnson T.P. (2014). Sources of Error in Substance Use Prevalence Surveys. *Int Sch Res Notices* 1-21.

Johnston, L. D., Miech, R. A., O'Malley, P. M., Bachman, J. G., Schulenberg, J. E., & Patrick, M. E.(2018). Monitoring the Future national survey results on drug use: 1975-2017: Overview, key findings on adolescent drug use. Ann Arbor: Institute for Social Research, The University of Michigan.

Juvonen J., Martino S., Ellickson P., Longshore D. (2007). But others do it! Do misperceptions of schoolmate alcohol and marijuana use predict subsequent drug use among young adolescents? *Journal of Applied Social Psychology*, 37(4):740-758.

Lundborg P. (2006). Having the wrong friends? Peer effects in adolescent substance use. *Journal of Health Economics*, 25(2):214-233.

Manski C.F. (2000). Economic analysis of social interactions. *Journal of Economic Perspectives* 14(3):115-136.

Nakajima R. (2007). Measuring Peer Effects on Youth Smoking Behaviour. *The Review of Economic Studies*, 74(3);897-935.

Committee on Developing a Strategy to Reduce and Prevent Underage Drinking; Bonnie RJ, O'Connell ME, editors. Reducing Underage Drinking: A Collective Responsibility. Washington (DC): National Academies Press (US); 2004. 2, Social, Health, and

Economic Consequences of Underage Drinking. Retrieved from: https://www.ncbi.nlm.nih.gov/books/NBK37611/

Pedersen E.R., Osilla K.C., Miles J.N.V., Tucker J.S., Ewing B.A., Shih R.A., D'Amico E.J. (2017). The Role of Perceived Injunctive Alcohol Norms in Adolescent Drinking Behavior, *Addictive Behaviors* 67:1-7.

Perkins H.W. (2014). Misperception is Reality: The Reign of Error About Peer Risk Behaviour Norms Among Youth and Young Adults. Xenitidou M. and B. Edmonds (Eds.) The Complexity of Social Norms, Computational Social Sciences. Switzerland: Springer International Publishing, pp. 11-36.

Prentice D., Miller D. (1993). Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, 64(2):243-256.

Rees C., Wallace D. (2014). The myth of conformity: Adolescents and abstention from unhealthy drinking behaviors. *Social Science & Medicine*, 108:34-45.

Schultz P.W., Nolan J.M., Cialdini R.B., Goldstein N.J., Griskevicius V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18:429-34.

Strashny A. (2014). Age of Substance Use Initiation among Treatment Admissions Aged 18 to 30. The CBHSQ Report: July 17, 2014. Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration, Rockville, MD

Thomas W., Thomas D. (1928). The child in America. New York, NY: Alfred Knopf.

Table 1: Descriptive statistics, N=4,557

Variable	Mean	SD	Min	Max
Drinking, W1	0.153	0.360	0	1
Smoking, W1	0.240	0.427	0	1
Marijuana use, W1	0.123	0.329	0	1
Drinking, W2	0.181*	0.385	0	1
Smoking, W2	0.319*	0.466	0	1
Marijuana use, W2	0.146*	0.353	0	1
Drinking misperception score, W1	0.234	0.468	-1	1
Smoking misperception score, W1	0.0370	0.452	-1	1
Marijuana use misperception score, W1	0.0820	0.389	-1	1
Average friends' drinking, W1	0.173	0.332	0	1
Average friends' smoking, W1	0.268	0.394	0	1
Average friends' marijuana use, W1	0.141	0.303	0	1
Under-estimated friends' drinking, W1	0.100	0.300	0	1
Correctly estimated friends' drinking, W1	0.494	0.500	0	1
Over-estimated friends' drinking, W1	0.407	0.491	0	1
Under-estimated friends' smoking, W1	0.206	0.404	0	1
Correctly estimated friends' smoking, W1	0.560	0.496	0	1
Over-estimated friends' smoking, W1	0.235	0.424	0	1
Under-estimated friends' marijuana use, W1	0.117	0.322	0	1
Correctly estimated friends' marijuana use, W1	0.659	0.474	0	1
Over-estimated friends' marijuana use, W1	0.223	0.417	0	1
Grade, W1	9.476	1.484	7	12
Age, W1	14.89	1.545	11	19
Male	0.478	0.500	0	1
Black	0.179	0.383	0	1
Other race	0.0120	0.107	0	1
Hispanic	0.160	0.367	0	1
PVT score	100.6	14.20	12	135
Ln(pretax family income)	3.616	0.883	-2.303	6.907
Both biological parents present	0.585	0.493	0	1
Mother or father has college degree	0.376	0.484	0	1
Religious	0.609	0.488	0	1
Older siblings present	0.514	0.500	0	1
Alcohol easily available at home	0.299	0.458	0	1
Cigarettes easily available at home	0.301	0.459	0	1
Illegal drugs easily available at home	0.0320	0.176	0	1
Years old when moved	0.429	0.495	0	1
Parents chose school	8.083	5.591	0	19

Note: * Statistically significantly different from W1 mean, $p<0.01$.

Table 2 Misperception and substance use by gender

	Males, N=2,180		Females, N=2,377	
Variable	Mean	SD	Mean	SD
Drinking misperception score, W1	0.237	0.475	0.231	0.461
Smoking misperception score, W1	0.0320	0.472	0.0410	0.433
Marijuana use misperception score, W1	0.0830	0.404	0.0810	0.374
Drinking, W2	0.219	0.414	0.146	0.353
Smoking, W2	0.334	0.472	0.305	0.461
Marijuana use, W2	0.166	0.372	0.127	0.334

Table 3: Regression estimates of the effect of misperception about friends' drinking

Panel A: Everyone	Lagged model		Fixed effects model	
Drinking misperception score, W1	0.268*** β	(0.021)	0.139***	(0.025)
Average friends' drinking, W1	0.376***	(0.040)	0.199***	(0.043)
Age, W1	0.019	(0.014)		
Male	0.049**	(0.021)		
Black	0.022	(0.024)		
Other race	-0.076	(0.052)		
Hispanic	0.056*	(0.030)		
PVT score	0.001	(0.001)		
Ln(pretax family income)	0.002	(0.009)		
Both biological parents present	-0.016	(0.017)		
Mother or father has college degree	-0.010	(0.024)		
Religious	-0.031*	(0.018)		
Older siblings present	0.025*	(0.014)		
Alcohol easily available at home	0.015	(0.019)		
Cigarettes easily available at home	0.047**	(0.019)		
Illegal drugs easily available at home	0.052	(0.041)		
Parents chose school	0.007	(0.017)		
Years old when moved	-0.001	(0.001)		
School fixed effects	Included			
Individual fixed effects			Included	
Wave 2 dummy			0.035***	(0.011)
Observations	4557		5254	
R-squared	0.225		0.043	

Panel B: Males	Lagged model		Fixed effects model	
Drinking misperception score, W1	0.300***	(0.041)	0.165***	(0.037)
Average friends' drinking, W1	0.384***	(0.069)	0.208***	(0.057)

Observations	2180		2536	
R-squared	0.316		0.049	

Panel C: Females	Lagged model		Fixed effects model	
Drinking misperception score, W1	0.236*** $^{\beta}$	(0.027)	0.108***	(0.034)
Average friends' drinking, W1	0.361***	(0.038)	0.191***	(0.065)
Observations	2377		2718	
R-squared	0.234		0.038	

Note: Statistical significance * p<0.1, ** p<0.05, *** p<0.01. "β" indicates statistically significant difference (p<0.05) in coefficients between the misperception score and average friends' behavior. "γ" indicates a statistically significant difference (p<0.05) in coefficients on misperception score between males and females. The dependent variable is the binary indicator variable for having drank alcohol at least once a month in the previous 12 months. All models are estimated using linear regressions. All estimates are weighted using Add Health W2 longitudinal survey weights. Models stratified by gender in panels B and C include all of the covariates from the models in panel A.

Table 4: Regression estimates of the effect of misperception about friends' smoking

Panel A: Everyone	Lagged model		Fixed effects model	
Smoking misperception score, W1	0.313*** $^{\beta}$	(0.025)	0.222***	(0.037)
Average friends' smoking, W1	0.450***	(0.028)	0.265***	(0.049)
Age, W1	-0.003	(0.016)		
Male	-0.010	(0.022)		
Black	-0.109***	(0.033)		
Other race	0.154*	(0.086)		
Hispanic	0.099***	(0.036)		
PVT score	-0.000	(0.001)		
Ln(pretax family income)	0.005	(0.011)		
Both biological parents present	-0.033	(0.027)		
Mother or father has college degree	-0.022	(0.021)		
Religious	-0.015	(0.020)		
Older siblings present	0.057***	(0.018)		
Alcohol easily available at home	-0.013	(0.026)		
Cigarettes easily available at home	0.069***	(0.021)		
Illegal drugs easily available at home	-0.044	(0.044)		
Parents chose school	-0.050**	(0.023)		
Years old when moved	0.003	(0.002)		
School fixed effects	Included			
Individual fixed effects			Included	

Wave 2 dummy			0.079***	(0.013)
Observations	4557		5250	
R-squared	0.234		0.084	

Panel B: Males	Lagged model		Fixed effects model	
Smoking misperception score, W1	0.345***	(0.035)	0.216***	(0.048)
Average friends' smoking, W1	0.396***	(0.047)	0.247***	(0.066)
Observations	2180		2536	
R-squared	0.277		0.089	

Panel C: Females	Lagged model		Fixed effects model	
Smoking misperception score, W1	0.285*** β	(0.031)	0.230***	(0.057)
Average friends' smoking, W1	0.495***	(0.035)	0.289***	(0.072)
Observations	2377		2714	
R-squared	0.304		0.081	

See note to table 3.

Table 5: Regression estimates of the effect of misperception about friends' marijuana use

Panel A: Everyone	Lagged model		Fixed effects model	
Marijuana use misperception score, W1	0.316*** β	(0.024)	0.219*** β	(0.036)
Average friends' marijuana use, W1	0.394***	(0.030)	0.312***	(0.054)
Age, W1	-0.006	(0.011)		
Male	0.015	(0.015)		
Black	0.003	(0.027)		
Other race	-0.033	(0.078)		
Hispanic	0.070**	(0.029)		
PVT score	0.001**	(0.001)		
Ln(pretax family income)	-0.007	(0.012)		
Both biological parents present	-0.043***	(0.014)		
Mother or father has college degree	0.011	(0.016)		
Religious	-0.046***	(0.015)		
Older siblings present	0.027**	(0.013)		
Alcohol easily available at home	0.004	(0.019)		
Cigarettes easily available at home	0.035**	(0.016)		
Illegal drugs easily available at home	0.089*	(0.047)		
Parents chose school	-0.030**	(0.013)		
Years old when moved	-0.001	(0.001)		
School fixed effects	Included			
Individual fixed effects			Included	
Wave 2 dummy			0.007	(0.010)
Observations	4557		5236	
R-squared	0.225		0.077	

Panel B: Males	Lagged model		Fixed effects model	
Marijuana use misperception score, W1	0.366***	(0.035)	0.282***	(0.053)
Average friends' marijuana use, W1	0.388***	(0.047)	0.396***	(0.082)
Observations	2180		2520	
R-squared	0.316		0.115	

Panel C: Females	Lagged model		Fixed effects model	
Marijuana use misperception score, W1	0.252*** β γ	(0.038)	0.150*** γ	(0.047)
Average friends' marijuana use, W1	0.380***	(0.041)	0.213***	(0.068)
Observations	2377		2716	
R-squared	0.259		0.044	

See note to table 3.

Table 6: Regression estimates stratified by level of misperception in W1: underestimation or overestimation

Drinking, underestimation	Lagged model		Fixed effects model	
Misperception score, W1 (everyone)	0.291***	(0.048)	0.157***	(0.031)
Misperception score, W1 (males)	0.330***	(0.073)	0.182***	(0.046)
Misperception score, W1 (females)	0.189**	(0.084)	0.133***	(0.041)
Drinking, overestimation				
Misperception score, W1 (everyone)	0.247***	(0.023)	0.119***	(0.026)
Misperception score, W1 (males)	0.274***	(0.040)	0.152***	(0.038)
Misperception score, W1 (females)	0.236***	(0.029)	0.080**	(0.033)

Smoking, underestimation	Lagged model		Fixed effects model	
Misperception score, W1 (everyone)	0.416***	(0.047)	0.258***	(0.046)
Misperception score, W1 (males)	0.497***	(0.048)	0.224***	(0.060)
Misperception score, W1 (females)	0.311***	(0.076)	0.297***	(0.071)
Smoking, overestimation				
Misperception score, W1 (everyone)	0.247***	(0.030)	0.210***	(0.041)
Misperception score, W1 (males)	0.236***	(0.046)	0.240***	(0.054)
Misperception score, W1 (females)	0.256***	(0.039)	0.168***	(0.064)

Marijuana, underestimation	Lagged model		Fixed effects model	
Misperception score, W1 (everyone)	0.378***	(0.061)	0.206***	(0.043)
Misperception score, W1 (males)	0.457***	(0.076)	0.242***	(0.069)
Misperception score, W1 (females)	0.287***	(0.084)	0.173***	(0.051)
Marijuana, overestimation				
Misperception score, W1 (everyone)	0.302***	(0.030)	0.202***	(0.038)
Misperception score, W1 (males)	0.345***	(0.049)	0.248***	(0.055)
Misperception score, W1 (females)	0.233***	(0.041)	0.148***	(0.052)

Note: Statistical significance * p<0.1, ** p<0.05, *** p<0.01. The underestimation sub-sample includes respondents who either underestimation or correctly estimated friend's substance use. The overestimation sub-sample includes respondents who either

overestimation or correctly estimated friend's substance use. See Table 3 for the list of covariates included in the models. All estimates are weighted using Add Health W2 longitudinal survey weights.