Published in final edited form as:

J Quant Criminol. 2017 March ; 33(1): 131–155. doi:10.1007/s10940-016-9286-9.

A Factor Analytic Model of Drug-Related Behavior in Adolescence and Its Impact on Arrests at Multiple Stages of the Life Course

Matthew D. Phillips¹

¹Department of Criminal Justice and Criminology, UNC Charlotte, Charlotte, NC, USA

Abstract

Objectives—Recognizing the inherent variability of drug-related behaviors, this study develops an empirically-driven and holistic model of drug-related behavior during adolescence using factor analysis to simultaneously model multiple drug behaviors.

Methods—The factor analytic model uncovers latent dimensions of drug-related behaviors, rather than patterns of individuals. These latent dimensions are treated as empirical typologies which are then used to predict an individual's number of arrests accrued at multiple phases of the life course. The data are robust enough to simultaneously capture drug behavior measures typically considered in isolation in the literature, and to allow for behavior to change and evolve over the period of adolescence.

Results—Results show that factor analysis is capable of developing highly descriptive patterns of drug offending, and that these patterns have great utility in predicting arrests. Results further demonstrate that while drug behavior patterns are predictive of arrests at the end of adolescence for both males and females, the impacts on arrests are longer lasting for females.

Conclusions—The various facets of drug behaviors have been a long-time concern of criminological research. However, the ability to model multiple behaviors simultaneously is often constrained by data that do not measure the constructs fully. Factor analysis is shown to be a useful technique for modeling adolescent drug involvement patterns in a way that accounts for the multitude and variability of possible behaviors, and in predicting future negative life outcomes, such as arrests.

Keywords

Drug use; Drug selling; Factor analysis; Arrests

Introduction

The relationship between drug-related behaviors and arrest is both widely researched and well documented (Bennett et al. 2008; Kinlock et al. 2003). This is with good reason, as either can be quite damaging to an individual. Drug behaviors carry risks of violence, and other criminal involvement (DeLisi et al. 2013; Korcha et al. 2014). An arrest can have a

profound impact on a person's life, either indirectly such as imparting negative stigma, or in more direct manners such as making it more difficult to get a job (Brame et al. 2012). Drug-involved individuals risk such disruptions by exposing themselves to arrest, and the literature is remarkably consistent in agreeing that the risk is profound (Dave 2008; DeLisi et al. 2015; Warner and Coomer 2003). Considering the number of crimes committed and the number of persons arrested, the overall risk of arrest is low. However, when compared to other crimes, the risk of arrest for drug behaviors is quite high (Langan and Levin 2002; Mosher 2001). Drug-involved individuals then disproportionately expose themselves to high risk of arrest for drug behaviors. It is less understood, however, the extent to which drug involvement may carry risk for arrests that are not related to drugs offenses.

The available literature on drug-related behaviors and experiences with the criminal justice system make it clear that many drug behaviors or offenses are related to various criminal sanctions, particularly arrest (Bennett et al. 2008; Dorsey and Middleton 2010; Rhodes et al. 2007). What is not clear is how the patterns of drug involvement relate to an individual's arrest history. That is, the literature at present cannot describe how individual variations in drug behaviors affect that individual's accrual of arrests through the life course. Use of a particular substance or habitual drug selling may carry an increased risk of arrest, but what of the totality of behaviors? The various drug-related behaviors in question are not mutually exclusive, nor do they happen in isolation. Understanding how an individual's complete drug behavior history influences experiences with the criminal justice system is therefore a critical goal.

An extremely common obstacle extant in the literature is that data adequate to assess the totality of behaviors are rare. Drug-related behavior is a rare phenomenon (relative to other forms of crime such as larceny), and obtaining substantial variation to permit statistical analyses is a challenge. To overcome this challenge, many data sets draw on what are referred to as "captive samples" (see Nurco et al. 2008). Captive samples refer to samples of individuals held in custody, such as arrestees or jail or prison inmates. Inmates are then asked retrospective questions regarding their drug offending (as well as other forms of offending). When looking at drug-related behavior and arrest or incarceration, this creates a statistical problem known as endogenous stratification—the samples used in estimation are selected conditional on the value of the endogenous or dependent variable. Methods have been devised to correct for endogenous stratification (see Rhodes et al. 2007), but ideally one would want a prospective sample that is not selected conditional on having an arrest or being incarcerated (and that has substantial variation on the drug offending measures).

A prospectively collected data set appropriate to leverage in studying these phenomena would capture information on individuals with multiple measures of drug use and drug selling (e.g. substances used and sold, the frequency of use and sales, etc.) and would capture every instance of arrest for each individual. Such a data set would include a robust battery of drug offending measures to allow for an exploration of how differing patterns of drug-related behavior affects subsequent arrest histories. The current study uses a data set providing exactly this type of information to better understand the totality of drug-related behaviors during the period of adolescence, considering a host of drug measures, including

drug use, drug selling, the substances involved in both, the age of initiation into such behaviors, and others.

The goals of this study are twofold. First, we intend to use factor analysis to develop a unique pattern or typology of drug-related behaviors for each individual in the data set. Traditional factor analysis searches for clusters of variables to create groups of respondents. We move beyond this paradigm to instead create groups of (drug offending) behaviors, which are used to empirically describe each respondent in our data set according to their drug-related behavior history as it evolved over the period of adolescence. The literature demonstrates that remarkable heterogeneity exists among drug-involved persons, and developing latent dimensions of drug-related behaviors through factor analysis allows us to exploit that heterogeneity. By first identifying how many dimensions of drug-related behaviors in adolescence exist in our sample, and then by calculating a score for each subject on each dimension (or factor), we create a unique description of drug-related behavior during adolescence for each subject. These combinations of scores for each subject preserve the heterogeneity in the sample.

Second, we use the factor scores for each subject on each dimension to predict the number of non-drug arrests at three phases in the life course; the end of adolescence, young adulthood, and adulthood. In these models, we control for prior arrests as well as current drug-related behaviors, so as to isolate the effect of adolescent drug behaviors on arrests at multiple points. Because the literature suggests disparities in drug behaviors between genders, and because the data considered herein are overrepresented by males, we segregate the analyses by male and female subjects. This separation allows us to examine how adolescent drug-related behavior affects males and females in different ways.

Results from this study demonstrate the utility of factor analysis to capture and describe the heterogeneity of drug-involved adolescents, and to better understand the impact of adolescent drug involvement on arrests later in life. The remainder of the paper is organized as follows. In the next section we review the relationship between drug-related behaviors, crime, and arrests, as well as previous efforts to create typologies of drug-involved individuals. We then discuss the data and methods used, prior to presenting the results of empirical analyses. In the final section we discuss the conclusions, limitations of our research, and directions for future work.

Drug-Related Behaviors, Arrests, and Typologies

The literature makes it clear that drug-related behaviors are related to antisocial behaviors, delinquency, and criminality, including violent offending (DeLisi et al. 2015). An individual's choice to engage in such offenses creates opportunities for that individual to experience criminal sanctions such as arrest. However, the precise mechanisms leading from drug-related behaviors to criminal offending are complex, and often depend on the form and frequency of drug behaviors.

Criminal offenders tend to be generalists, rather than specialists (Altschuler and Brounstein 1991), and thus individuals who are prone to engage in drug-related behaviors are apt to

engage in a variety of criminal offenses. Thus, there is an association between drug use and other forms of criminality. Most commonly, it is thought of for its relation to violent crime. For example, individual drug use is consistently associated with robbery and assault (Baumer et al. 1998; Bennett et al. 2008; Kinlock et al. 2003), and weapon-carrying (Lizotte et al. 2000), yielding the potential for severe violence and injury (Korcha et al. 2014). A few studies suggest drug use may be related to homicide offending (Goldstein 1985; Hagelstam and Hakkanen 2006; McLaughlin et al. 2000; Shaw et al. 2006), but the relationship is not as clear or consistent as that between drug use and less serious violent offenses. In addition to violent crime, property crime such as vandalism, burglary, and breaking and entering is also related to drug use (Altschuler and Brounstein 1991; Bennett et al. 2008).

Individual drug use shows a consistent relationship with many criminal behaviors, including violent and property crime. But so too does drug selling. Much as we see with drug use, drug selling shows a strong association to crime, particularly violent crime. Phillips (2012) demonstrated that drug selling was able to predict level of violence, or violence incidence, while drug use was only able to predict prevalence of violence. Lizotte et al. (2000) show that a high level of drug selling predicts carrying a gun. Violence and guns have reciprocal effects, especially in drug markets. Violence encourages the participants to turn to guns and gangs for protection, which in turn increases the levels of violence (Braga et al. 2008). Drug selling has been shown to be a remarkably common activity for gang members (Thornberry et al. 2003). Drug markets create reliably violent circumstances, a finding well established in the literature (Fleisher 2006; Goldstein 1985; Miller 2001; Valdez and Sifaneck 2006). Further, the violence can vary with the particular drugs being sold. Selling cocaine and crack cocaine, for example, is an especially violent enterprise (Bennett et al. 2008).

Altschuler and Brounstein (1991) found that juveniles who sold drugs were likely to be criminal “generalists,” committing both crimes against persons as well as property. The violent crimes in this study included carrying a weapon, using a weapon to threaten violence, robbery, and various forms of assault (including aggravated and sexual assault). The property crimes included vandalism, auto theft, burglary, breaking and entering, and buying or selling stolen goods. The conclusion that drug selling is related to the above list of crimes is quite consistent with conclusions presented in Goldstein (1985). Goldstein found that the crime and violence surrounding drug markets is most typically “systemic” (meaning that it is intrinsic to drug markets, and used as an instrument of market regulation) or “economic-compulsive” (meaning that it is used to obtain money for illegal substances).

Singularly, drug use and drug selling consistently show negative impacts on an individual’s life through involvement in violence and property crime. However, when an individual is involved with both of these behaviors the negative outcomes, most notably criminal activity, become much more pronounced. In the sample of adolescents studied by Altschuler and Brounstein (1991), subjects who used and sold drugs were the most likely to commit both crimes against persons (violent crimes) and against property, and at the greatest rate. Further, selling drugs is an often natural progression from using drugs for many individuals (Johnson et al. 1995). This not only compounds the probability of observing a kind of comorbidity in which an individual is engaged in both drug use and drug selling, and thus being subject to their combinative effect on criminal activity, it also highlights the importance of empirical

research allowing for both forms of drug-related behaviors to be measured. Thornberry et al. (2003) found that gang membership facilitated the increase of drug use and drug selling,¹ and that while drug use declined after leaving the gang, drug selling endured. Similarly, gang membership is associated with increases in serious violence, drug selling, and serious theft among gang-involved youth (Gordon et al. 2014). As with the combination of drug use and drug selling, it is therefore of critical importance that gang involvement be measured when studying the impact of drug-related behaviors.

The relationship between drug-related behaviors and delinquency and criminality seems to be especially pronounced during adolescence (Passini 2012; Brook et al. 2011; Monahan et al. 2014). Despite a well-documented association between drug-related behaviors and delinquency and criminal offending, the substantial heterogeneity among drug-involved adolescents (DeLisi et al. 2013, 2015; Brook et al. 2011) creates differential risks for exposure to arrest. Particularly chronic or heavily drug-involved adolescents engage in the highest levels of delinquent offending and are therefore most susceptible to criminal arrest. In studying the ways in which adolescent drug-related behaviors may influence arrests at multiple points in the life course, it is then necessary to measure not only the type but also the level of drug-related behaviors in which an adolescent engages. Having discussed how drug-related behaviors are associated with criminality generally, we now consider how drug-related behaviors are associated with arrests specifically.

In nearly any study using samples drawn from arrestee populations, drug offenses are near ubiquitous (Mosher 2001; Dave 2008), and an incredible proportion of all arrestees have reported drug use at some point in their life (Langan and Levin 2002; Lyman and Potter 2007). Mosher (2001) notes that drug offenses represented the largest category of arrests for the biggest cities in the United States as of the turn of the century. Almost a quarter of federal inmates report being motivated by drugs to commit the offense for which they are currently serving their sentence (Dorsey and Middleton 2010). The high proportion of drug users among arrestees exists not only in the United States, but in other cultures as well. According to Maden et al. (1992), 43 % of an arrested sample in the United Kingdom had used an illicit drug prior to arrest, with the most common substances being marijuana, heroin, and methamphetamine.

Drug selling is also an extremely common offense in arrested samples. Langan and Levin (2002) followed persons released from custody and found that 64.2 % were rear-rested for a drug trafficking offense, compared to 67.2 % being rearrested for a drug possession offense. In fact, the arrest rates for drug offending are so high that Warner and Coomer (2003) concluded that official drug arrest data is a valid indicator of the level of drug activity for small geographic locations such as neighborhoods. Rosenfeld and Decker (1999) echo this conclusion for city-level data, for cocaine and heroin use.

The substances involved are salient in determining the risk of arrest. Drug offenders unilaterally face a substantial hazard of being arrested, and the more addictive drugs, like

¹This finding speaks to a larger potential for criminality to increase an individual's embeddedness with a deviant subculture that normalizes drug use and drug selling. Certainly, for some individuals this may well be the case. As will be discussed in the methods section, this study is focused on the direction of association that posits drug-related behaviors influencing criminality.

cocaine, or more recently, methamphetamine, carry higher arrest risks (Cartier et al. 2006; Harrison and Gfoerer 1992; Rhodes et al. 2007). But the frequency of drug use is also highly relevant in understanding arrest rates. Harrison and Gfoerer (1992) analyzed data from the 1991 National Household Survey on Drug Abuse and find that arrest rates are highest for persons most heavily involved in drug use, as measured by frequency of use. The lowest arrest rates were found in individuals who did not use alcohol or drugs in the year prior to the survey. Even after controlling for demographic characteristics such as age, race, gender, marital status, income, and education, the drug use measures (frequency of using marijuana and cocaine in the past year) were significantly related to arrest for criminal behavior. This was true not only for drug arrests, but also for arrests for both violent and property crime, a finding echoed above. Rhodes et al. (2007) examined the drug use of arrestees and found that the arrest rate increases with the intensity of drug use.² Again, these arrests were not always for drug offenses. Compared to persons who do not use drugs (other than marijuana), low intensity drug users have arrest rates that are 18 % higher; medium intensity users have 40 % higher arrest rates, and high intensity users have 43 % higher rates. Chronic drug users (those who used cocaine, heroin, or methamphetamine at least four times in the month prior to arrest) had higher arrest rates than all other types of drug users (Rhodes et al. 2007).

Thus we see that the literature overwhelmingly supports the notion that drug-related behavior in its various forms is inextricably tied to elevated risk of arrest. Uniformly, conclusions support the idea that drug users and drug sellers are quite likely to experience arrest, and that the risk of arrest (and arrest rate) increases with both the frequency of offending and the type of the substances involved. However, the literature is limited in that it does not allow for simultaneous study of the myriad behaviors in the drug behavior collective. In other words, extant studies consider some small selection of drug behaviors rather than such behavior in its totality. The current study extends the literature by using a battery of drug-related behaviors to create drug behavior typologies that are unique to each individual subject, and uses those patterns to predict arrests at various stages of the life course. This novel application of factor analysis to empirically encapsulate adolescent's drug-related behavior by leveraging a robust data set capable of simultaneously measuring behaviors currently accounted for separately in the literature is the advancement presented herein.

Typologies of Drug-Involved Adolescents

To explore the patterns of drug use and drug selling during adolescence, typologies of drug-related behaviors are created using a robust longitudinal data set. Using typologies as heuristic and empirical constructs in drug offense research is quite common and particularly useful. It has long been recognized that drug offenders are not a homogenous group (see Nurco et al. 2008). The behaviors particular to individuals vary widely, and typologies therefore are useful devices to separate the heterogenous group of drug-involved adolescents into categories that more accurately group adolescents similarly to one another.

²Here, intensity is derived from frequency of use of crack cocaine, powder cocaine, heroin, and crystal methamphetamine, taken from data gathered in the Arrestee Drug Abuse Monitoring (ADAM) program.

There are several examples of empirical research using simple drug behavior typologies (Caulkins et al. 1998; French et al. 2001; Newcomb 1994). When both drug use and drug selling are considered together, these typically take the form of categories such as no drug involvement, drug use only, drug selling only, and both drug use and drug selling (Altschuler and Brounstein 1991; Black and Ricardo 1994; Ricardo 1994). Even when studies examine only drug use or only drug selling, typologies can still be useful. Newcomb (1994), for example, grouped drug users into marijuana users only, cocaine users only, and polydrug users. French et al. (2001) distinguished between chronic and non-chronic drug users in their analysis. Typologies of drug selling also exist. Caulkins et al. (1998) broke sellers into various types, according to the substances they sold. Fagan and Chin (1990) examined only sellers of cocaine, but grouped them according to whether they sold powdered cocaine, crack cocaine, or both, and whether they sold alone or in groups.

The categorization of drug offenders through typologies therefore is not a novel concept. Although numerous extant studies in the literature categorize drug offenders based on variables such as substance used, frequency of drug use, using drugs alone or in groups, substances sold, selling drugs in a group, and the like, most of these studies create categories of drug users separate from drug sellers. That is, studies tend to focus on only one family of drug behaviors. Often this is due to data constraints. Robust data sets including measures on a broad spectrum of drug behaviors are rare. Nonetheless, studying the patterns of drug use and drug selling that exist in such a data set would represent a substantial addition to the literature in that it could be leveraged to offer a more complete view of drug-related behavior during adolescence.

Generating a typology of drug-involved adolescents according to their patterns of use and selling is important not only because there are multiple types of drug-involved adolescents, but also because substantial differences are likely to exist between the various types (Brook et al. 2011; DeLisi et al. 2013; Kinlock et al. 2003). For example, the consequences of using marijuana once a month may differ greatly from those of using cocaine every day. Similarly, the effects of selling small amounts of drugs infrequently are likely not equivalent to the effects of heavy involvement in the retail drug markets. The current project compares types of users to other types of users (and sellers to sellers). It is unique in that it also compares drug users to drug sellers, and to persons who do both. More completely, it classifies adolescents according to where they fall along both the drug use and drug selling dimensions. Finally, it develops novel typologies based on a method not commonly used in the literature, although it is highly adept for this problem; factor analysis.

Methods

Data

This study uses data from the Rochester Youth Development Study (RYDS), an on-going longitudinal study of a panel of youth at high risk for violence and delinquency. The RYDS project began in 1988, and selected 1000 seventh and eighth graders in the Rochester (New York) Public School System to be interviewed, along with a parent or guardian. To date, RYDS has completed fourteen interviews for the panel, reaching subject ages in the early 30 s.

The original RYDS sample was stratified on two dimensions in order to select subjects who were at high risk for violence and delinquency. First, males were oversampled (75 vs. 25 %) as they are more likely than females to engage in serious and violent offenses³ (Blumstein et al. 1986; Huizinga et al. 1992). Second, students living in areas of the city with high residential arrest rates were also oversampled. This was based on the assumption that adolescents who live in such areas are at greater risk for offending than students living in areas with lower residential arrest rates. High residential arrest areas were identified by assigning each census tract in Rochester a resident arrest rate which reflected the proportion of the tract's total adult population arrested by the Rochester Police Department in 1986.

Data collection for the RYDS was completed in several phases. Phase 1 includes the first 9 waves of data collection, and collectively captures the period of adolescence. From Wave 1 to Wave 9, subjects (G2) were interviewed every 6 months from the spring of 1988 to the fall of 1996. In addition, one of their parents (G1) was also interviewed from Waves 1–8. RYDS contains information gathered from school records and social services records covering the entirety of Phase 1. By Wave 9, our subjects' average age was 18 years old. Following a 2-year gap in data collection, Phase 2 began in 1994. Phase 2, covering the stage of young adulthood, included annual interviews for three waves (Waves 10, 11, and 12). Subject ages at these waves averaged 21, 22, and 23, respectively. The final phase of data collection focusing on the G2 subjects, Phase 3, contains two waves (13 and 14), completed in 2003 and 2005. Phase 3 contains data during adulthood, when subjects are 29 and 31 years old, on average. At all waves of Phases 1, 2, and 3, data were gathered from police⁴ records for our G2 subjects.

A principal strength of the RYDS data is that it can bring to bear a host of drug-related behavior variables for each individual sampled.⁵ The drug behavior variables are included in the delinquency section of the RYDS survey. Collectively, these variables cover the range of possible drug use and drug selling activities. The measures ask the respondent about their behaviors since the date of the last interview. Since the interview waves were spaced approximately 6 months apart, the respondents were therefore reporting on their behavior in the previous 6 months. Unlike some other studies asking subjects to report on all prior behavior, the RYDS data are bounded temporally. Further, compared to the studies that do use multiple waves of data collection and temporal bounds, RYDS uses a shorter temporal bound than is often used. This makes the data less susceptible to the problems of retrospective data collection found in the literature (Roberts et al. 2005).

The subject panel is 68 % African-American, 17 % Hispanic, and 15 % White. These proportions are quite close to what was expected given the population characteristics of Rochester schools and the decision to oversample high-risk youth. Subject attrition in the

³Because of this oversampling, subsequent analyses are disaggregated by gender.

⁴Both from Rochester Police Department (RPD) and the New York State Department of Criminal Justice Services (DCJS). Including both of these agencies assures that events occurring outside of Rochester will be included. Unfortunately, but data outside New York State are not available.

⁵It should be noted that while the RYDS interviews ask each subject a battery of questions regarding drug use and selling, the prevalence of such behaviors in the sample is still low. The overwhelming majority of drug use, for example, is the use of marijuana. Nearly all of the rest is cocaine use. Nonetheless, previous drugs and crime research using RYDS data has proven quite fruitful, despite not having extreme variation on the drug offending measures.

RYDS is quite low when compared to other longitudinal studies. From Waves 2–12, we experienced only about 1 % attrition per year. At Wave 12, 85 % (846) of the original 1000 subjects were re-interviewed. Attrition increased slightly during Phase 3, which included roughly 785 subjects. A formal test of subject attrition within RYDS revealed that the subjects retained did not significantly differ from those not retained on multiple dimensions, including gender, social class, family structure, drug use, delinquency, property crime, and violent crime (see Krohn and Thornberry 1999).

Factor Analytic Typologies

As discussed in the literature review above, the creation of typologies of drug offenders is not a novel concept. However, the work reviewed above created typologies of drug offenders, with categories populated by individuals. The current work will take a different approach, instead creating typologies of drug-related behaviors. That is, the typologies created are classes of behaviors, not classes of actors. This idea will be elaborated upon below.

One key advancement made by the current study is the ability to create typologies of drug behavior that seeks to capture the totality of drug involvement during adolescence. Previous research has been limited in that data were available for only one or few dimensions of drug involvement (Caulkins et al. 1998; Fagan and Chin 1990; French et al. 2001; Newcomb 1994). The typologies that could be created based on those data thus were overly simplistic and constrained. Even when both drug use and drug selling are considered together, the typologies were rudimentary (as above, including use only, selling only, or both). Finally, because they were mostly based on cross-sectional data, the typologies cannot account for changes in behaviors over time. The complexity of the drug offending career remains hidden in the simplistic typologies that exist in the literature up to this point. The drug offending typologies created in the current project are nuanced, and include a more complete spectrum of drug behaviors. Because they are based on longitudinal data, they account for patterns of change for individuals.

A second key advancement in this research is a notable departure from traditional drug offending research employing typologies. The typologies created here are patterns of *drug-related behaviors* instead of *drug-involved adolescents*. The difference is subtle, but quite important. Traditional typologies group offenders into categories based on various behavioral traits and predilections. The typologies used here do not reflect groups of individuals. Rather, they capture the underlying or latent dimensions of drug behaviors that are present to greater or lesser extents in every individual. Under this conceptualization, a single respondent is not categorized into a single category of behavior. Instead, each respondent is evaluated for their level of involvement in each of the dimensions of drug-related behavior. These dimensions are not simple binary indicators of whether the respondent engaged in the behavior, but are scored on a continuum. For example, a person will not be categorized as a “marijuana user,” but would instead be measured on the degree to which he or she uses marijuana, in addition to being measured on all of the other latent dimensions. For the sake of convenience these underlying latent dimensions are referred to

as typologies. As stated in Rummel (1970, p. 21), “factors can be considered a typology, classifying or categorizing phenomena according to their inter-relationships.”

To create the typologies based on the drug use and drug selling data in RYDS, factor analysis will be used. Factor analysis is a mathematical procedure that searches for and uncovers systematic dependencies and correlations, synthesizing thousands of measurements and resolving them into distinct patterns of occurrence (Abdi 2003; Thompson 2004). Factor analysis has many applications, and can be used to explore a content area, structure domains, classify or reduce data, define relationships, test hypotheses, formulate theories, make inferences, and more. For the purposes of the current project, factor analysis is used for classifying data. More specifically, it is used to develop an empirical typology of drug behaviors. Factor analysis is well suited to this purpose as it groups interdependent variables into descriptive categories and shows how individuals cluster (Rummel 1967, p. 450; Thompson 2004).

The patterns or factors that emerge from this technique represent groups of variables, not groups of individual subjects. This is why it is therefore important to conceptualize the factors as dimensions of drug behavior, rather than descriptions or categories individuals. As above, while traditional typology research was interested in placing individuals into various categories based on their behavior, the current research uses behavior to describe individuals along multiple dimensions of drug-related behaviors via factor analysis. The factors represent latent traits of individuals with respect to one form of criminality, namely drug offending. The ability to capture and describe these latent patterns is a crucial contribution to the literature made by this study. The ability to measure each respondent’s score continuously for each emergent dimension provides a total picture of drug behavior, representing a key advancement in this study.

The results obtained from factor analysis depend heavily on the variables the researcher includes. Careful selection of those variables is thus a crucial step in the process. Here, the goal is to use factor analysis to uncover common patterns of drug use and drug selling. The current study considers variables in three categories; drug use variables, drug selling variables, and other variables known to be salient in the drugs milieu. Since the typologies are intended to be patterns of drug behavior in adolescence, these variables will be gathered from the Phase 1 data in RYDS (Waves 1–9), self-reported by the G2 subjects.⁶

The drug use variables⁷ to be included in the typologies are; age of onset of drug use (age at initiation), number of substances used, number of times used⁸ marijuana, number of times

⁶A note on missing data is necessary prior to continuing. For the drug-related behaviors captured in the factor analysis below, it was possible to use the full sample of 881 subjects who remained in the RYDS study through the end of Phase 1. Missing data on these behaviors was caused by skip patterns in the data collection instrument. During interviews, subjects were first asked if they had engaged in a particular behavior. If the subject said they had, they were then asked follow-up questions, such as how many times they had engaged in the behavior. If the subject said they had not, the follow-up questions were skipped and are therefore missing. Since the methods used here involved counting the number of times a subject engaged in a variety of behaviors, if the subject was missing on a variable because he or she responded previously that they had not engaged in that behavior, they were recoded to a zero on the count variable. Models below are estimated on the full sample of subjects who remained in the study by the end of the phase.

⁷More correctly, these represent concept areas, rather than specific variables. Each of these concepts are measured for each subject for each wave of data collection from Waves 1–9. When estimating the factor analysis, it will be fruitful to include both wave-specific and cross-wave variables, and to measure each of these concepts or constructs in multiple ways. This is because factor analysis requires there to be more cases than variables, and more variables than factors. Factor analysis does not extract the total possible factors that

used hard drugs,⁹ number of waves drugs were used (duration of drug use), number of waves was marijuana used, and number of waves hard drugs were used. The drug selling variables will include age of onset of drug selling, number of substances sold, number of times sold marijuana, number of times sold hard drugs, number of waves drugs were sold (duration of drug selling), number of waves marijuana was sold, number of waves hard drugs were sold, and highest amount of drugs sold.¹⁰ This is not an exhaustive list, but should give the reader a good feel for the variety of variables that are used in the factor analysis.¹¹

Emergent from the process of factor analysis are factor scores. The case's actualized value on each variable is multiplied by its loading on that pattern. The sum of these loading-times-data products yields the factor score. Cases that have high values on the variables entering that pattern will have high factor scores. Cases that have low values on the variables entering that pattern will have low factor scores. Factor scores can then be interpreted as new data for the cases. They can be used to compare cases on the patterns, and can be used as inputs in subsequent analyses.

Recall that in the current study the factors capture underlying or latent dimensions of subjects' patterns on drug behavior. The factor scores are then the empirical scores or quantifications of each subject's participation (or abstention) in each of the emergent factors or dimensions. The factors are not mutually exclusive;¹² each subject receives a score on each factor. This represents another advancement made by the current research. Previous typology research was capable of classifying individuals chiefly on their participation in at most a few behaviors. The current project has argued that this method is an insufficient way to measure the spectrum of an individual's drug behavior.

The argument here is that factor analysis holds high utility in drug research in that it can first identify the correct number of dimensions along which drug offending should be described, and second can provide a score on every dimension for each subject in the data set. Further, those scores will not be artificially reduced into stunted categories. Having done so, the study will further test the utility of the factor analytic model of drug offending by determining the predictive validity of the factor scores in a multivariate context. Specifically, the factor scores will be entered as exogenous variables predicting the individual subjects' number of accrued arrests at Phase 1 (at the end of adolescence), Phase 2 (early adulthood), and Phase 3 (adulthood). The outcome measure of official arrests was chosen because of the

exist in the data. It instead stops when no additional significant or meaningful variance remains to be extracted (Rummel 1970, p. 169). To yield a sufficient number of factors for these purposes then, multiple measures of many constructs are necessary to include meaningful variation. The variables discussed in this section are intended to give the reader an understanding of the concept areas that will be included, rather than a precise list of variables and their construction. This caveat applies equally to the drug selling variables, and variables known to be salient.

⁸Count variables such as this will be measured both across Phase 1 and specific to waves. Alternatively, these could be converted in frequencies per a given unit of time such as week or month (λ). As above, the goal is to provide sufficient variation to yield multiple factors describing adolescent drug behavior patterns.

⁹Id est, drugs other than marijuana.

¹⁰Dollar amount, either during a wave or across the Phase.

¹¹All data on drug behavior measures were drawn from the self-reported items in the RYDS dataset. These data allowed for roughly 300 unique drug measures to be created for each subject, although the majority showed remarkably little variation.

¹²Orthogonal rotation imposes the requirement that factors be mutually uncorrelated, however oblique rotation explicitly recognizes correlation between factors.

overwhelming support in the literature of the notion that involvement in the various forms of drug behaviors is positively associated with arrests.

All procedures and analyses presented below were performed using SAS (version 9.4) and SAS/STAT (version 13.1).

Factor Analysis

Before a factor analysis could be performed it was first necessary to assemble the most extensive possible list of variables pertaining to drug behaviors captured in the RYDS interviews during Phase 1 of the project. These variables were then studied through correlation matrices and entered into a series of factor analytic models to begin the process of isolating the communality and patterns in the data. This process let two themes emerge. First, that the across-phase variables were more powerful descriptors of behaviors than within-wave variables. This came as little surprise. After all, the goal was to describe as complete as possible a picture of drug behavior during adolescence. Measuring behaviors in only one wave is ineffective as the behaviors, and their occurrence, change rapidly. This variation is precisely what the current project is after. The second emergent theme was that the data contain less variation on the substances involved than would be preferred. That is, variables that were specific to particular drugs involved (e.g. cocaine, crack, heroin, etc.) were typically less useful than variables that collapsed drugs into two categories; marijuana and hard drugs. Admittedly, this is unfortunate. However, the simple fact is that in this data set, insufficient numbers of subjects used or sold each of the hard drugs to have statistical variation to exploit on their own.¹³ For this reason, the reader will notice the final factor analysis included “hard drug” variables instead of variables pertaining to individual drugs in this category.¹⁴

All told, the factor analysis model included a total of 33 variables. From these variables, factor analysis extracts a large set of factors. In fact, it extracts as many factors as there are variables. The factors are rank ordered according to the proportion of the variance each factor explains in the variables, with the first factor explaining the most variance. However, only a small subset of factors is retained for further consideration. These original factors are known as the “unrotated” factors,¹⁵ or factor patterns.

Eigen values were also calculated for the factors. Eigen values express the number of variables variance explained by each factor, and are not dependent on the factors being rotated. Eigen values are commonly displayed in Scree plots, with the factor number along the bottom axis and the eigen values along the vertical axis. Factors are considered to not add to the model when their eigen values fall much below 2.0. At this point, the factor is not explaining any more variance than the original variables. Recall that the goal of factor analysis is to collapse variables into more parsimonious descriptions. Factors with eigen

¹³Marijuana, by contrast, provided ample variation.

¹⁴With 3 exceptions, detailed and discussed below.

¹⁵The purpose of the unrotated factor analysis is to determine the correct number of factors. The first unrotated factor pattern identifies the largest pattern of relationships in the data. The second unrotated factor pattern identifies the next largest pattern of relationship in the data, independent of the first pattern. The third unrotated factor pattern identifies the next largest pattern, independent of the first and second patterns, and so on. Thus the amount of variation in the data explained each successive pattern decreases.

values <2 defeat this purpose, and can be discarded as either irrelevant or nonexistent (that is, they can be assumed to reflect measurement error; Abdi 2003; O'Rourke et al. 2013). The Scree plot in Fig. 1 shows the results from this process.¹⁶ The factor analysis from this point on relies on four identified factors.¹⁷

Having estimated an unrotated factor analysis and identified four meaningful factors for this data set, the next step is to perform a rotation on the factors that were retained. In geometric terms, rotation means that the axes (representing the factors) are rotated (around the origin) to more closely fit the vectors (the clusters of interrelated variables). Rotating the factors in no way changes the vectors, it only changes perspective on them. Therefore rotation is used to make interpretation of the factors easier and more reliable.¹⁸

The first rotation performed is an orthogonal rotation.¹⁹ There are several methods of orthogonal rotation, the most popular of which is Varimax rotation²⁰ (Comrey and Lee 2013; O'Rourke et al. 2013). A Varimax solution gives each factor a small number of large factor loadings, and a large number of small factor loadings. This is a simple solution because each original variable is correlated with one factor, so each factor represents a small number of variables. As its name would imply, in orthogonal rotation, the factors remain at right angles to each other. However, the factors themselves may be correlated with one another. Orthogonal rotation allows the factors to rotate in space around the origin, on the condition that the factors maintain orthogonal relationships to each other. This imposes the assumption that factors are uncorrelated with one another, typically an unrealistic assumption.

Oblique rotation, by contrast, not only allows the factors to rotate around the origin, but also to change their angular distance to each other to account for correlations among factors. In other words, the factors are free to take any position in the factor space. This is a more accurate reflection of real-world relationships. Obliquely rotated factors can be easily interpreted by looking at the correlations between the rotated axis and the original variables. These correlations are the factor loadings (Abdi 2003). In obliquely rotated factors, the factor loadings are near zero when a variable is not involved in the pattern, and close to 1.0 when a variable is almost perfectly involved in a pattern. As there are several methods of orthogonal rotation, so too are there several methods of oblique rotation (Comrey and Lee 2013; O'Rourke et al. 2013). One of the most common, and the method used here, is the Promax rotation.²¹ The Promax factor loadings solution is shown in Table 1.

¹⁶As identified in the plot, the first factor has an eigen value of 11.35. Thus it explains the variance of a little more than 11 variables. Similar, with its eigen value of 4.41, the second factor explains the variance of about 4 and half variables. The eigen values for the third and fourth factors are 3.40 and 2.72, respectively. However, the fifth factor has an eigen value of only 1.82. It explains the variance of less than two variables.

¹⁷For completeness, the subsequent analyses were estimated with 5 factors. The fifth factor added practically nothing to the analysis, and results were unchanged. Using four factors is therefore the more parsimonious choice.

¹⁸Reliable here means that the results can be replicated with different data sets (Abdi 2003).

¹⁹When vectors are highly positively related, the angular distance between them will be very small (i.e. close to 0°), and when they are highly negatively related, the angular distance will be near the other extreme (i.e. near 180°). When vectors are unrelated to one another, the angular distance between them is 90°.

²⁰Other methods include the Varimin, Quartimax, and Equimax. Because orthogonal rotation is not the focal rotation, these are not discussed.

²¹This method proceeds in two steps. In the first step, a target matrix is identified using orthogonal rotation whose entries (values) are raised to a power between 2 and 4. Here, the first step rotation method is the Varimax rotation, and the entries are raised to the third power, forcing the values to become bi-polar. The second step is to compute a least squares fit to the Varimax solution.

Results

Factor Scores and Typologies of Drug-Related Behavior

As shown in Table 1, there are four distinct factors emergent from the factor analysis process. These factors represent the underlying dimensions or latent patterns of drug behavior in adolescence that were the target of this project. Once developed, the factors can be given symbolic labels according to the variables they each capture. Naming them *descriptively* gives meaning to their content and more important gives clues to factor content found in other studies (Rummel 1967, p. 471).

Factor 1 can then be descriptively named “Drug Selling.” The variables that load strongly on the Drug Selling factor each pertain to the various aspects of drug selling measured by the original variables. For example, the average number of times a subject sold drugs in a given wave, the most number of times a subject sold hard drugs in a wave, and the most times a subject sold marijuana in a given wave are all found in this factor or dimension. These three variables have very high factor loadings, meaning the correlation between the rotated factor axis and the variables is quite high.²² However, it is equally important to note the variables with large negative factor loadings. These variables were the age of initiation for marijuana selling and the age of initiation for hard drug selling.²³ The negative sign indicates that subjects who sold drugs the most tended to get involved in drug selling early in adolescence (i.e. they had very low scores on the age of initiation variables, which is to say they were younger). What is interesting about this factor is that both marijuana and hard drug variables loaded equally strongly. The interpretation of this result is that within the RYDS sample, the selling of marijuana clustered with the selling of hard drugs and the two were empirically identical in this dimension.

The second factor can be descriptively named “Use of Marijuana.” As the name suggests, and as seen in the table, the variables that loaded on the second factor all describe the subject’s use of marijuana during adolescence. Facets of this dimension include the average number of times marijuana was used in a wave, the most number of times marijuana was used in a wave, and the total number of times marijuana was used during Phase 1. Also found in this factor is the most number of times a subject used any drug. This is not surprising, since this variable was driven by the use of marijuana. In other words, marijuana was the overwhelming favorite for drug use in the RYDS sample. It is important to note that this dimension only describes marijuana use, not hard drug use. Whereas the drug selling variables clustered regardless of whether marijuana or hard drugs were being sold, this is not the case for drug use. Separate dimensions emerged for marijuana use and hard drug use.

The hard drug use dimension is encompassed by Factor 3, and can be labeled “Use of Hard Drugs.” Examples of variables that loaded on this factor include the average number of times hard drugs were used in a wave, the total number of times hard drugs were used in

²²Alternatively, the cosine of the angle between them is small.

²³Also included in this factor were two dummy variables for age of initiation for marijuana selling and age of initiation for hard drug selling. The factor analysis process discards cases listwise if a case is missing on any variable. This means that if a subject never sold hard drugs or marijuana, they would be deleted from the sample. To correct for this, subjects who never initiated were recoded on these variables to their current age. The dummy variables remove this correction, and their factor loadings can be ignored.

Phase 1, the most number of times hard drugs were used in a wave, and the number of times cocaine, crack cocaine, and heroin were used in Phase 1. Especially interesting is that these drugs had significant variation to load individually, despite the small number of users in the RYDS sample. This suggests that use of the three drugs are quite distinct phenomena. As in Factor 1, the age of initiation of hard drug use loaded with a negative factor loading. Subjects with high marks on hard drug use tended to start use early in adolescence.

The fourth and final factor can be descriptively labeled “Age of Initiation.” The clustered variables comprising this dimension are the age of initiation to drug use (any drug use), the age of initiation to marijuana use, and to a lesser extent the age of initiation to hard drug use. The weaker factor loading for hard drug use implies that the age of initiation for hard drug use was less-highly correlated with the factor. This means that the users most typically initiated drug use with marijuana, a finding that is in no way surprising.

Identifying the factors is an important piece of this research, but the central thrust is to use these factors to describe the patterns of drug-related behavior during adolescence. An appropriate question then is, “who do the factors describe?” That is, how can the factor scores be used to explore real samples of individuals? To begin this discussion, one point should be articulated. A factor score of zero does not indicate that a subject does not participate in the factor, or that the pattern does not describe the subject. Factor scores are standardized scores with a mean of zero and standard deviations of one. Thus, a subject with a score of zero on a particular factor falls on the average of the factor.

It is possible to quickly look at the factor patterns to identify roughly how many subjects in the sample can be described as either high or low on the factors. In factor analysis, it is common convention to use a factor score of 0.30 or more to identify scores as being high (or –0.30 or less as low). Due to the overwhelming number of males in the sample,²⁴ males and females will be presented separately and in terms of percentages. These results are shown in Table 2. There are 94 males and 15 females who have high factor scores for the Drug Selling factor (14 and 6 %, respectively). Many more subjects were low on their Drug Selling factor scores, with 224 males (35 %) and 105 females (43 %). In the Use of Marijuana factor (Factor 2), there are 70 (10 %) males and 13 (6 %) females with high factor scores, and 30 (5 %) males and 9 (4 %) females with low factor scores. In other words, marijuana use is more common than not in the RYDS sample. Looking at the Use of Hard Drugs factor, there are 45 males and 9 females (7 and 4 %) with high factor scores, but 19 males and 4 females with low factor scores (3 and 2 %). With such low percentages of subjects who are low on this dimension, we can assert that at least some use of hard drugs was the norm in the RYDS sample, even though persistent or chronic hard drug use was not common.²⁵

The final factor operates in the reverse direction. Low scores on the factor indicate the subjects initiated drug use early. Thus a low score on the factor is the indicator of problematic behavior. There were 230 males (36 %) and 103 females (42 %) with low scores in this dimension. 401 males (63 %) and 140 females (57 %) had high scores on the

²⁴Recall that the original RYDS sample is 75 % male.

²⁵This finding adds to the understanding of the RYDS subjects, long known to be prone to the use of marijuana. The current research indicates that at least sporadic use of hard drugs is not as uncommon as previously thought.

dimension, and initiated drug use late in adolescence.²⁶ It is notable that nearly all of the subjects, both male and female, fell into one of the extremes, initiation drug use early in adolescence or much later. Table 2 presents the numbers and percentages of subjects with high and low factor scores on each of the four factors for easy reference.

At this point the reader will hopefully see the utility of the factor score method of measuring drug-related behavior during adolescence. It identifies succinct and concrete dimensions of drug offending, including common and unique aspects of behaviors. It also provides an empirical basis for the number of factors. Three factors would have been insufficient to describe the variation in the data, but there was not enough variation to warrant five. Finally, it provides each subject with a precise score on all four factors or patterns.

Using Factor Scores to Predict Arrests

To further gauge the utility of the factor analytic model in assessing an individual's drug behavior in adolescence, the factor scores were next used to predict the subjects' number of accrued arrests. By far, the most well-documented, well-researched, and well-supported criminal justice sanction (with respect to drug-related behaviors) is that of arrest.²⁷ Whereas the drug-related behavior measures rely on data self-reported by the G2 subjects, arrest data are derived from official measures from the police record portions of RYDS.²⁸ Here, the focal outcome will be the number of arrests accrued by a subject during the phase. Based on empirical research and theoretical principles, the more heavily drug-involved subjects are hypothesized to have more arrests than those less involved at each phase.

To begin this analysis, the factor scores for each subject were entered into an OLS regression model predicting the G2's number of arrests²⁹ accrued during Phase 1 (ending roughly at age 18). Because drug-related behaviors would be expected to show a relationship with drug-related arrests, the dependent variable in the models contains only arrests for non-drug-related offenses. Also included in this model were several control variables; namely dummy variables to indicate whether the G2 was black or Hispanic, a dichotomous indicator of whether the subject completed high school or a GED (coded 1 = yes and 0 = no), and finally a dichotomous indicator of whether the subject was receiving welfare at the end of Phase 1 (coded 1 = yes and 0 = no). This model is presented as Eq. 1. The joint test of significance statistic presented in the table is a value indicating the statistical significance of the 4 factor scores together in predicting the outcome, number of Phase 1 arrests. This is a statistical test to determine if the four factor scores together have a significant (that is, non-zero) association with the dependent variable.³⁰ In other words, the joint significance test is an

²⁶Or, not at all. Dummy variables account and control for this possibility.

²⁷As is explained below, the arrests considered in the empirical models here exclude drug-related arrests, as the relationship between drug-related behaviors and drug arrests is already so well documented.

²⁸Because agreement between self-reported and official arrests in RYDS has been shown to be quite high (Krohn et al. 2013), using official data here presents little to no bias.

²⁹Arrests in this data set are highly skewed to the right end of the distribution. There are several corrections for this non-normal distribution, including Poisson or negative binomial regression. Poisson regression assumes the conditional variance of the outcome variable be equal to the conditional mean. In all of the arrest outcomes in this project, this assumption is violated. Specifically, the variances were much larger than the mean. The simplest correction proposed in situations like these is to take the natural log of the dependent variable, and this correction was used here. All arrests are in logged units. Because the remainder of the continuous outcome variables exhibit approximately normal distributions, ordinary least squares regression is appropriate.

indicator of whether the factor score dimensions together are significant predictors of the outcome.

$$\text{Arrests}_i = \alpha + \beta_1(F1_i) + \beta_2(F2_i) + \beta_3(F3_i) + \beta_4(F4_i) + \beta_5(\text{Black}_i) + \beta_6(\text{Hispanic}_i) + \beta_7(\text{HS}_i) + \beta_8(\text{Welfare}_i) + \varepsilon_i$$

(1)

As the reader will recall, the original RYDS sample was stratified based on gender. There are roughly 3 males for each female in the sample. Further, gender was a highly salient variable in the review of the literature presented above. That is, various drug offending behaviors do not impact males in the same way as they impact females. For both of these reasons, subsequent analyses were performed separating the male sample from the female sample.

The first panel of Table 3 presents the results of an OLS regression model predicting the number of arrests accumulated in Phase 1 for males. The joint significance test at the bottom of the table indicates that jointly the factor scores are significantly related to the outcome, here the number of arrests. Consistent with this, we can see that three of the four factor scores are independently statistically significant predictors of arrest (Factors 1, 3, and 4). The Factor 1 (drug selling) coefficient of 0.064 indicates that subjects with high factor scores on the drug selling dimension are predicted to have more arrests than their peers with lower factor scores. Similarly, the Factor 3 (Use of Hard Drugs) coefficient of 0.062 indicates that subjects with high factor scores on the hard drug use dimension are also predicted to have more arrests. Finally, the Factor 4 (Age of Initiation) coefficient of -0.133 indicates that subjects who initiated drug use early are predicted to have more arrests than peers who initiated later. The R^2 for this model tells us the model explains just above 18 % of the variance in Phase 1 arrests for males.

The second panel of the table uses the factor scores for drug behavior dimensions to predict the number of arrests incurred by the end of Phase 2, when subjects are roughly 21–23 years old. These arrests do not include any arrests incurred during Phase 1, any arrests for drug-related offenses, and the model includes the same control variables including in the Phase 1 model.³¹ The model also adds control variables for current phase drug use and drug selling. Here again, the factor scores are jointly significant as predictors of arrests at Phase 2 for

³⁰This test amounts to an analysis of variance, distributed as F, with $n-k$ degrees of freedom where n equals the sample size and k equals the number of independent variables. The null hypothesis for the test is one of no relationship, or that none of the factor scores are significantly different than zero.

³¹For completeness, models identical to those presented below were also estimated including measures for drug use and drug selling at Phase 2 and Phase 3. These measures were created by counting the number of use and selling incidents for each subject across the phases. The rationale for their inclusion was that if past drug offending is predicted to be related to these outcomes, then it is reasonable to ask if contemporaneous drug use and drug selling could have a direct impact on the same outcomes. Surprisingly, these measures were never significant in any model, with one exception. Drug use was a significant predictor of arrests at Phase 2 for males. In all cases, the sizes of the coefficients were miniscule in comparison, all less than 0.002. Further, including these measures did not mediate the significance of the adolescent drug offending factor scores in any case. The likely reason is that by Phase 2 and Phase 3, there are so few subjects who continued to use and sell drugs that the variables did not have statistical power. For these reasons, the measures were dropped from the final models presented herein. At the introduction to the project, the competing hypotheses of offenders becoming entrenched in drugs versus aging out were discussed. The results here support the idea that the vast majority of the RYDS subjects aged out of crime by the time they reached adulthood.

male subjects. However, once the control variables are added to the equation, the factor scores are no longer statistically significant. Arrests in Phase 2 are predicted by the number of arrests during Phase 1. Because the factor scores were significant predictors of Phase 1 arrests, the factor scores thus have an indirect effect on Phase 2 arrests, but not a direct effect. The R^2 of this model, 0.17, indicates that the regressors explain roughly 17 % of the variance in Phase 2 arrests.

The results of the models using factor scores to predict arrests for male subjects at Phase are shown in the third panel of Table 3. The factor scores are jointly significant in predicting Phase 3 arrests.³² As seen in the Phase 2 model, the factor scores become insignificant once the control variables are added to the model. The only significant predictors are current drug use and current drug selling. Increases in both of these behaviors are shown to be related to a higher predicted number of non-drug-related arrests during Phase 3. The R^2 of 0.19 suggests that while the regressors explain almost 20 % of the variance, by this time in adulthood the influence of adolescent drug-related behaviors has all but disappeared. This is supported by the fact that the factor scores are not independently significant in predicting arrests at Phase 3. Substantively, this can be interpreted to mean that the direct impact of adolescent drug selling and drug use decays, with more temporally proximate behaviors having the most influence.

We now turn our attention to the female subjects. Structurally, the models predicting the number of arrests for the females are identical to those predicting outcomes for males. The OLS model predicting Phase 1 arrests for females is shown in the first panel of Table 4. We see the factor scores are jointly significant in predicting arrests. Independently, only Factor 2 (marijuana use) is a significant predictor of non-drug-related arrests. This is not surprising given the relatively small proportion of female subjects involved in either drug selling or hard drug use. The coefficient of 0.330 indicates that subjects with higher factor scores on this dimension are predicted to have incurred more arrests during Phase 1 than their counterparts with lower factor scores. Parallel to what was seen with the male arrest model, the R^2 is sizable, as the model explains nearly 15 % of the variance in Phase 1 arrests.

The second panel of Table 4 presents the model predicting the number of arrests incurred during Phase 2. The factor scores for drug-related behavior are jointly significant in predicting these arrests. After including all controls, Factors 1 (drug selling)³³ and 2 (marijuana use) individually are statistically significant predictors of arrests at this Phase. Diverging from what was seen with male subjects, this indicates that adolescent drug behaviors for females have direct effects on new arrests for non-drug-related offenses in early adulthood. These effects are present even after controlling for current phase drug behaviors and prior arrests. Similar to the male subjects, arrests at Phase 2 for females were predicted by arrests at Phase 1. Approximately 24 % of the variance in Phase 2 arrests is explained by this model, a reasonably high figure. For females in early adulthood, the drug behavior dimensions carry a direct relationship with the arrests in early adulthood.

³²Again, these are only arrests incurred during Phase 3.

³³The reader will recall that Factor 1 was not a significant predictor of arrests for female subjects at Phase 1. The significant finding here may signal that for problematic female offenders, drug selling has a longer lasting impact. We are grateful to an anonymous reviewer for suggesting this possibility.

Using the factor scores to predict non-drug-related arrests at Phase 3, controlling for arrests at Phases 1 and 2 and host of other measures, it is clear that by this point in the females' lives the adolescent drug dimensions hold no sway. The full model results are shown in the third panel of Table 4. As the reader can see, the factor scores are not jointly significant as predictors of Phase 3 arrests, nor were any factor scores independently significant. From this we can infer that females are relatively insulated from their adolescent drug-related behavior by the time they reach their 30 s. In other words, females are being arrested at this stage for reasons that are generally independent of their adolescent drug histories. The drug behavior dimensions have no direct relationship with arrests at this stage in life for female subjects, a finding consistent with their male counterparts.

Conclusions

The development and exploration of common patterns of drug-related behavior during adolescence was accomplished through factor analysis. The overwhelming majority of prior work was built on simple categorization of drug-involved individuals, and this categorization ignored the rich variation that exists among such persons. While common patterns among drug-involved individuals exist, the results here suggest there are markedly more subtleties than are usually sought in research. Treating individuals in a purely categorical way yields stilted results. To improve on this strategy, factor analysis was used to develop a more robust typology of drug-related behavior. Four factors emerged, representing four latent or underlying dimensions of drug-related behavior. These dimensions were then explored to give an idea of how prevalent the various patterns are in the RYDS sample. However, this study cautiously advocates against the categorization of individuals. Rather, it is argued here that scoring each subject on each dimension—and by extension creating a unique pattern for each subject—is an important advancement made here.

The factor scores were consistent predictors of non-drug-related arrests at the first phase of data collection, for both male and female subjects. However, looking at the individual factor scores suggests that the mechanisms operate through different dimensions for males than for females. For example, the Factor 3 (hard drug use) and Factor 4 (early initiation) coefficients were significant in predicting male arrests, but never significant in predicting female arrests. Instead, female arrests were predicted by factor scores on the marijuana use dimension at Phases 1 and 2. Importantly, the drug behavior factor scores were salient predictors of non-drug-related arrests for males only through the end of Phase 1. The factor scores remained significant into Phase 2 for females. This suggests that for females, the implications of adolescent drug-related behaviors are more enduring and disruptive. This is likely because females involved in drug selling during adolescence were the more problematic offenders, and this behavior had lasting salience.

The results from the regression models demonstrated that the factor scores were able to predict the outcomes in consistent ways. They also showed the benefit of using measures as encompassing as the factor scores. Simple categorizations would miss the remarkable variation that exists. For example, if one measured only drug use, and ignored drug selling, one would miss the contribution of that important dimension. The early age of initiation to drug-related behavior is rarely included in empirical analyses, but emerged as one of the

most consistent predictors. Similarly, if a drug use category collapsed marijuana use with hard drug use, their unique contributions would be entirely masked. It also argues that measures should be tailored to the data set at hand.

The current study argues that factor scores offer a more complete and holistic way to capture the depth and breadth of drug behavior than the more traditional typology approach. The factor analysis technique is an empirically driven tool to identify commonalities and disparities in the universe of drug-related behaviors. Without first identifying these dimensions in the data, the researcher who categorizes individual subjects is doing so in a somewhat arbitrary way, and may be missing key dimensions of behavior. Further, the factor scores can map out precise patterns of drug behavior that are uniquely descriptive of individual subjects. The scores are rich descriptors of each dimension, and represent marked improvements over the limited or stilted values more commonly encountered in the literature. In sum, the factor analytic model used here is offered as a benchmark, a powerful tool that can be leveraged by future research studying drug-related behaviors.

Of course, the project here has its own weaknesses. Because the data were limited to a single study in the US, external validity is hampered. Replications of this work in other cities where data is available, or indeed with national samples, would prove quite interesting. If results were replicated, one could argue the results shown here are widespread. If not, then the impact of drug-related behaviors on the life course may exhibit regional variations. Similarly, the study used only adolescent drug behaviors, and did not use factor scores to model adult drug behaviors. Whether the results are applicable to only drug-involved adolescents or to drug-involved persons of all ages remains to be seen. Lastly, the factor analytic model of drug-related behavior was limited by the variation on the drug measures included in the RYDS sample. As noted above, there was comparatively little variation on drug use measures for hard drugs. It is feasible that with greater variation, more factors (representative of dimensions of drug-related behaviors) could have emerged from the analysis.

For all these reasons, the current research is by no means intended to be a final word on the subject. Rather, it is hoped that it will be used as part of the foundation on which build a stronger and more well developed literature. Having shown the benefit of measuring drug-related behavior in adolescence in a more complete fashion, future work can build on the efforts here to continue scientific research on drug behavior in all of its forms. Having considered the impact of drug-related behavior on arrests, future work would benefit from exploring its impact in other domains of life. A more complete view of drug behavior and the life course will require additional research. The current project is offered as the next step in science reflecting this small corner of human behavior.

Acknowledgments

Support for the Rochester Youth Development Study has been provided by the Office of Juvenile Justice and Delinquency Prevention (86-JN-CX-0007), the National Institute on Drug Abuse (R01DA005512), the National Science Foundation (SBR-9123299), and the National Institute of Mental Health (R01MH63386). Work on this project was also aided by grants to the Center for Social and Demographic Analysis at the University at Albany from NICHD (P30HD32041) and NSF (SBR-9512290). We are most grateful to the many helpful comments and suggestions given by the anonymous reviewers on earlier versions of this paper.

References

- Abdi, H. Factor rotations in factor analysis. In: EdLewis-Beck, M.Bryman, A., Futing, T., editors. Encyclopedia of social sciences research methods. Thousand Oaks: Sage; 2003.
- Altschuler DM, Brounstein PJ. Patterns of drug use, drug trafficking, and other delinquency among inner-city adolescent males in Washington, DC. *Criminology*. 1991; 29(4):589–622.
- Baumer E, Lauritsen JL, Rosenfeld R, Wright R. The influence of crack cocaine on robbery, burglary, and homicide rates: a cross-city, longitudinal analysis. *J Res Crime Delinq*. 1998; 35(3):316–340.
- Bennett T, Holloway K, Farrington D. The statistical association between drug misuse and crime: a meta-analysis. *Aggress Violent Behav*. 2008; 13(2):107–118.
- Black MM, Ricardo IB. Drug use, drug trafficking, and weapon carrying among low-income, African-American, early adolescent boys. *Pediatrics*. 1994; 93(6):1065–1072. [PubMed: 8197009]
- Blumstein, A., Cohen, J., Roth, JA., Visser, CA. Criminal careers and “career criminals”. Washington: National Academy Press; 1986.
- Braga AA, Pierce GL, McDevitt J, Bond BJ, Cronin S. The strategic prevention of gun violence among gang-involved offenders. *Justice Q*. 2008; 25(1):132–162.
- Brame R, Turner MG, Paternoster R, Bushway SD. Cumulative prevalence of arrest from ages 8–23 in a national sample. *Pediatrics*. 2012; 129:21–27. [PubMed: 22184650]
- Brook DW, Brook JS, Rubenstein E, Zhang C, Saar NS. Developmental associations between externalizing behaviors, peer delinquency, drug use, perceived neighborhood crime, and violent behavior in urban communities. *Aggress Behav*. 2011; 37(4):349–361. [PubMed: 21544831]
- Cartier J, Farabee D, Prendergast ML. Methamphetamine use, self-reported violent crime, and recidivism among offenders in California who abuse substances. *J Interpers Violence*. 2006; 21(4): 435–445. [PubMed: 16501213]
- Caulkins JP, Johnson B, Taylor A, Taylor L. What drug dealers tell us about their costs of doing business. Heinz Research, paper 43. 1998
- Comrey, AL., Lee, HB. A first course in factor analysis. London: Psychology Press; 2013.
- Dave D. Illicit drug use among arrestees, prices and policy. *J Urban Econ*. 2008; 63(2):694–714.
- DeLisi M, Angton A, Behnken MP, Kusow AM. Do adolescent drug users fare the worst? Onset type, juvenile delinquency, and criminal careers. *Int J Offender Ther Comp Criminol*. 2013
- DeLisi M, Vaughn MG, Salas-Wright CP, Jennings WG. Drugged and dangerous prevalence and variants of substance use comorbidity among seriously violent offenders in the United States. *J Drug Issues*. 2015
- Dorsey, TL., Middleton, P. Drugs and crime facts. Bureau of Justice Statistics, NCJ 165148. Washington: US Department of Justice; 2010.
- Fagan J, Chin K-L. Violence as regulation and social control in the distribution of crack. *Natl Inst Drug Abuse Res Monogr Ser*. 1990; 103:8–43.
- Fleisher, MS. Youth gang social dynamics and social network analysis: applying degree centrality measures to assess the nature of gang boundaries. In: Short, JF., Hughes, LA., editors. Studying youth gangs. Lanham MD: Rowan Altamira; 2006. p. 85-98.
- French MT, Roebuck C, Alexandre PK. Illicit drug use, employment, and labor force participation. *South Econ J*. 2001; 68(2):349–368.
- Goldstein PJ. The drugs/violence nexus: a tripartite conceptual framework. *J Drug Issues*. 1985; 15(4): 493–506.
- Gordon RA, Rowe HL, Pardini D, Loeber R, White HR, Farrington DP. Serious delinquency and gang participation: combining and specializing in drug selling, theft, and violence. *J Res Adolesc*. 2014; 24(2):235–251. [PubMed: 24954999]
- Hagelstam C, Hakkanen H. Adolescent homicides in Finland: offence and offender characteristics. *Forensic Sci International*. 2006; 164(2):110–115.
- Harrison L, Gfoerer J. The intersection of drug use and criminal behavior: results from the national household survey on drug abuse. *Crime Delinq*. 1992; 38(4):422–443.
- Huizinga, D., Morse, BJ., Elliott, DS. The national youth survey: an overview and description of recent findings. Boulder: University of Colorado; 1992.

- Johnson BD, Golub A, Fagan J. Careers in crack, drug use, drug distribution, and nondrug criminality. *Crime Delinq.* 1995; 41(3):275–295.
- Kinlock TW, O’Grady KE, Hanlon TE. Prediction of the criminal activity of incarcerated drug-abusing offenders. *J Drug Issues.* 2003; 33:897–920.
- Korcha RA, Cherpitel CJ, Witbrodt J, Borges G, Hejazi-Bazargan S, Bond JC, Ye Y, Gmel G. Violence-related injury and gender: the role of alcohol and alcohol combined with illicit drugs. *Drug Alcohol Rev.* 2014; 33(1):43–50. [PubMed: 24261437]
- Krohn MD, Thornberry TP. Retention of minority populations in panel studies of drug use. *Drugs Soc.* 1999; 14:185–207.
- Krohn MD, Lizotte AJ, Phillips MD, Thornberry TP, Bell KA. Explaining systematic bias in self-report measures: factors that affect the under- and over-reporting of self-reported arrests. *Justice Q.* 2013; 30(3):501–528.
- Langan, PA., Levin, DJ. Bureau of Justice Statistics Special Report. Washington: US Department of Justice; 2002. Recidivism of prisoners released in 1994.
- Lizotte AJ, Krohn MD, Howell JC, Tobin K, Howard GJ. Factors influencing gun carrying among young urban males over the adolescent-young adult life course. *Criminology.* 2000; 38(3):811–834.
- Lyman, MD., Potter, GW. *Drugs in society.* 5th. Cincinnati: Anderson Publishing Co; 2007.
- Maden A, Swinton M, Gunn J. A survey of pre-arrest drug use in sentenced prisoners. *Br J Addict.* 1992; 87:27–33. [PubMed: 1543938]
- McLaughlin CR, Daniel J, Joost TF. The relationship between substance use, drug selling, and lethal violence in 25 juvenile murderers. *J Forensic Sci.* 2000; 45(2):349–353. [PubMed: 10782954]
- Miller, J. *One of the guys: girls, gangs, and gender.* New York: Oxford University Press; 2001.
- Monahan KC, Rhew IC, Hawkins JD, Brown EC. Adolescent pathways to co-occurring problem behavior: the effects of peer delinquency and peer substance use. *J Res Adolesc.* 2014; 24(4):630–645. [PubMed: 25506186]
- Mosher C. Predicting drug arrest rates: conflict and social disorganization perspectives. *Crime Delinq.* 2001; 47(1):84–104.
- Newcomb MD. Drug use and intimate relationships among women and men: separating specific from general effects in prospective data using structural equation models. *J Consult Clin Psychol.* 1994; 62(3):463–476. [PubMed: 8063973]
- Nurco, DN., Kinlock, TW., Hanlon, TE. The drugs-crime connection. In: Inciardi, JA., McElrath, K., editors. *The American drug scene.* 5th. New York: Oxford University Press; 2008. p. 357–370.
- O’Rourke, N., Psych, R., Hatcher, L. *A step-by-step approach to using SAS for factor analysis and structural equation modeling.* Cary: SAS Institute; 2013.
- Passini S. The delinquency–drug relationship: the influence of social reputation and moral disengagement. *Addict Behav.* 2012; 37(4):577–579. [PubMed: 22285118]
- Phillips MD. Assessing the impact of drug use and drug selling on violent behavior in a panel of delinquent youth. *J Drug Issues.* 2012; 42(3):298–316. [PubMed: 26889079]
- Rhodes W, Kling R, Johnston P. Using booking data to model drug user arrest rates: a preliminary to estimating the prevalence of chronic drug use. *J Quant Criminol.* 2007; 23(1):1–22.
- Ricardo IB. Life choices of African-American youth living in public housing: perspectives on drug trafficking. *Pediatrics.* 1994; 93(6):1055–1059. [PubMed: 8197007]
- Roberts J, Mulvey EP, Horney J, Lewis J, Arter ML. A test of two methods of recall for violent events. *J Quant Criminol.* 2005; 21(2):175–193.
- Rosenfeld R, Decker SH. Are arrest statistics a valid measure of illicit drug use? The relationship between criminal justice and public health indicators of cocaine, heroin, and marijuana use. *Justice Q.* 1999; 16(3):685–699.
- Rummel RJ. *Understanding factor analysis.* J Confl Resolut. 1967; 11(4):444–480.
- Rummel, RJ. *Applied factor analysis.* Evanston: Northwestern University Press; 1970.
- Shaw J, Hunt IM, Flynn S, Amos T, Meehan J, Robinson J, Bickley H, Parsons R, McCann K, Burns J, Kapur N. The role of alcohol and drugs in homicides in England and Wales. *Addiction.* 2006; 101(8):1117–1124. [PubMed: 16869841]

- Thompson, B. Exploratory and confirmatory factor analysis: understanding concepts and applications. Washington: American Psychological Association; 2004.
- Thornberry, TP., Krohn, MD., Lizotte, AJ., Smith, CA., Tobin, K. Gangs and delinquency in developmental perspective. Cambridge: Cambridge University Press; 2003.
- Valdez, A., Sifaneck, SJ. Getting high and getting by: dimensions of drug selling behaviors among American Mexican gang members in South Texas. In: Egley, A. Maxson, CL. Miller, J., Klein, MW., editors. The modern gang reader. 3rd. Los Angeles: Roxbury Publishing Company; 2006. p. 296-310.
- Warner BD, Coomer BW. Neighborhood drug arrest rates: are they a meaningful indicator of drug activity? A research note. J Res Crime Delinq. 2003; 40(2):123-138.

Fig. 1.
Scree plot showing eigen values by factors

Table 1

Obliquely rotated factor patterns (Promax rotation)

	Factor 1	Factor 2	Factor 3	Factor 4
Average No. of times sold hard drugs per wave	0.835			
Most No. of times sold hard drugs per wave	0.781			
Most No. of times sold marijuana per wave	0.758			
Total No. of times sold hard drugs, Phase 1	0.730			
Average No. of times sold marijuana per wave	0.727			
Total No. of times sold marijuana, Phase 1	0.721			
Largest hard drug selling group	0.687			
Average size of hard drug selling group	0.671			
Most money made selling marijuana	0.438			
Largest marijuana selling group	0.429			
Average size of marijuana selling group	0.366			
Most money made selling hard drugs	0.348			
Age of initiation to marijuana selling	-0.767			
Marijuana selling dummy variable	-0.781			
Age of initiation to hard drug selling	-0.792			
Hard drug selling dummy variable	-0.797			
Most No. of times used marijuana per wave		0.953		
Most No. of times used all drugs per wave		0.947		
Total No. of times used marijuana, Phase 1		0.938		
Average No. of times used marijuana per wave		0.932		
Average No. of times used all drugs per wave		0.914		
Average No. of times used hard drugs per wave			0.901	
Total No. of times used hard drugs, Phase 1			0.899	
Most No. of times used hard drugs per wave			0.854	
Total No. of times used cocaine, Phase 1			0.804	
Total No. of times used crack, Phase 1			0.570	
Total No. of times used heroin, Phase 1			0.564	
Age of initiation to hard drug use			-0.616	
Hard drug use dummy variable			-0.614	
Age of initiation to drug use				0.943
Drug use dummy variable				0.941
Age of initiation to marijuana use				0.940
Marijuana use dummy variable				0.938

Table 2

Number of males and females with high and low factor scores

	Factor 1	Factor 2	Factor 3	Factor 4
<i>High scores</i>				
Males	94	70	45	230
Females	15	13	9	103
<i>Low scores</i>				
Males	224	30	19	401
Females	105	9	4	140

N (males) = 638; N (females) = 243

Table 3

OLS regression model using factor scores to predict arrests at Phases 1, 2, and 3, male subjects only

	Phase 1 arrests			Phase 2 arrests			Phase 3 arrests		
	Coef.	SE	B	Coef.	SE	B	Coef.	SE	B
F1 (drug selling)	0.065**	(0.023)	0.124	0.031	(0.025)	0.060	0.005	(0.016)	0.018
F2 (marijuana use)	0.014	(0.022)	0.029	0.016	(0.022)	0.035	0.009	(0.015)	0.032
F3 (hard drug use)	0.062**	(0.019)	0.128	0.000	(0.020)	0.001	0.010	(0.012)	0.040
F4 (early initiation)	-0.133***	(0.023)	-0.231	-0.034	(0.026)	-0.059	-0.011	(0.017)	-0.033
Black	0.251***	(0.059)	0.209	0.127*	(0.064)	0.104	0.054	(0.041)	0.077
Hispanic	0.187*	(0.072)	0.123	0.132	(0.078)	0.086	0.041	(0.050)	0.045
Completed HS/GED	-0.166**	(0.057)	-0.113	-0.074	(0.060)	-0.052	-0.057	(0.038)	-0.066
Welfare receipt	0.081	(0.077)	0.040	0.081	(0.080)	0.040	-0.008	(0.053)	-0.007
Current drug use	-	-	-	0.000	(0.000)	0.003	0.036*	(0.018)	0.096
Current drug selling	-	-	-	0.000	(0.000)	0.000	0.219***	(0.037)	0.285
Phase 1 arrests	-	-	-	0.113***	(0.015)	0.326	0.014	(0.010)	0.073
Phase 2 arrests	-	-	-	-	-	-	0.030	(0.011)	0.125
(Intercept)	0.264***	(0.068)	-	0.248**	(0.075)	-	0.077**	(0.048)	-
N	638			524			494		
R ²	0.182			0.172			0.194		
Joint test of significance of factor scores (F value)	26.38***			7.10***			6.20***		

* $p < 0.05$;

**

*** $p < 0.01$;

 $p < 0.001$

Table 4
OLS regression model using factor scores to predict arrests at Phases 1, 2, and 3, female subjects only

	Phase 1 arrests			Phase 2 arrests			Phase 3 arrests		
	Coef.	SE	B	Coef.	SE	B	Coef.	SE	B
F1 (drug selling)	0.031	(0.052)	0.040	0.156**	(0.058)	0.181	-0.029	(0.028)	-0.072
F2 (marijuana use)	0.330**	(0.099)	0.294	0.230*	(0.127)	0.178	-0.053	(0.049)	-0.102
F3 (hard drug use)	-0.053	(0.076)	-0.059	0.063	(0.089)	0.064	0.009	(0.036)	0.024
F4 (early initiation)	-0.015	(0.024)	-0.044	-0.035	(0.027)	-0.089	0.009	(0.012)	0.058
Black	0.090	(0.096)	0.105	0.128	(0.109)	0.129	-0.043	(0.047)	-0.104
Hispanic	0.060	(0.108)	0.060	0.131	(0.121)	0.114	-0.074	(0.052)	-0.155
Completed HS/GED	-0.182**	(0.057)	-0.209	-0.080	(0.067)	-0.080	0.024	(0.028)	0.059
Welfare receipt	0.068	(0.045)	0.098	0.078	(0.053)	0.094	0.007	(0.022)	0.019
Current drug use	-	-	-	0.001***	0.000	0.247	-0.010	(0.015)	-0.046
Current drug selling	-	-	-	-0.061*	0.025	-0.168	0.277***	(0.050)	0.360
Phase 1 arrests	-	-	-	0.040	0.040	0.069	0.038*	(0.016)	0.162
Phase 2 arrests	-	-	-	-	-	-	0.045**	(0.013)	0.236
(Intercept)	0.214*	(0.106)	-	0.100	(0.123)	-	0.018	(0.052)	-
N	243			215			204		
R ²	0.150			0.239			0.242		
Joint test of significance of factor scores (F value)	5.46***			9.25***			0.28		

* $p < 0.05$;

** $p < 0.01$;

*** $p < 0.001$