2019

EXAMINING THE UTILITY OF BEHAVIORAL ECONOMIC DEMAND IN ADDICTION SCIENCE

Justin Charles Strickland
University of Kentucky, justrickland@uky.edu
Digital Object Identifier: https://doi.org/10.13023/etd.2019.030

STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained needed written permission statement(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine) which will be submitted to UKnowledge as Additional File.

I hereby grant to The University of Kentucky and its agents the irrevocable, non-exclusive, and royalty-free license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless an embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student's advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student's thesis including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Justin Charles Strickland, Student

Dr. William W. Stoops, Major Professor

Dr. Mark T. Fillmore, Director of Graduate Studies

EXAMINING THE UTILITY OF BEHAVIORAL ECONOMIC DEMAND IN ADDICTION SCIENCE

DISSERTATION

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Arts and Sciences at the University of Kentucky

By
Justin Charles Strickland

Lexington, Kentucky

Director: Dr. William W. Stoops, Professor of Psychology

Lexington, Kentucky

2019

ABSTRACT OF DISSERTATION

EXAMINING THE UTILITY OF
BEHAVIORAL ECONOMIC DEMAND IN ADDICTION SCIENCE

The marriage of perspectives from behavioral economic theory and learning theory has the potential to advance an understanding of substance use and substance use disorder. Behavioral economic demand is a central concept to this interdisciplinary approach. Evaluating demand in the laboratory and clinic can improve previous research on the relative reinforcing effects of drugs by accounting for the multi-dimensional nature of reinforcement rather than viewing reinforcement as a unitary construct. Recent advances in the commodity purchase task methodology have further simplified the measurement of demand values in human participants. This dissertation project presents a programmatic series of studies designed to demonstrate the utility of using a behavioral economic demand framework and the purchase task methodology for understanding substance use disorder through basic and applied science research. Experiments are presented spanning a continuum from theoretical and methodological development to longitudinal work and clinical application. These experiments demonstrate three key conclusions regarding behavioral economic demand. First, behavioral economic demand provides a reliable and valid measure of drug valuation that is applicable to varied drug types and participant populations. Second, behavioral economic demand is a stimulus-selective measure specifically reflecting valuation for the commodity under study. Third, behavioral economic demand provides incremental information about substance use in the laboratory and clinical setting above and beyond traditional measures of reinforcer valuation and other behavioral economic variables. These findings collectively highlight the benefits of behavioral economic demand and provide an important platform for future work in addiction science.

KEYWORDS: Behavioral Economics; Demand; Reinforcer; Substance Use Disorder

Justin Charles Strickland

January 18 2019
Date

EXAMINING THE UTILITY OF
BEHAVIORAL ECONOMIC DEMAND IN ADDICTION SCIENCE

By

Justin Charles Strickland

William W. Stoops
Director of Dissertation

Mark T. Fillmore
Director of Graduate Studies

January 18 2019
Date

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the caring, attentive, and thoughtful support of my advisor, Dr. William Stoops. I would like to express my deepest appreciation for his contributions to my professional and personal growth from this mentorship.

I would also like to thank other committee members on this dissertation as well as the many other individuals throughout graduate school that have provided keen insights into scientific interests and goals, including Drs. Joshua Lile, Craig Rush, Carrie Oser, Joshua Beckmann, Brady Reynolds, Michele Staton, David Fardo, Michael Wesley, Ryan Lacy, and Mark Smith.

I would also like to thank past and current members of the Laboratory of Human Behavioral Pharmacology for support in data collection efforts. In particular, I would like to thank Drs. Levi Bolin and Katherine Marks for their support throughout graduate school as mentors, collaborators, and friends.

Finally, I would like to thank Joanne, David, and Kevin Strickland and Jessica Annonio for their love and encouragement throughout this process.

This research was supported by the National Science Foundation 1247392 [JCS], graduate research grants from the Psi Chi Honor Society [JCS], pilot research funds from the University of Kentucky Department of Behavioral Science and Center on Drug and Alcohol Research [JCS], and professional development funds from the University of Kentucky [WWS].

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

GENERAL INTRODUCTION

Introduction

Over half (52.1%) of United States adults 18 or older reported lifetime illicit drug use in 2017 and a fifth (19.3%) reported past year use (Center for Behavioral Health Statistics 2018). A majority of adults also reported past month alcohol use (55.9%) and a quarter reported past month tobacco use (24.2%). These estimates of substance use prevalence are mirrored by those showing high and relatively stable patterns of substance use disorder. Among United States adults 18 or older, 6.8 million met criteria for a substance use disorder and 14.0 million for an alcohol use disorder according to 2017 estimates (Center for Behavioral Health Statistics 2018). Substance use is also highly comorbid with other physical and mental health problems, frequently exacerbating the symptoms and trajectories of these conditions (Kessler et al. 1996; Regier et al. 1990; Stein 1999). Estimates from 2017, for example, showed that individuals reporting past year mental illness also reported rates of substance use disorder that were six-fold higher than those without mental illness (8.6% versus 1.4%) and rates of alcohol use disorder that were three-fold higher (12.6% versus 4.1%) (Center for Behavioral Health Statistics 2018). These values from the National Survey on Drug Use and Health (NSDUH) likely represent a conservative estimate of point prevalence given that homeless and incarcerated populations are not included, thereby suggesting that even higher rates of substance use and comorbidities likely occur in the United States (Caulkins et al. 2015a; Caulkins et al. 2015b).

The public health impact of substance use is also clear and staggering. For example, the annual economic costs of excessive drinking are estimated at $250 billion (Sacks et al. 2015) and 5.3% of global mortality is attributable to alcohol consumption (World Health Organization 2018). Tobacco cigarette use remains the leading cause of

preventable death globally and nationally and contributes to nearly 500,000 deaths per year in the United States alone (US Department of Health Human Services 2014). Similar high economic costs ($193 billion annually) and health harms are observed for illicit substance use (National Drug Intelligence Center 2011). It is without question that substance use is relevant to any public health dialogue, especially when individuals experiencing some direct or indirect impact of substance use represent a sizable minority, if not the majority, of the population.

Despite great efforts to address substance use disorder in the past decades, current approaches lack universally effectiveness, pose numerous treatment barriers, and still result in high rates of relapse (e.g., Czoty et al. 2016; Priester et al. 2016). For example, Czoty and colleagues (2016) found that although 64 putative medications had been tested in over 100 randomized, placebo-controlled clinical trials for cocaine use disorder, none had advanced to the stage of FDA approval. Even when approved and effective behavioral and pharmacological treatments are available, systemic and systematic barriers such as stigma, individual vulnerabilities, and service availability result in high rates of underutilization (e.g., only 12% of individuals in need of substance use treatment received it in 2017) (Center for Behavioral Health Statistics 2018; Priester et al. 2016). The application of alternative and innovative theoretical approaches is therefore essential for identifying novel prevention and treatment targets as well as for enhancing access and adherence to those effective strategies that do exist.

The marriage of perspectives from the pharmacological, psychological, and economic sciences has the potential to advance our understanding of substance use disorder in this way (Bickel et al. 2016a; MacKillop 2016). These interdisciplinary approaches broadly propose that substance use disorder is associated with clear patterns of maladaptive decision-making and choice. For example, the reinforcer pathology model suggests that substance use is characterized by a persistent high

valuation for drugs of abuse combined with an excessive preference for immediate reinforcers over long-term health consequences (Bickel et al. 2017). Such integration of tools and theory across disciplines exemplifies a multi-faceted approach for generating novel and holistic insights to address substance misuse and to improve and enhance these underlying disciplines understanding of behavior and choice.

A key principle at the intersection of pharmacology, psychology, and economics is behavioral economic drug demand. Demand is operationally defined as the consumption of a good at a given cost. A demand curve describes this functional relationship across a range of costs (i.e., unit prices). The concept of demand simply translates to psychology and the experimental analysis of behavior when a good is defined as a reinforcer (e.g., food, drugs) and a cost defined as an operant requirement (Hursh and Roma 2013). Evaluating demand has the potential to advance previous research on the relative reinforcing effects of drugs by accounting for the multi-dimensional nature of reinforcement rather than viewing reinforcement as a homogenous and unitary construct (Hursh and Silberberg 2008; Johnson and Bickel 2006). Demand can also be easily measured in the human laboratory or clinic using the commodity purchase task. Participants are asked in this procedure to report hypothetical consumption of a good (e.g., alcohol drinks) across a range of prices (e.g., $0.01, $1.00/drink). Such an approach is particularly appealing because of its cost and time efficiency and adaptability for populations for whom drug self-administration is not ethically or practically feasible (e.g., patients in residential treatment; those with medical contraindications).

Although existing applications of demand and the purchase task methodology to drug-taking behavior have produced promising results, this literature is still in its infancy and basic and applied science gaps still exist. **The overarching framework for this dissertation project is the utilization of behavioral economic demand in basic and**

applied research to better understand substance use disorder. These studies span a continuum from theoretical and methodological development to clinical application. Below, I begin by providing a general overview of the history of behavioral economics and behavioral economic demand in psychological science, broadly, and addiction science, specifically. I then describe the procedures for evaluating behavioral economic demand with a particular emphasis on the purchase task procedure and existing studies evaluating its psychometric properties. A brief review of the literature utilizing the purchase task procedure to evaluate pharmacological, environmental, and individual differences associated with behavioral economic demand is then provided. I conclude this introduction by presenting an overview of the aims of this dissertation as well as the five primary experiments.

Behavioral Economics and Psychological Science

Classical and neo-classical economic theories posit that economic actors and markets operate rationally and that decisions are executed based on a rationality assumption (e.g., to maximize economic gain). Central to this approach is the expected utility hypothesis. This hypothesis posits individual's maximize expected utility of choice when presented with uncertain outcomes (i.e., maximize subjective value multiplied by expected probability) (Von Neumann and Morgenstern 2007). This mathematical representation of expectation helped to guide and formalize a principal theory of rational choice in economic and human decision-making. However, the expected utility hypothesis was not without contest. One of the most prominent and systematic demonstrations of expected utility violations was the Allais Paradox (Allais 2008). Economist Maurice Allais demonstrated that individuals often reversed decisions when presented with gambles of common outcomes. For example, consider the following gambling pairs:

Gamble 1: A = $1000 at 100%; B = $1000 at 89%, $5000 at 10%, and $0 at 1%

Gamble 2: A' = $1000 at 11% and $0 at 89%; B' = $5000 at 10% and $0 at 90%

Expected utility theory posits that reducing each gamble by removing common outcomes results in identical gambles for each pair (i.e., identity gambles):

A/A' = $1000 at 11%

B/B' = $5000 at 10% and $0 at 1%

Individuals should then by an expected utility hypothesis show equivalence in choice across these two gambling sets. Instead, Allais and others observed that individuals tended to prefer A to B and B' to A' thereby demonstrating choice reversals in direct opposition to expected utility hypotheses.

Cognitive psychologists formed a cohesive challenge to rational economic approaches in the 1960s arguing that findings such as the Allais paradox and others from the psychological literature demonstrated clear violations of rationality assumptions. Parts of these challenges were codified in 1979 in the seminal paper "Prospect Theory: An Analysis of Decision Under Risk" by Daniel Kahneman and Amos Tversky. Kahneman and Tversky (2013) argued that cognitive psychology could be used to explain deviations from neoclassical theories and that suboptimal behavior is a consequence of systematic choice biases that depart from traditional economic decisions (i.e., expected utility decisions). This and related texts spurred the growth of behavioral economics and an attempt to explain through psychological factors choice and decision-making.

Contemporary theoretical models in addiction science share a common interest with these behavioral economic models by positing that addiction is a disorder directly related to choice and decision-making. For example, Lamb and Ginsburg's (2017) Behavioral Allocation Disorder (BAD) approach argues that substance use behaviors should be framed as decisions to allocate behavior to drug use over other more prosocial alternatives. Similarly, Heyman (2009) argues in his text *Addiction: A Disorder of Choice*

that choice models, such as matching law and hyperbolic discounting, provide robust prediction of substance use patterns and that addiction is ultimately an example of typical everyday choice, albeit a self-destructive one (see also Heyman 2013). These accounts parallel the diagnostic criteria for substance use disorder in which behaviors relevant to choice and decision-making are heavily featured (e.g., an individual's use of more drug than intended; unsuccessful efforts to control one's drug use; using drugs to the exclusion of other activities) (American Psychiatric Association 2013). These theories and clinical features make the application of behavioral economics and its focus on choice at the bounds of rationality a logical one.

This use of economic decision-making to understand drug-taking behavior is not without historical precedent. Pioneers in the field of the experimental analysis of behavior understood the benefits of borrowing a framework developed in this challenge to rational economics. Howard Rachlin, for example, effectively applied concepts from economics in his contributions to learning theory, including the matching law and discounting processes (Rachlin 1974, 1980, 2006; see related contributions from Herrnstein 1961; Herrnstein 1990). Rachlin argued in this body of work that concepts such as substitutability from economic demand theory could be used to help understand behavioral allocation, in particular decisions made in response to environmental fluctuations in response cost (Rachlin et al. 1976). Steven Hursh presented similar arguments in the early 1980s arguing that concepts such as elasticity and economic substitutability could be used to understand choice behavior (Hursh 1980). Such ideas of applying economic frameworks to psychological theories of choice were later extended to drug self-administration and abuse liability testing (Bickel et al. 1990; Bickel et al. 1991; Hursh 1991).

The last decade has witnessed a renewed interest in the use of economic theory and behavioral economics to understanding drug valuation and reinforcement. This

resurgence can be partially traced to the proliferation of the purchase task methodology in the human laboratory and clinic and the development of new models for testing drug demand (see reviews by Hursh and Roma 2013; Koffarnus and Kaplan 2017; MacKillop 2016). The following section reviews this emerging literature on behavioral economic demand and the methods used to collect and analyze demand data.

Behavioral Economic Demand

The quantitative analysis of demand is fundamental to understanding consumer choice in microeconomic theory. However, such applications have only recently gained a widespread popularity and application in addiction science. Extending theories and principles used to explain demand for traditional commodities to drug-taking behavior is logical considering the shared interests of economists and psychologists, including the value of goods (i.e., reinforcers) and how behavior is allocated under constraint (i.e., operant choice). Demand is operationally defined as the consumption of a good at a given cost and a demand curve describes this functional relationship across a range of costs (i.e., unit prices). These concepts easily translate to psychology and the experimental analysis of behavior when a good is defined as a reinforcer (e.g., food, drugs) and a cost defined as the operant requirement on a particular schedule of reinforcement (Bickel et al. 2014; Hursh 1984; Hursh and Roma 2013; Rachlin et al. 1976). The unit price for a particular drug commodity may be defined by the operant requirement needed to obtain that drug (i.e., unit price = responses required/dose). Manipulating either the dose delivered or work necessary to deliver that dose changes this unit price. Subsequent observation of responses across a range of unit prices then provides a means of generating demand functions effectively and efficiently. Put in these terms, the extensive of economic literature regarding mechanisms of demand effectively translates to research conducted in the behavioral pharmacology laboratory and clinical setting.

The promise of behavioral economic demand compared to traditional measures of drug self-administration (e.g., response rate, infusions delivered, breakpoint) is the efficient isolation of behavioral mechanisms underlying drug effects (Hursh and Roma 2013). Evaluating demand in this way accounts for and describes a multi-dimensional nature of reinforcement rather than viewing reinforcement as a unitary construct (Hursh and Silberberg 2008; Johnson and Bickel 2006). Such an isolation of behavioral mechanisms is foundational to behavioral pharmacology (Thompson and Schuster 1968) and is consistent with recent appeals for research determining the behavioral mechanisms mediating drug-taking behavior and drug effects (Pitts 2014).

Mathematical Models of Demand

Theoretical and empirical accounts support the notion that demand curves functionally capture two behavioral mechanisms underlying substance use: 1) demand intensity and 2) demand elasticity (Bidwell et al. 2012; Hursh and Silberberg 2008; Mackillop et al. 2009). Demand intensity represents the consumption of a commodity at a theoretical unit price of zero or when the commodity is free and is thought to represent a hedonic set point of consumption. Demand elasticity reflects how sensitive the consumption of a good is to changes in price. Other measures of demand, such as breakpoint (i.e., price point at which consumption drops to zero), O_{max} (i.e., maximum expenditure), and P_{max} (i.e., price point at maximum expenditure), cluster with demand elasticity in factor analytic studies, which is not surprising given that these measures are derivatives of the elasticity value (Aston et al. 2017; Bidwell et al. 2012). Demand intensity and demand elasticity alter the shape and position of a demand curve and independently influence drug consumption. This is important given that understanding the specific mechanism(s) by which manipulations affect drug-taking behavior is critical for the design and dissemination of interventions to address substance use disorders.

Mathematical representations of demand have developed in a relative parallel

fashion to those procedures used to generate those curves. One of the most popular of these equations, the exponential equation, plots consumption as a nonlinear function of price, demand intensity, and demand elasticity (Hursh and Silberberg 2008):

$$\log_{10}Q = \log_{10}(Q_0) + k(e^{(-\alpha * Q0 * C)}-1)$$

Where Q = consumption at a given price; Q_0 = derived demand intensity (consumption at a hypothetical zero price); k = a constant that denotes the range of consumption values in $\log_{10}$ units; C = commodity price; and α = derived essential value (a measure of demand elasticity). Greater values of Q_0 indicate greater consumption at a theoretical price of zero or greater demand intensity. Higher values of α indicate a greater demand elasticity or greater change in consumption with change in unit price. An intervention to address substance use will ideally decrease Q_0 and/or increase α, thereby decreasing demand intensity and/or increasing demand elasticity, respectively.

More recently, the exponentiated equation has been proposed as an alternative equation (Koffarnus et al. 2015). Zero consumption values (i.e., prices at which no commodity is purchased, commonly observed at high prices) present quantitative challenges when applying the exponential equation (Koffarnus et al. 2015; Yu et al. 2014). This is because the exponential model requires the logarithmic transformation of consumption, represented by the left side of the equation, which is mathematically impossible for zero. One common solution to this problem is replacing zeros with small, non-zero values (e.g., 0.1, 0.01, or 0.001) before fitting the exponential model. The selection of replacement values, however, has considerable effects and can differentially impact the resulting outcomes given that the differences between 0.1, 0.01, and 0.001 are sizable when considered on a logarithmic scale. Koffarnus and colleagues (2015) developed a modified equation in which both sides of exponential model are raised to the power of 10:

$$Q = Q_0 * 10^{k(e^{(-\alpha * Q0 * C)} - 1)}$$

The “exponentiated” model removes the need for logarithmic transformation and allows for the inclusion of zero in the model fitting. This modification improved fits in both the initial demonstration using cigarette demand data (Koffarnus et al. 2015) as well as for alcohol, cigarette, and cocaine demand in an independent follow-up evaluation (Strickland et al. 2016b).

Other alternatives to analyze demand data have also recently been explored. These approaches include those using mixed-effect modeling in traditional nonlinear forms (Yu et al. 2014) or using left-censored (Liao et al. 2013), two-part (Zhao et al. 2016), or Bayesian approaches (Ho et al. 2018). These methods are designed to address many of the problems identified above (e.g., zero consumption values) and have proved beneficial in simulation analyses. However, the appeal of these procedures is currently limited due to factors such as modeling that can only be conducted using certain statistical programs, strong assumptions about the nature of zero consumption responses (e.g., that these responses represent undetected consumption versus desired abstinence), and the lack of independent follow-ups testing these models in alternative data sets and incrementally above existing methods.

Measures of Behavioral Economic Demand

A number of procedures have been developed to assess demand in animal subjects and human participants. Common among these methods is the manipulation of unit price, typically by changing the dose delivered and/or ratio requirement on the active schedule of reinforcement. For example, if unit price is functionally defined as lever presses needed to obtain 1.0 mg of cocaine then increasing the dose delivered while maintaining the fixed-ratio (FR) requirement will decrease the unit price (i.e., “more bang for the buck”). On the other hand, increasing the FR requirement while holding the dose

constant will increase the unit price (i.e., “pay more for the same good”). Not only is this concept of unit price useful for laboratory research, but it may also be easily applied to traditional purchasing behavior wherein increased cost per unit good (e.g., cost per pack of cigarettes; per bottle of wine) functionally increases unit price.

Either between- or within-session methods may be used to manipulate the FR requirement or dose delivered to generate demand curves. When using between-session techniques, the FR requirement or dose delivered will vary between each session to manipulate unit price (e.g., Johnson and Bickel 2006; Peitz et al. 2013). These procedures are appealing because they limit the possibility for carryover effects. However, between-session demand curves are also liable to extraneous variables that cause daily fluctuations in behavior and are less cost and time effective. In contrast, within-session procedures vary unit price within a single session by using systematic “bins” or components (see review by Bentzley et al. 2013). In the animal laboratory, within-session manipulations are typically accomplished using a “threshold procedure” in which subjects complete successive 10 minute components with progressively decreasing doses delivered to increase unit price (Bentzley et al. 2014; Bentzley and Aston-Jones 2017). The threshold procedure has proved useful for addressing existing gaps in preclinical research given its ability to generate demand curves within a single session allowing for high-throughput and high-resolution evaluation of changes in demand as a function of individual subject characteristics (e.g., hormonal fluctuations; acute drug administration).

Commodity purchase tasks, a form of within-session demand curves, have become the most popular method to examine economic demand in the human laboratory and clinic (see reviews by Kaplan et al. 2018; MacKillop 2016; Reed et al. 2013). The commodity purchase task is a questionnaire in which individuals are asked to report consumption of specific commodities (e.g., cigarettes) across changes in price per unit

(e.g., $0.01, $0.10, $1.00 per cigarette.). Although some parameters can differ across different task implementation (e.g., price range or time of purchase; see review by Kaplan et al. 2018), most purchase tasks utilize a consistent set of instructions. These instructions typically specify that choices are to occur in a closed-economy (i.e., this is the only opportunity to purchase the commodity), that the commodity cannot be stockpiled or sold later and all that is purchased must be consumed, and that the participant has the income available to make these purchases that they usually would. Hypothetical choice is frequently used, but correspondence is generally good between hypothetical and realized outcomes, supporting the validity of assessing hypothetical drug commodity choices (Amlung et al. 2012; Amlung and MacKillop 2015; Wilson et al. 2016).

One of the first studies to evaluate demand using the purchase task determined simulated (hypothetical) heroin and cigarette demand in outpatients recruited from a buprenorphine clinic (Jacobs and Bickel 1999). Although few studies expanded upon this initial demonstration in the years immediately following, purchase tasks have proliferated in the past decade. This proliferation has led to the use of the purchase task procedure to assess demand for a diverse and growing list of substances, including alcohol, nicotine (e.g., cigarettes, e-cigarettes), cannabis, cocaine, and prescription drugs (e.g., Amlung and MacKillop 2015; Aston et al. 2015; Bruner and Johnson 2014; MacKillop et al. 2008; Pickover et al. 2016; Stoops et al. 2016).

Most of the extant literature has focused on own-price elasticity and demand (i.e., demand for a substance in isolation of other commodities). However, recent research has expanded these efforts to include evaluation of cross-price elasticity and economic substitutability (e.g., Johnson et al. 2017b; Murphy et al. 2016; Peters et al. 2017; Tucker et al. 2017). Cross-price demand represents the responsiveness of quantity demanded for a good as a function of the change in price of another good. Cross-price elasticity

then is the mathematical relationship between price-changes in the price-manipulated commodity and demand for the alternative price-fixed commodity. Commodities may function as a substitute meaning that as the price increases for the price-manipulated good that consumption increases for the alternative (i.e., positive cross-commodity elasticity). Coca-Cola® and Pepsi® are prototypic substitutes because as the price increases for one of these products a consumer would presumably increase consumption of the alternative (i.e., the products *substitute* for one another). Commodities may also function as complements meaning that as the price increases for one good that consumption decreases for the alternative (i.e., negative cross-commodity elasticity). Hotdogs and hotdog buns are prototypic complements because as the price increases for one of these products, a consumer would presumably decrease consumption of the alternative (i.e., the products *complement* one another). The purchase task methodology may be simply adapted to index cross-commodity demand by evaluating consumption of price-varying and price-fixed commodities in concert with one another.

One of the clear benefits of the purchase task is the ease and efficiency of administration for a variety of research purposes. These tasks are similar to the threshold procedure used in the animal laboratory in that unit price is varied within a single session. Transformation of the price-level consumption from purchase tasks into demand curves allows for the examination of specific behavioral mechanisms of demand, including intensity and elasticity as reviewed above. Hypothetical choice also allows for the use of drug purchase tasks in populations for whom other measures of drug use, such as drug self-administration, are impractical or not ethically feasible. This is particularly relevant for the clinical application of the purchase task procedure given that these populations, including treatment-seeking patients or those with medical contraindications, represent a substantive portion of the population to whom

interventions development efforts should generalize.

Purchase Task Psychometrics

The proliferation of purchase task research has also resulted in a growing literature demonstrating the psychometric properties of these tasks. A review of studies using the alcohol purchase task provides a comprehensive example of this evidence supporting the underlying psychometric properties of the purchase task procedure (see additional review in Kaplan et al. 2018; MacKillop 2016). Note that similar results have been observed for the cigarette purchase task, albeit in a less comprehensive manner (e.g., Few et al. 2012; MacKillop et al. 2008; Murphy et al. 2011).

Several studies have supported test-retest reliability of the alcohol purchase task in college populations (Acuff and Murphy 2017; Murphy et al. 2009). For example, one study demonstrated acceptable test-retest reliability over a one-month period in college drinkers (r_{xx} = .67 for demand intensity and r_{xx} = .71 for demand elasticity) (Acuff and Murphy 2017). Other studies have demonstrated the construct validity of the task by revealing associations between alcohol demand and alcohol use frequency and severity in college students (Murphy and MacKillop 2006; Murphy et al. 2009) and community samples (Amlung et al. 2017a). Morris and colleagues (2017) have more recently shown that the alcohol purchase task retains this construct validity when translated to the online crowdsourcing platform Amazon Mechanical Turk (mTurk). Specifically, demand intensity and elasticity in that study were associated with use severity (i.e., AUDIT scores) after controlling for relevant demographic covariates. A meta-analysis of this literature by Kiselica and colleagues (2016) reached a similar conclusion regarding construct validity reporting an association between demand and alcohol-related measures with a larger effect size for demand intensity (r = .34 to .51) than elasticity (r = -.11 to -.20). This meta-analysis also concluded, however, that the alcohol purchase task may provide only limited incremental validity over other established measures of alcohol

use (e.g., the AUDIT) and future work is needed to test this assertion more thoroughly.

Few studies have evaluated the relationship between purchase task and self-administration data. Those studies that do exist have shown good correspondence of purchase task data with the outcomes from traditional procedures. For example, Chase and colleagues (2013) found that measures of cigarette demand were predictive of choices made in a concurrent cigarette-chocolate choice task. Similarly, Amlung and colleagues (2012) showed that estimated consumption on a hypothetical alcohol purchase task closely aligned with actual consumption during a later self-administration period. Other research has demonstrated similarities in responding for hypothetical and incentivized purchase tasks in order to support construct validity (Amlung et al. 2012; Amlung and MacKillop 2015; Wilson et al. 2016). In one study, alcohol demand on hypothetical and incentivized tasks were highly correlated at both individual price points and for overall demand metrics (Amlung et al. 2012). These relationships were later replicated with a similar design conducted in an independent sample of heavy drinkers (Amlung and MacKillop 2015). Wilson and colleagues (2016), however, did report some differences between real and hypothetical purchase data for cigarette demand. These differences primarily reflected lower elasticity parameters in the hypothetical condition, whereas intensity values were similar in magnitude. Additional work is needed to replicate these effects in larger and more heterogeneous samples as well as to determine if these findings of the relationship between hypothetical and incentivized demand generalize beyond alcohol and cigarette use.

Although the clinical relevance of demand is still under investigation, preliminary evidence suggests that alcohol demand may help to identify behavioral mechanisms underlying effective interventions (Bujarski et al. 2012) or function as prognostic variables predicting treatment success (MacKillop and Murphy 2007; Murphy et al. 2015). In a group of college student drinkers, for example, changes in alcohol demand

intensity observed immediately following a brief intervention were predictive of drinking behavior at one-month follow up (Murphy et al. 2015).

Pharmacological, Individual Difference, and Environmental Effects on Demand

The following sections review research examining pharmacological, individual difference, and environmental influences on behavioral economic demand. A particular focus is placed on the purchase task methodology given its prominent role in research with human participants. Corroborating evidence from alternative procedures and/or the preclinical literature is also offered, as available. This review is intended to highlight ways in which the purchase task methodology has been validated through replication of well-described effects as well as existing gaps in the clinical application of demand in addiction science.

Pharmacological Effects on Alcohol Demand

A number of studies have examined pharmacological influences on alcohol demand in human participants. In this regard, several individual differences in comorbid drug use have been related to alcohol demand. For example, college students who regularly smoke show decreased demand elasticity for alcohol relative to those who do not smoke (Yurasek et al. 2013). Similar increased alcohol demand among individuals reporting tobacco cigarette use was observed in another study conducted with a community sample indicating that these results were not limited to college populations (Amlung et al. 2017a). In another study, greater demand intensity and lower demand elasticity were observed in college students who reported frequent use of caffeinated alcoholic beverages (Amlung et al. 2013).

Alcohol demand also appears sensitive to direct pharmacological manipulation. For example, one study found that acute doses of alcohol designed to increase blood alcohol concentration (BAC) to 0.10% increased intensity, maximum expenditure, and breakpoints during the ascending, but not descending, limb of the alcohol curve (Amlung

et al. 2015a). This finding was consistent with an earlier study demonstrating increases in alcohol demand breakpoints and P_{max} following intravenous alcohol administration (Bujarski et al. 2012). This latter study also evaluated the effects of the opioid antagonist naltrexone on alcohol demand and revealed an attenuation of alcohol demand intensity, breakpoint, and expenditure in the naltrexone group following both alcohol and placebo pre-treatment (Bujarski et al. 2012). This finding is important because it suggests that the purchase task assay is sensitive to the demonstrated effective pharmacotherapy naltrexone (Maisel et al. 2013; Rosner et al. 2010).

Pharmacological Effects on Cigarette and Nicotine Demand

A considerable body of work has also evaluated pharmacological variables affecting cigarette demand. Most of these studies have examined approved pharmacological treatments for smoking cessation, including the monoamine transport inhibitor bupropion and the nicotinic partial agonist varenicline. For example, one study investigated the effects of bupropion treatment on cigarette demand during the one-week prior to initiation of a smoking cessation attempt (Madden and Kalman 2010). Bupropion failed to alter demand for cigarettes at that one-week time point, however changes in demand elasticity at one week were predictive of cigarette abstinence at treatment follow-up (i.e., 10 weeks later). Four studies have examined the effects of varenicline on cigarette demand with mixed findings (Green and Ray 2018; McClure et al. 2013b; Murphy et al. 2017; Schlienz et al. 2014). In the first of these studies, one week of varenicline exposure increased demand elasticity relative to placebo (McClure et al. 2013b). Another study observed significant reductions in demand intensity following a one-week run up of varenicline or nicotine replacement patch medication prior to a quit attempt date (Murphy et al. 2017). Notably, the magnitudes of reduction in demand intensity in that study were predictive of length of abstinence at 1 and 3 months. Significant reductions in O_{max}, but no changes in other demand outcomes, were observed in a third

study in which participants received 10-days of varenicline treatment (Green and Ray 2018). A final study found no differences between varenicline and placebo treatment with respect to changes in demand, with both groups showing similar magnitude increases in demand elasticity and decreases in demand intensity over a four-week trial (Schlienz et al. 2014). These discrepant outcomes could be due to differences in analytic strategies, attrition rates, and study setting. Further tests directly manipulating these parameters in larger samples are necessary to test these possibilities.

Pharmacological Effects on Cocaine Demand

To date, the only human laboratory study evaluating pharmacological effects on cocaine demand tested the safety and tolerability of acute intranasal cocaine (0 mg to 80 mg) during maintenance on a range of doses of phendimetrazine (0 mg to 210 mg/day), a weak monoamine releaser and prodrug for the more potent monoamine releaser phenmetrazine (Stoops et al. 2016). Although phendimetrazine was safe and tolerable when combined with cocaine, this putative pharmacotherapy did not alter cocaine demand on a cocaine purchase task. Future human laboratory studies and clinical trials will be important for establishing the predictive validity of the cocaine purchase task for assessing therapeutic efficacy.

Consistent with its popularity in the broader animal self-administration literature, cocaine demand has received a great deal of attention in the animal laboratory. This research has provided evidence for the influence of acute drug pretreatments on demand intensity and elasticity, with the underlying goal of revealing behavioral mechanisms by which putative therapeutics might mediate beneficial effects. For example, one study found that acute treatment with haloperidol increased drug demand intensity but decreased P_{max} values, consistent with haloperidol's antagonist effects at dopamine receptors (Oleson et al. 2011). Treatment with other pharmacological agents, including the serotonin-reuptake inhibitor fluoxetine and $GABA_B$ agonist baclofen, also

decreased P_{max} in that study. In contrast, *d*-amphetamine dose-dependently increased P_{max} but did not affect demand intensity. Other investigators have examined novel pharmacological targets that could modulate cocaine demand. One novel target is the trace amine-associated receptor (TAAR) 1, implicated in modulating dopaminergic and glutamatergic activity (Thorn et al. 2014). Acute treatment with the TAAR-1 agonist RO50263397 increased cocaine demand elasticity, suggestive of a therapeutic effect at higher unit prices (Thorn et al. 2014). Acute oxytocin treatments have also demonstrated potential efficacy for treating cocaine use disorder by reducing demand intensity and increasing demand elasticity in rodents (Bentzley et al. 2014). In another study from that laboratory, transient inactivation of the subthalamic nucleus with the $GABA_A$ agonist muscimol did not alter demand intensity but produced a large increase in demand elasticity (Bentzley and Aston-Jones 2017).

Pharmacological Effects on Opioid Demand

Only one study has examined pharmacological variables influencing demand for opioids in the human laboratory. In that study, individuals with an opioid use disorder that were maintained on buprenorphine were treated with the noradrenergic autoreceptor antagonist yohimbine (Greenwald et al. 2013). Yohimbine decreased elasticity of demand for hydromorphone but did not affect demand intensity. This finding is consistent with the notion that stress, in this case a pharmacologically mediated stressor, increases drug demand (see *Stress* section below for more details).

Several studies have examined demand for μ opioid agonists in animal models. For example, no changes in heroin demand were observed following sub-chronic tetrahydrocannabinoltreatment (THC) (Solinas et al. 2004). In another study, acute treatment with the orexin-1 receptor antagonist SB-334867 reduced demand intensity and increased demand elasticity for the short-acting mu opioid remifentanil (Porter-Stransky et al. 2015). Treatment with a morphine-conjugate vaccine also increased

heroin demand intensity in another study indicative of an antagonist effect (Raleigh et al. 2014).

Behavioral Economic Demand in Vulnerable Populations

The purchase task procedure has also proved valuable for studying drug valuation and variations in this valuation for a variety of vulnerable populations. For example, Higgins and colleagues (2017b) demonstrated the validity of the purchase task procedure for estimating cigarette demand among pregnant women. This study found that cigarette demand was also prospectively predictive of the likelihood of making a quit attempt during pregnancy supporting this measures clinical utility. A similar experiment successfully applied the purchase task procedure to the study of very-low nicotine content cigarette demand in three vulnerable populations (i.e., women of reproductive age, opioid-dependent individuals, and individuals with affective disorders) (Higgins et al. 2017a). Another study found that in a community recruited sample cigarette demand was higher among individuals reporting symptoms of past-year psychopathology (e.g., emotional disorder) (Farris et al. 2017). These findings indicate the flexibility and sensitivity of the purchase task procedure in populations of varying health backgrounds and those with contraindications to traditional drug self-administration procedures (e.g., pregnant women).

Impulsivity and Demand

Several studies have evaluated the relationship between impulsivity and measures of drug demand intensity and elasticity. The premise for these tests is an extensive body of research demonstrating a connection between impulsivity and its underlying constructs (e.g., sensation seeking, delay discounting) with drug use (see review by Bardo et al. 2013). In the human laboratory, most studies have examined the relationship between impulsivity and measures of alcohol demand. With respect to alcohol demand, several studies have shown increased demand persistence (e.g., higher breakpoints) in

individuals with higher levels of self-reported negative urgency and sensation seeking (Amlung et al. 2013; Gray and MacKillop 2014; Skidmore and Murphy 2011). These and other studies have also shown that measures of demand intensity are positively related to these facets of impulsivity (Amlung et al. 2013; Gray and MacKillop 2014; Smith et al. 2010). In one such study, demand intensity and maximum expenditure (i.e., O_{max}) moderated the relationship between measures of impulsivity (i.e., negative urgency and sensation-seeking) and drinks per week (Smith et al. 2010). Specifically, greater alcohol demand predicted a stronger positive association between impulsivity and weekly alcohol consumption. Fewer studies have examined the relationship between behavioral measures of impulsivity (e.g., delay discounting) and demand. However, the available literature suggests a positive relationship between discounting and alcohol demand intensity, but not elasticity (MacKillop et al. 2010a). Another study failed to find a significant relationship between monetary delay discounting and cannabis demand, however did show that each uniquely associated with cannabis use severity and frequency, respectively (Aston et al. 2016).

Stress and Demand

One study evaluated the effects of acute stress induction on alcohol demand and found increases in multiple measures, including intensity, breakpoint, and maximum expenditure (Amlung and MacKillop 2014). This finding was replicated in a later study that used a personalized stress manipulation to increase alcohol demand intensity and decrease elasticity relative to neutral mood induction (Owens et al. 2015). The relationship between stress induction and alcohol expenditure in that study was also moderated by a genetic polymorphism in the gene for corticotrophin releasing hormone-binding protein (*CRH-BP*). Specifically, a subset of individuals defined by their genotype at this locus showed a greater increase in alcohol expenditure following stress induction than those with the alternative genotypes. This finding is notable given a previously

demonstrated role of *CRH-BP* in modulating the relationship between stress and drug intake in animal models (Wang et al. 2005; Wang et al. 2007). Taken together, the outcomes of these studies are consistent with the larger body of research implicating stress and negative mood in alcohol use and misuse (Brown et al. 1995; Levy 2008; Ramo and Brown 2008).

Drug-Related Cues and Demand

The impact of drug-related cues on drug demand is grounded in the cue reactivity and incentive motivational literature. These theories posit that repeated drug use sensitizes pathways associated with the attribution of salience, motivation, and reward and that with repeated associative pairing of drugs and drug-related cues, incentive salience and motivation transfers to these drug-paired stimuli (Robinson and Berridge 1993).

Studies in the human laboratory have similarly investigated the impact of drug-related cues on alcohol and cigarette demand. For example, several studies have demonstrated increased demand intensity and decreased demand elasticity for alcohol following alcohol cue presentation (Amlung and MacKillop 2014; Hochster et al. 2018; MacKillop et al. 2010b; but see Amlung et al. 2012). Consistent effects have been observed for smoking cues with respect to cigarette demand (Acker and MacKillop 2013; MacKillop et al. 2012). In the first of these studies, exposure to tobacco-related cues reduced demand elasticity for cigarettes (MacKillop et al. 2012). In a later study using virtual reality to present smoking cues, a decrease in demand elasticity was replicated as well as an increase in demand intensity observed (Acker and MacKillop 2013). The only study to evaluate cue-reactivity and cannabis demand found increases in demand intensity and decreases in demand elasticity following cue exposure (i.e., handling cannabis cigarettes) (Metrik et al. 2016).

Similar support for the relevance of cues is observed in the animal laboratory. One study, for example, examined the effects of audiovisual cues associated with cocaine self-administration on cocaine demand (Bentzley and Aston-Jones 2015). Removal of the light and tone cues significantly increased demand elasticity while having no effect on demand intensity. This finding suggests that the presence (or absence) of drug-related cues can modulate drug demand in animal models.

Environmental Influences: Contingencies and Concurrent Commodities

Manipulating environmental contingencies also alters drug demand. For example, Skidmore and Murphy (2011) examined the effects of next-day responsibilities (e.g., tests, classes) on alcohol demand in a college-aged, non-clinical sample. Next-day responsibilities produced robust increases in alcohol demand elasticity and decreases in intensity. These effects have been replicated in other studies in which similar reductions in alcohol demand are observed with increases in next-day responsibilities among college students (Berman and Martinetti 2017; Gentile et al. 2012; Murphy et al. 2014). Another study evaluated the effects of punishment and alternative reinforcers (i.e., money) on hydromorphone demand in heroin-dependent, buprenorphine stabilized participants (Greenwald 2010). In that study, punishment (i.e., loss of money) and access to an alternative reinforcer increased demand elasticity for hydromorphone. As would be expected from learning theory and choice mechanisms, these findings suggest that the availability and nature of other contingencies in the environment play a crucial role in determining drug demand.

Recent research has also utilized cross-commodity procedures to evaluate the impact of concurrently available, alternative reinforcer price on demand and cross-price elasticity. These cross-commodity tasks generally present participants with a situation in which the price of one commodity (e.g., cigarettes) is manipulated while the price of a concurrently available commodity (e.g., e-cigarettes) is held constant. Several studies

have evaluated the cross-commodity relationship between e-cigarettes and other nicotine-containing products (Johnson et al. 2017b; Snider et al. 2017; Stein et al. 2018a). One study, for example, showed that e-cigarettes could serve as a superior substitute for tobacco cigarettes as compared to nicotine gum (Johnson et al. 2017b). Another study observed similar results wherein e-cigarettes functioned as substitutes for tobacco cigarettes and that the magnitude of this substitution was related to e-cigarette history (i.e., individuals with a more extensive history of e-cigarette use showed greater substitution) (Snider et al. 2017).

Two studies have evaluated cross-price elasticity related to cannabis use. One of these studies found no evidence that cannabis and tobacco substituted or complemented one another (Peters et al. 2017). Another study evaluated demand for cannabis from illicit (i.e., a dealer) and licit (i.e., a dispensary) sources (Amlung et al. 2019). Legal cannabis was considered a superior product as well as a better substitute for illicit cannabis products. These studies collectively demonstrate the sensitivity of demand to concurrently available reinforcers and relevance of evaluating cross-price elasticity for drug commodities.

Summary and Dissertation Aims

This introduction has reviewed the literature on the application of consumer demand theory to understanding behavioral mechanisms underlying substance use. Recent advances in the purchase task methodology have led to a proliferation of research on drug demand in human participants. Applying economic demand functions to drug-taking behavior in this way presents the promise of uncovering novel mechanisms by which drug use persists in the face of numerous negative social, economic, and health consequences. This literature provides ample and consistent evidence that drugs of abuse conform to basic principles of economic demand; namely, that demand decreases with increases in unit price and that demand is adequately explained by mathematical

functions. These studies have also provided evidence for the reliability and validity of the alcohol and cigarette purchase tasks and expected relationships with known pharmacological and behavioral moderators of substance use and misuse (e.g., stress, drug-related cues).

More work is needed, however, to understand significant gaps in this literature, including the more extensive study of illicit substance use and the predictive and incremental clinical utility of these procedures. The majority of research has focused on alcohol and cigarette purchase and less is known about the psychometric properties of these procedures for evaluating illicit substance use. Similarly, many of the existing studies evaluating task psychometrics have relied upon college student samples or those in a college community making generalizations to a broader population difficult. Predictive and clinical utility also remains a significant limitation given the lack of longitudinal research and challenges to the incremental significance of alcohol and other drug demand.

The overarching framework for this dissertation project is the utilization of behavioral economic demand in basic and applied research on substance use disorder. This framework will be leveraged to advance prior research on behavioral economic demand in two key ways. First, this dissertation will advance prior research by further demonstrating psychometric strength and utility in the prospective prediction of substance use. Second, these studies will extend the purchase task methodology to less commonly evaluated substances and further demonstrate the utility of this procedure across varied drug classes. These advances will be addressed in four aims across five experiments. These aims are designed to programmatically evaluate the application of behavioral economic demand in addiction science research from the level of theory development to the level of clinical application. As such, these aims will collectively describe and demonstrate the means by which behavioral economic demand may

advance existing theories of substance use and provide predictive information about substance use within and outside an intervention context.

All experiments will also utilize the online crowdsourcing platform Amazon Mechanical Turk (mTurk). mTurk is an emerging research tool that allows for the effective and efficient sampling of large numbers of diverse research participants. The second chapter of this dissertation provides a comprehensive overview of the existing literature using mTurk in addiction science research. This review will therefore provide the setting by which mTurk is used in the remainder of the dissertation research.

The following presents a brief overview of each aim and experiment(s) designed to achieve that aim.

The first aim of this dissertation is to describe the contribution of behavioral economic demand to addiction science theory. This aim will be achieved through the conduct of Experiment 1 that tests the unique relationship between demand and cannabis use severity and dependence. Individuals reporting cannabis use and controls will complete a battery of behavioral economic and substance use measures, including cannabis demand and delay discounting. The hypotheses are: 1) that cannabis behavioral economic demand will uniquely predict frequency and quantity of cannabis use and 2) that cannabis delay discounting will uniquely predict cannabis use severity. These outcomes would support existing reinforcer pathology models predicting a unique role for demand and discounting in aspects of substance use disorder (see review by Bickel et al. 2017).

The second aim of this dissertation is to provide novel assessments of the psychometric properties underlying the purchase task procedure. This aim will be primarily accomplished through the conduct of Experiments 2 and 3, although aspects of other experiments will also address issues related to psychometrics.

Experiment 2 will demonstrate the stimulus-selectivity of commodity purchase tasks when evaluating behavioral economic demand for drug commodities. As reviewed above, several studies have evaluated the basic psychometric properties of alcohol and cigarette purchase tasks for evaluating demand (e.g., test-retest reliability). However, no studies have systematically evaluated the stimulus-selectivity of purchase tasks to demonstrate that demand metrics are specific to valuation of the commodity under study. Participants will complete alcohol and soda purchase tasks (Experiment 2a) or cigarette and chocolate purchase tasks (Experiment 2b) and demand metrics compared to self-reported use behaviors. The hypotheses are 1) demand outcomes will closely associate with commodity-similar variables (e.g., alcohol demand to weekly alcohol use) and 2) demand outcomes will not closely associate with commodity-different variables (e.g., alcohol demand to weekly soda use). These findings will support stimulus-selectivity by showing that the commodity under study is the primary determinant of the behavioral allocation and choice.

Experiment 3 will develop, refine, and comprehensively validate a battery of behavioral economic measures relevant to prescription opioid use. The majority of research on behavioral economic demand has focused on alcohol, cigarettes, and, to a lesser extent, cannabis. Little research has evaluated behavioral economic demand for other illicit substances, broadly, and non-medical use of prescription drugs, specifically. This study will determine the utility of a variety of purchase task procedures (e.g., single commodity and cross-commodity demand) for use in prescription opioid research. Participants reporting non-medical prescription opioid use will complete this battery of measures at two times separated by one month to establish temporal reliability. The hypotheses are: 1) behavioral economic measures of prescription opioid use will show good construct validity and 2) measures will be test-retest reliable.

The third aim of this dissertation is to establish the predictive and incremental validity of behavioral economic demand for describing prospectively measured substance use. This aim will be primarily accomplished through the conduct of Experiment 4. Experiment 4 will evaluate the relationship between longitudinal patterns of alcohol consumption and the behavioral economic measures of alcohol demand, delay discounting, and proportionate alcohol-related reinforcement. Participants reporting past week alcohol use will complete an 18-week longitudinal study in which daily alcohol use is reported during weekly assessments. Behavioral economic demand will be collected at baseline and post-study surveys to further establish temporal reliability. The hypotheses are: 1) behavioral economic demand will uniquely predict alcohol consumption above and beyond other behavioral economic measures (e.g., delay discounting) and standard alcohol use measures (e.g., AUDIT) and 2) behavioral economic demand will be test-retest reliable.

The fourth aim of this dissertation is to demonstrate the utility of behavioral economic demand for interventions development research. This aim will be accomplished through the conduct of Experiment 5. Experiment 5 will evaluate two forms of Internet-delivered training, working memory training and inhibitory control training, for reducing alcohol consumption. Participants will be randomized to one of these two training conditions or a control group and complete daily training sessions for two weeks. Alcohol consumption prior to, during, and following the intervention period will be measured. A battery of measures evaluating behavioral economic demand for alcohol and other goods will also be collected at baseline and post-intervention follow-ups. Behavioral economic measures will be tested as prognostic indicators of and surrogate measures for intervention success. The hypotheses are: 1) behavioral economic demand will function as a prognostic variable related to reductions in alcohol

consumption and 2) changes in behavioral economic demand over the intervention period will function as a surrogate measure for changes in alcohol consumption.

The remaining chapters will reflect manuscripts from the proposed dissertation experiments that are published, in preparation, or under review. This presentation will result in some repetition of the literature reviewed above as well as elsewhere within the introductions of these manuscripts. Similarly, descriptions of the methodological procedures will be described for each chapter as in the submitted or accepted manuscript source.

A general conclusion chapter will follow that summarizes the collective results of these studies and critical directions for future work.

Chapter 2

THE USE OF CROWDSOURCING IN ADDICTION SCIENCE RESEARCH: AMAZON MECHANICAL TURK

(Strickland and Stoops 2019)

Introduction

Human laboratory, clinical trial, community intervention, and epidemiological approaches have traditionally guided the conduct of addiction science research with human participants. Studies from these perspectives have provided insights into basic science (e.g., mechanism or theory development) and applied science (e.g., interventions development) questions relevant for substance use and misuse. Nevertheless, the sampling procedures typically used for these forms of research present well-documented and persistent challenges. Participant recruitment and retention are notorious problems in human laboratory studies and clinical trials (e.g., Backinger et al. 2008; Del Boca and Darkes 2007; Gul and Ali 2010; Hansen et al. 1985; Howard and Beckwith 1996). Slow participant accrual also means that it is often challenging for human work to keep pace with that conducted in the animal laboratory, thereby making translational and collaborative research difficult. Even after devoting extensive financial resources and time to recruiting participants, small samples can lead to underpowered studies that lack the number of participants needed for appropriate statistical comparisons (Button et al. 2013; Ioannidis 2005). Problems with geographic and/or demographic homogeneity and subsequently reduced generalizability may also result from sampling that occurs at a single or limited number of sites (e.g., one addiction clinic or university) (Del Boca and Darkes 2007). Although these concerns are not unique to addiction science, such problems are often compounded when working with hard-to-reach populations, such as individuals reporting illicit substance use or those with specific behavioral histories (e.g., injection drug use).

An emerging sampling methodology positioned to supplement existing research practice, as well as to advance current methods, is crowdsourcing. Crowdsourcing refers to the completion of tasks through a flexible, open call to a large number of people (Estelles-Arolas and Gonzalez-Ladron-De-Guevara 2012; Howe 2006). The last decade has witnessed the development and refinement of open Internet crowdsourcing markets, one popular source being Amazon Mechanical Turk (mTurk; also commonly abbreviated as AMT, MTurk, or MTURK) (Bohannon 2016). This period has also observed a dramatic growth in the use of mTurk to conduct research in psychological and other health sciences. In fact, the number of manuscripts indexed in PsycINFO under the terms “Mechanical Turk” or “mTurk” increased nearly 4-fold in the 4-year span from 2014 to 2017 (Figure 2.1).

mTurk’s “Internet laboratory” presents numerous strengths, such as the rapid and cost-effective sampling of diverse and potentially hard-to-reach participants, that may help to offset limitations related to traditional sampling methods. The primary objective of this review is to describe the utility of using crowdsourcing and mTurk for research relevant to addiction science. This objective will be accomplished by first reviewing the historical context of crowdsourcing that led to its current use in academic research. Next, evidence supporting the validity of mTurk for clinical and behavioral science, broadly, and addiction science, specifically, will be examined. A summary of existing mTurk studies evaluating substance use and misuse will then be provided to highlight the realized and potential applications of mTurk for addiction science researchers. Finally, best practice recommendations for the conduct of crowdsourced research are offered as well as remaining questions that future research will be well positioned to address.

The Historical Context of Crowdsourcing

The phrase crowdsourcing may be traced to 2006 when *Wired* Editors Jeff Howe and Mark Robinson coined the term referring to the use of the Internet to “outsource work to

the crowd" (Howe 2006). Multiple definitions have been presented since, but all share a common idea of creating an open call to the public in order to solve a specific problem (Estelles-Arolas and Gonzalez-Ladron-De-Guevara 2012). This application of crowdsourcing is present in varied aspects of personal and professional life. For example, Wikipedia can be considered one of the most successful crowdsourced projects, wherein the efforts of many individuals were (and still are) relied upon for the curation of online encyclopedia articles. A particularly compelling example from the biomedical community was the solution in less than three weeks of the protein structure of a retroviral protease that had remained unsolved by scientists for over a decade (Khatib et al. 2011). Lofty goals, such as the creation of a free online encyclopedia or answers to otherwise boggling scientific problems, may be accomplished with the division and aggregation of responsibilities through crowdsourcing.

One popular crowdsourcing option to emerge in the past decade is mTurk (Amazon 2018). Amazon initially developed mTurk as an online labor market that allowed businesses to outsource problems to a human workforce. This idea was inspired by a need to complete simple tasks and other problems that computers are unable to accomplish, are inefficient and error-prone at solving, or are able to do only after extensive and/or complex coding (e.g., transcribing receipts, categorizing items). This "human machine" was designed to effectively and efficiently complete these problems, akin to the origin of the name Mechanical Turk, an 18th century chess-playing machine that was covertly operated by a human chess master inside the automaton (Amazon 2018; Morrison and Morrison 1997).

Tasks on mTurk are created by *requesters* and presented as *Human Intelligence Tasks* (HITs) that *workers* can complete (see Table 2.1 for commonly used terms in crowdsourcing/mTurk work and their academic research analogs). This work is incentivized by compensation for each HIT completed. Amazon also collects fees from

the requester as a percentage of this wage to help maintain the service as well as generate a profit from it (currently 40% that is paid for by the requester, not taken out of the worker's earnings). Launched in November 2005, mTurk has rapidly grown over the past decade and Amazon now boasts over 500,000 users from 190 countries (Amazon 2018).

The completion of these simple tasks remained mTurk's primary use in the years following its launch. However, psychological scientists soon realized the practical benefits for generating convenience samples afforded by mTurk (Mason and Suri 2012). Early adopters in the research community drew clear parallels between the sampling pool available on mTurk and undergraduate psychology participant pools that are often used for convenience sampling. These individuals argued that, unlike psychology participant pools, mTurk provided a sample with greater demographic and geographic diversity and potential improvement upon the often used W.E.I.R.D. samples (i.e., Western, Educated, Industrialized, Rich, and Democratic samples) (Henrich et al. 2010; Landers and Behrend 2015; Mason and Suri 2012). This rationale combined with a rapid rate of data collection at relatively low cost has helped motivate the spread of mTurk through scientific disciplines.

Some of the first studies using mTurk for research purposes belonged to the cognitive and industrial/organizational psychology literatures (e.g., Crump et al. 2013; Keith et al. 2017). Personality, clinical, and social psychologists soon also adopted mTurk as a sampling tool (e.g., Chandler and Shapiro 2016; Miller et al. 2017). More recent years have seen a spread of mTurk to widely varying fields, such as education research (Follmer et al. 2017), cancer biology (Lee et al. 2017), and theoretical biology (Rand 2012). Although the type and purpose of research may differ by discipline, common benefits such as enhanced participant diversity, reduced cost, and improved rates of data collection are often cited as motivating factors behind using mTurk.

Evaluating the Validity of mTurk Samples

The following section provides an overview of seminal work evaluating the validity of mTurk for psychological and addiction research. Several excellent reviews have recently addressed aspects of these and related issues (Chandler and Shapiro 2016; Keith et al. 2017; Woods et al. 2015). Therefore, rather than provide a comprehensive review of this large (and rapidly expanding) literature, we instead highlight representative publications and those directly relevant for addiction science research. All original data presented in this review (i.e., Figure 2.2) were collected under protocols approved by the University of Kentucky Institutional Review Board (IRB #15-1110 "*Using Online Sampling to Examine Population Data for Cognitive-Behavioral Tasks*").

Demographics and Survey-Taking Behavior of mTurk Participants

mTurk is a form of non-probability convenience sampling that results in samples with a demographic composition that differs in some ways from nationally representative probability samples (Chandler and Shapiro 2016; Landers and Behrend 2015). Several studies have attempted to capture how mTurk samples may systematically deviate from the demographic characteristics of the United States population. The primary findings of this research are that mTurk samples tend to be younger, more educated, less religious, and more liberal as well as less likely to be married, a racial minority, or fully employed (e.g., Berinsky et al. 2012; Huff and Tingley 2015; Paolacci and Chandler 2014). Other research has demonstrated that samples may depart from nationally representative sources with respect to health behaviors, for example reporting lower rates of influenza vaccination, asthma, and exercise and higher rates of depression (Walters et al. 2018).

The ideal samples for generalizable research are probability samples that are representative of the United State population. However, this is seldom, if ever, achieved outside of large-scale (and expensive) national survey data. The appropriate point of comparison for demographic representativeness is then likely comparisons with other

conventionally used and viable convenience sampling methods (for more discussion of this issue see Landers and Behrend 2015). Several studies have compared mTurk samples with traditional convenience samples to show equivalence and, in some cases, superiority. For example, one study found that mTurk samples were more representative of the United States population than college student samples or those drawn from college towns, at least for the purposes of political science research (Berinsky et al. 2012). Another study in political science demonstrated similarities in occupational and geographic characteristics between an mTurk sample and the Cooperative Congressional Election Survey (a nationally stratified survey of United States adults) and found that demographic correspondence improved in younger cohorts (Huff and Tingley 2015).

Other studies have demonstrated similarities in participant responding across sampling methods. For example, one study found statistical equivalence in stress and sleep measures collected on mTurk and in a college sample (Briones and Benham 2017). Another study found some statistically significant differences on an emotion classification task between samples drawn from mTurk and those drawn from college campuses or online forums (Bartneck et al. 2015). However, it was argued that similarities in the distribution of responding and the relatively small magnitude effect of these differences observed meant that any deviations were unlikely to be practically meaningful. Another study found that self-admission of previous problematic responding (e.g., responding in socially acceptable ways, to “help” the researcher, or without paying attention) did not systematically differ between mTurk, community, and college samples (Necka et al. 2016). In fact, some research suggests that mTurk samples pay *more attention* to the tasks at hand perhaps because of the extensive experience with and expectation of attention checks in the mTurk community (Hauser and Schwarz 2016, ; see further discussion of attention and validity checks below). mTurk has also shown

some superiority in direct comparisons with other online recruitment methods, such as Facebook or email Listservs, with one study in family science showing improved demographic diversity at lower cost and higher speed of collection for the mTurk sampling method (Dworkin et al. 2016).

mTurk participants in some studies have reported slightly higher rates of substance use than those recorded in nationally representative studies (e.g., the National Survey on Drug Use and Health [NSDUH]). For example, data collected by our group for mTurk studies conducted in 2016 indicated higher rates of lifetime illicit drug use among mTurk participants (N = 5269) than those observed for data collected in a nationally representative sample during the same time period (e.g., 61.5% reporting lifetime cannabis use on mTurk versus 47% nationally; Figure 2.2). These data are consistent with other studies indicating higher rates of recent illicit drug use reported by mTurk participants. For example, one study conducted in 2015 found rates of past month cocaine use (4.3%) that exceeded estimates from nationally representative sources collected in the same year (~0.8%) (Strickland and Stoops 2015; but see Caulkins et al. 2015a; Caulkins et al. 2015b, for concerns about the conservative nature of national estimates). Results from another study reported similar high rates of recent cannabis use among mTurk participants relative to the general population in 2013 (10.6% versus 7.6%) (Shapiro et al. 2013). These estimates should be taken with caution given the problems associated with generalizing point estimates from non-probability sampling methods such as mTurk. However, the reported rates of illicit drug use on mTurk do provide evidence that individuals with varying substance use histories may be sampled through the platform. This argument is bolstered by research studies (described in greater detail below) evaluating participants reporting substance use across many drugs and drug classes, including alcohol, cannabis, cigarettes, e-cigarettes, cocaine, hallucinogens, heroin, methamphetamine, and prescription opioids (e.g., Dunn et al.

2016a; Dunn et al. 2016b; Johnson et al. 2015; Koffarnus et al. 2015; Mellis et al. 2017; Peters et al. 2017; Rass et al. 2015a; Strickland and Stoops 2015).

Scale Psychometrics

Other studies have supported the validity of mTurk data collection by demonstrating scale reliabilities and factor structures of common psychological scales that are consistent with traditional sampling methods (e.g., Behrend et al. 2011; Buhrmester et al. 2011; Feitosa et al. 2015; Kim and Hodgins 2017; Shapiro et al. 2013). For example, personality researchers have observed a reliable five-factor solution and strong internal consistency (e.g., Cronbach's $\alpha > .80$) for the Big Five Inventory on mTurk as typically recorded in laboratory and clinical samples (Behrend et al. 2011; Feitosa et al. 2015). Strong test-retest reliabilities ($r_{xx} > .80$) have also been described for the Big Five Inventory and other widely used personality and clinical measures, such as the Beck Depression Inventory and Brief Experiential Avoidance Questionnaire (e.g., Buhrmester et al. 2011; Shapiro et al. 2013). Similarly, high rates of consistency (> 95%) have been observed for demographic measures taken at multiple survey locations or over multiple measurement periods (e.g., Mason and Suri 2012; Rand 2012).

Recent data also indicate the reliability and validity of common substance use scales when used on mTurk. Some of the most convincing and comprehensive evidence comes from a recent study in participants with a history of alcohol use, cannabis use, or problematic gambling (Kim and Hodgins 2017). Participants in that study completed a battery of standardized measures (e.g., the Alcohol Use Disorder Identification Test [AUDIT]) at two time periods separated by one week. High internal consistency, test-retest reliability, and stability of diagnostic categories were observed for most scales over this one-week period. These rates were also comparable to those observed in other laboratory-based research with the exception that internal consistency (Cronbach's $\alpha = .75$) and test-retest values (ICC = .72) were slightly lower for the WHO-ASSIST in

cannabis users. Participants also reported that they found it easier to answer honestly about sensitive questions on mTurk than in an interview setting (mean rating of 6 on a 7-point scale [Strongly Disagree-Strongly Agree]). Additional research is needed to test and confirm the reliability and validity of other common measures in these and other substance-using populations. However, the results of this study provide promising support for the use of common substance use measures on mTurk.

Demonstration of Common Psychological Phenomena

Other studies have evaluated the validity of mTurk data collection by examining widely documented psychological phenomena in the online setting. The premise for these studies is that similar effects and effect size estimates should be observed online as in the laboratory setting thereby supporting the fidelity of mTurk for psychological research. For example, Crump and colleagues (2013) successfully replicated a variety of common experimental psychology outcomes, such as the Stroop effect (i.e., reaction time interference with incongruent stimuli pairs) and the Simon effect (i.e. faster reaction times when stimuli are spatially congruent). Failures to replicate other effects (e.g., masked priming using a short prime durations) were attributed to concerns related to technology, like lack of control over browser-based display properties, rather than problems specific to the mTurk participant pool. Similar results were observed in another study wherein open-source software was used to replicate classic psycholinguistic effects (e.g., filler-gap dependency processing) that were dependent on small differences in response time and precise response time estimates (Enochson and Culbertson 2015). A particularly compelling study conducted by Mullinix and colleagues (2015) compared 20 different political science experiments when evaluated on mTurk with effect sizes observed in a nationally representative population-based sample. Support for mTurk was found with 80.6% of effect sizes estimates (29 of 36) replicating on mTurk.

Particularly relevant to addiction science is a growing literature replicating and extending findings related to delay and probability discounting using the mTurk platform. When presented with choices that differ in delay, probability, and amount, individuals must weigh the relative benefits of such outcomes. The discounting of delayed monetary gains refers to the acceptance of a smaller, sooner reward (e.g., $500 now) over a later, larger reward (e.g., $1000 in three months) (Odum 2011). Alternatively, in probability discounting, the value of a reward is reduced as a function of the odds against receiving that reward. Discounting, and delay discounting in particular, has received extensive attention in theoretical accounts of substance use and it has been argued that excessive delay discounting represents a trans-disease process contributing to disease-related vulnerability (Bickel et al. 2012).

Several mTurk studies have emphasized discounting processes, in part due to an interest in episodic future thinking (EFT) within interventions development research (see *Interventions Development* section below). One large sample study replicated the well-described effect of higher delay discounting rates in smokers compared to non-smokers (Jarmolowicz et al. 2012). This study also found no differences in probability discounting between these groups, which was also consistent with prior literature and indicated that the observed effects were behaviorally specific when tested in the online platform. Another study sought to replicate six well-described effects in the discounting literature within the context of a novel question related to opportunity cost (Johnson et al. 2015). This attempt was largely successful and replicated at least five of six effects, such as the magnitude effect (i.e., steeper discounting for smaller delayed rewards than larger delayed rewards) and steeper discounting of consumable goods relative to money. An effect of smoking status was not observed in that study. However, this apparent discrepancy with prior literature was attributed to procedural differences between studies

and the use of a delay task with a maximal 24-hour delay in which differences between smokers and non-smokers had not been previously observed.

Taken together, the extant literature provides ample evidence for the reliability and validity of data obtained on mTurk. Although demographic characteristics may differ in some ways from nationally representative samples, these discrepancies are well documented and show some improvements over alternative convenience sampling methods. Other research has described similar psychometric properties and observed "known" psychological effects on mTurk as in laboratory literature. Fewer validation studies have systematically and specifically evaluated measures and behaviors related to substance use and use disorder. However, those studies that do exist provide initial support for the validity of mTurk for questions relevant to addiction science.

Applications of mTurk in Addiction Science

The application of mTurk in addiction science is a relatively recent development when compared to its uses in other areas of psychological science (Figure 2.1). The following section reviews this emerging literature that utilizes mTurk to answer questions relevant to substance use and misuse. This discussion is a narrative review of existing research rather than a systematic review using PRISMA guidelines and, therefore, is not an exhaustive review of all extant literature. The section is organized so as to focus on four of the broad approaches utilized, to date: 1) cross-sectional research and replication studies, 2) measure development, 3) longitudinal designs, and 4) interventions development.

Cross-Sectional Research and Replication Studies

One of the most popular uses of mTurk for psychological and addiction science is cross-sectional survey and basic cognitive-behavioral research. These studies may be conducted as independent experiments or combined with ongoing laboratory projects as a replication sample. This latter use is a particularly notable strength of online sampling

as it allows for the relatively rapid testing of effects observed in laboratory studies in a new and independent sample. Such replication attempts have become increasingly relevant in psychological sciences given recent challenges regarding reproducibility and related failures to replicate published findings (Nosek et al. 2015).

The ability of online crowdsourcing to accompany typical studies conducted with human participants can help enhance the overall rigor and generalizability of observed results. For example, Athamneh, Stein, & Bickel (2017) found that individuals with higher intentions to quit smoking showed lower delay discounting rates in two cross-sectional cohorts from the human laboratory and mTurk. The application of mTurk in this study was particularly noteworthy as two related, but independent and distinct tasks were used to evaluate discounting in each setting, demonstrating that the observed effects were not methodologically bound. A similar multi-sample approach was used in another study evaluating the relationship between drinking to cope and hazardous drinking in a college psychology pool and a non-college mTurk sample (Veilleux et al. 2014). That study observed a significant relationship between these variables in both samples thereby replicating previous research in college students and extending those findings to adults. Additionally, a novel mediation model explaining the relationship between negative affect intensity and drinking to cope through emotional clarity/strategies was supported in both samples (Veilleux et al. 2014).

mTurk has been widely utilized for the study of behavioral economic demand in substance using populations. A review of this emerging literature serves to provide several case examples of how mTurk may be used to complement in-person studies. Behavioral economic demand represents the orderly relationship between consumption of a good and its price (see reviews by Bickel et al. 2000; Hursh and Roma 2013; MacKillop 2016; Reed et al. 2013). Recent years have witnessed a growth in the human laboratory and clinical study of demand due in part to an increasing utilization of the

hypothetical commodity purchase task. This procedure asks participants to report hypothetical consumption of a good (e.g., alcohol) across a range of prices (e.g., $0.01, $1.00/drink) and is particularly appealing because of its cost and time efficiency as well as adaptability for populations with whom drug self-administration is not ethically or practically feasible (e.g., patients in residential treatment, those with medical contraindications to drug administration) (Jacobs and Bickel 1999; Kaplan et al. 2018). These attributes have also made the purchase task portable to mTurk thereby affording researchers the opportunity to index drug valuation in a remote, online setting.

One large sample study evaluated the validity of administering behavioral economic measures, including the commodity purchase task, on mTurk (Morris et al. 2017). A large sample of alcohol-using participants (N = 865) was recruited on mTurk and completed an alcohol purchase task to measure alcohol demand and a reinforcement survey schedule to measure proportionate alcohol reinforcement. Purchase task data were systematic and provided unique prediction of alcohol use severity supporting the convergent validity of this measure on mTurk, a finding that was consistent with extant laboratory and clinic research (see reviews of alcohol purchase task studies in Kaplan et al. 2018; MacKillop 2016). Another study evaluated the unique prediction of cannabis use by behavioral economic demand and delay discounting (Strickland et al. 2017b). Purchase task data were systematic for cannabis and alcohol commodities with cannabis demand uniquely predicting cannabis quantity-frequency and cannabis delay discounting uniquely predicting cannabis use severity (i.e., cannabis use disorder symptom counts). These findings replicated those observed in a prior laboratory study (Aston et al. 2016) supporting the validity of online data collection and highlighting its utility as a source for replication studies. Other studies have successfully used the purchase task to evaluate demand for alcohol (Kaplan et al. 2017; Noyes and Schlauch 2018) and cannabis (Peters et al. 2017) as well cocaine (Strickland et al. 2016c),

cigarettes (Koffarnus et al. 2015; Snider et al. 2017; Stein et al. 2018b; Strickland and Stoops 2017), and e-cigarettes (Johnson et al. 2017b; Snider et al. 2017) emphasizing the versatility of the online platform for studying varying drug classes. This ease and speed of data collection also allows for the study of parametric manipulations or other aspects of task design that may be overlooked when conducting in-person research in the interest of focusing on clinical applications. For example, research on mTurk has assessed a novel demand equation for purchase task data (Koffarnus et al. 2015), the stimulus-selectivity of the purchase task procedure (Strickland and Stoops 2017), and the influence of variations in task instructions on demand outcomes (Kaplan et al. 2017).

Other studies have leveraged mTurk for cross-commodity purchase task research. Cross-commodity tasks present participants with a situation in which the price of one commodity (e.g., cigarettes) is manipulated while the price of a concurrently available commodity (e.g., e-cigarettes, nicotine gum) is held constant. These procedures provide a measure of the extent to which commodities function as complements (i.e., as the price of one increases, consumption for the other decreases; hot dogs and hot dog buns, for example) or substitutes (i.e., as the price of one increases, consumption for the other increases; Coca Cola® and Pepsi®, for example). Two studies have evaluated the cross-commodity relationship between e-cigarettes and other nicotine-containing products on mTurk (Johnson et al. 2017b; Snider et al. 2017). The first study showed that e-cigarettes might serve as a superior substitute for tobacco cigarettes when compared to nicotine gum (Johnson et al. 2017b). The other study observed similar results wherein e-cigarettes functioned as substitutes for tobacco cigarettes and that the magnitude of this substitution was related to a participant's e-cigarette use history (i.e., individuals with a more extensive history of e-cigarette use reported greater substitution) (Snider et al. 2017). Another study evaluated cannabis and tobacco cigarettes and found no evidence that these commodities substituted for or complemented one another

(Peters et al. 2017). Taken together, this behavioral economic demand literature demonstrates the varied basic science and applied applications of mTurk for cross-section research as well as the ability to study drug valuation within an online context.

Corresponding laboratory and online studies may also be used to test the specificity of laboratory effects by recruiting relevant control groups. This approach was used in a series of experiments evaluating the relative rate of learning from positive and negative outcomes in cocaine users and controls (Strickland et al. 2016a). Cocaine users were first recruited for a laboratory study in which a reduced sensitivity to learning from positive relative to negative outcomes was observed on a probabilistic learning task. These effects were then replicated in an independent mTurk sample and specificity to a cocaine-use history demonstrated by also recruiting an online control sample.

Another apparent benefit of research conducted on mTurk is the ability to screen for and select samples with specific behavioral or health histories. This advantage can be especially useful for emerging trends in substance use whose profile has not yet been established, thereby making targeted community recruitment difficult. In line with this idea, a number of researchers have leveraged mTurk to sample electronic cigarette (e-cigarette) users. Large sample characterization studies have been conducted, such as one evaluating use patterns and perceptions of relative harm in dual e-cigarette and tobacco cigarettes (Rass et al. 2015a). Other studies have evaluated more specific aspects pertaining to this emerging and growing substance use trend, including relationships between tobacco cigarette smoking history and e-cigarette perceptions (Bauhoff et al. 2017), factors related to the use of e-cigarettes in women of reproductive age (Chivers et al. 2016), the effectiveness of advertisements for e-cigarettes as smoking cessation aids (Jo et al. 2018), the development and validation of a vaping craving questionnaire (Dowd et al. 2018), and predictors of using “vape” pens for cannabis administration (Morean et al. 2017).

Similar targeted recruitment strategies have been used to identify individuals with specific behavioral or health histories for which community sampling may yield low participant accrual and difficulties in generating adequately powered samples. For example, HIV+ smokers have been recruited to identify nicotine-related knowledge, perceived health risks of cigarette smoking, and predictors of cessation interests within this particularly vulnerable health group (Pacek et al. 2017a, 2017b). Crowdsourcing has also been used to rapidly recruit larger-than-typical samples of special populations, such as individuals with lifetime psychedelic use (Forstmann and Sagioglou 2017), chronic pain (Tompkins et al. 2016; Tompkins et al. 2017), men who have sex with men (Herrmann et al. 2015), and individuals with illicit drug use histories (Dunn et al. 2016a; Dunn et al. 2016b; Strickland and Stoops 2015). These latter two examples are particularly noteworthy given that snowball sampling is traditionally employed in human laboratory and clinical studies to recruit these populations, which can result in biased observations and greater homogeneity within the resulting samples (Biernacki and Waldorf 1981; Faugier and Sargeant 1997). Although attention to the limitations presented by self-report data is necessary (see further discussion of this issue below), mTurk may be used to access specialized populations for recruitment of larger size and more diverse samples than are typically afforded in community-based research.

Measure Development

Measure development efforts have benefited from using mTurk to rapidly generate large samples with relatively diverse substance use histories. Large samples may be utilized to develop a measure and its initial factor structure (exploratory factor analysis; EFA) or serve as a replication sample to determine the generalizability and factor invariance of a novel measure (confirmatory factor analysis; CFA). This approach has been used to develop measures for a variety of topics relevant to addiction science, such as therapeutic alliance during cigarette-cessation counseling (Warlick et al. 2018),

attitudes toward contraband cigarettes (Adkison et al. 2015), alcohol-related myopia (Lac and Berger 2013), and diagnostic testing for *DSM-5* caffeine use disorder (McGregor and Batis 2016).

A blend of in-person and online samples may also be used for measures development. An elegant pair of studies by Dunn and colleagues (2016a; 2016b) exemplifies this approach. Each study first recruited smaller clinical samples for EFA and then recruited larger and more geographically and clinically diverse samples from mTurk to determine factor invariance and conduct CFA. For example, one study developed a 3-factor Brief Opioid Overdose Knowledge (BOOK) questionnaire by first recruiting illicit opioid users (n = 147) from a single clinic in Baltimore for initial scale development (Dunn et al. 2016b). The internal validity of the scale was then confirmed by recruiting individuals from two independent clinic sites (n = 199) as well as a larger mTurk cohort of chronic pain patients receiving an opioid analgesic (n = 502). The ability to rapidly confirm factor structures and generalizability for novel measures is a clear strength of crowdsourcing for measures development research.

mTurk-assisted measure development has also received considerable attention in the study of behavioral addictions (i.e., non-drug related addictions). Novel food addiction measures, including the Yale Food Addiction Scale Version 2.0 (Gearhardt et al. 2016) and a brief version of this scale (Schulte and Gearhardt 2017), were recently developed on mTurk and subsequently used in cross-sectional research on the platform (Rainey et al. 2018). Similar methodological studies have used mTurk for studying technology-related behavioral addictions, such as mobile phone addiction (Bock et al. 2016; Contractor et al. 2017), social media addiction (Muench et al. 2015), and Internet gaming addiction (Beard et al. 2017). Although the debate over “behavioral addiction” versus the pathologizing of common behavior is beyond the scope of this review (see Kardefelt-Winther et al. 2017, for a relevant discussion of this issue), mTurk is becoming

an increasingly utilized resource for those interested in characterizing non-substance-related addictive disorders.

Longitudinal Research

The unique identifiers assigned to mTurk participants and easy-to-use interface also allow for the conduct of follow up assessments. Such test-retest designs have been used by researchers with diverse interests in psychological science. For example, Daly and Nataraajan (2015) observed response rates of 75% at two months, 56% at four months, and 47% at three months following a baseline assessment across three independent personality psychology experiments. Others have observed similar rates when recording at weeks or months times (Chandler et al. 2015; Shapiro et al. 2013). One particularly noteworthy study used an intensive daily diary approach to evaluate the relationship between electronics use and sleep quality (Lanaj et al. 2014). Some evidence for the feasibility of this approach was observed with response rates of 61% for surveys completed over 10 consecutive workdays in that study.

As noted above (see *Scale Psychometrics* section), Kim and Hodgins (2017) observed test-retest reliabilities for substance use measures similar to those observed for in-person research. One week follow up rates of 87% or greater were observed in that study for alcohol-using, cannabis-using, or problematic gambling participants. Two other studies have used intensive longitudinal methods (i.e., frequent or dense measurement such as daily diary or ecological momentary assessment) to evaluate alcohol consumption (Boynton and Richman 2014; Strickland and Stoops 2018b). In the first study, alcohol consumption was measured over a 14-day period using a daily diary design and findings commonly reported in the literature were observed supporting the validity of the approach (e.g., heavier drinking on the weekend) (Boynton and Richman 2014). Participants (N = 369) also completed 8.5 of the daily measurements (60.7%) on average providing some support for study feasibility. The second study extended these

preliminary findings by collecting weekly recordings of alcohol and soda use over an 18-week period (Strickland and Stoops 2018b). Participants (N = 278) reported that this design was acceptable (i.e., 94% indicated they would participate again). Feasibility was also indicated by an average completion rate of 73% of participants per week over the 18-week period (range: 64.1%-86.8%/week). Construct and external validity were further demonstrated through the replication of expected relationships that were specific to alcohol use and not observed for soda, such as heavier alcohol consumption by individuals with higher AUDIT scores and on weekends. These studies collectively provide preliminary support for the feasibility, acceptability, and validity of conducting longitudinal work with substance-using populations on mTurk.

Interventions Development

mTurk has also been utilized for recent interventions development work. The application of mTurk for interventions purposes may prove particularly useful because of the inherent similarities to Internet-based interventions (see reviews by Andersson and Titov 2014; Carroll and Rounsaville 2010; Dallery et al. 2015; Kurti et al. 2016, on the use of Internet-based interventions in psychology and addiction science). Internet-based interventions provide many benefits for substance use prevention and treatment, such as access to otherwise remote or hard-to-reach populations (e.g., rural and/or adolescent populations; Harris and Reynolds 2015; Reynolds et al. 2015; Stoops et al. 2009). Pilot testing and refinement of such interventions on mTurk is a particularly appealing application of crowdsourcing given the inherent portability for future large scale, Internet-based trials and dissemination.

An emerging body of literature has used mTurk to evaluate the effects and mechanisms underlying anti-smoking health warnings and other mass media messages targeting cigarette use. Some studies, for example, have evaluated the effects of manipulating cigarette packaging (e.g., packet label or the use of iconic images) on

perceptions of harm (e.g., Lazard et al. 2017; Leas et al. 2017; Pearson et al. 2016). One particularly innovative design exposed participants to FDA-proposed textual and pictorial warnings about smoking-related hazards, textual warnings with irrelevant images, or text-only warnings and evaluated cigarette use and feelings about smoking at baseline and a seven-day follow up (Shi et al. 2017). Participants exposed to the FDA-proposed warnings showed greater motivation to quit, fewer reported cigarettes smoked per day at seven-day follow up, and better memory for the warnings than those in the other two conditions. These findings suggested that images communicating smoking-related risk enhanced the persuasiveness of the proposed warnings. Another study employed a pre-post design to evaluate the impact of exposure to smokeless tobacco constituent information on risk and knowledge measures (Borgida et al. 2015). Exposure to information about the carcinogenic constituents of smokeless tobacco improved knowledge about the contribution of these components of tobacco to disease risk and acknowledgement that products may present varied levels of risk (e.g., medicinal nicotine replacement therapies present less risk than smokeless tobacco and cigarette products).

mTurk is also well suited for exploring the efficacy and mechanisms underlying brief interventions. Screening and brief interventions are commonly used in the clinical setting as a “first-line of defense” for prevention and treatment (Pilowsky and Wu 2013). This strategy is consistent with the broader idea of “Screening, Brief Interventions, and Referral to Treatment” or SBIRT (Madras et al. 2009). SBIRT proposes a comprehensive and integrated identification and treatment linkage for individuals at risk for or suffering from a substance use disorder.

Three studies have evaluated brief interventions for alcohol use on mTurk. One study evaluated the feasibility and acceptability of providing online feedback of alcohol use in older adults (50+) via personalized or normative feedback approaches, two brief

interventions with moderate effects on alcohol consumption (Kuerbis et al. 2017). Online feedback was deemed feasible and normative feedback outperformed personalized feedback for motivating changes in drinking patterns. Another study evaluated normative feedback to evaluate underlying mechanisms of change and found tentative support for changes in drinking behavior through a belief in the accuracy of feedback mechanism (Kuerbis et al. 2016). A third study evaluated the effects of personalized feedback intervention ("Check Your Drinking" Cunningham et al. 2009) on alcohol use at a 3-month follow up (Cunningham et al. 2017). High follow-up rates were observed at 3-months (85%). However, reductions in alcohol use with the personalized feedback intervention were only observed for one of four outcome variables (i.e., AUDIT consumption subscale).

An additional study evaluated delivery of a brief opioid overdose education intervention on mTurk (Huhn et al. 2018). Participants reporting prescription opioid use for pain completed two variants of opioid overdose education related to opioid effects, opioid overdose symptoms, and opioid overdose response. Overdose education increased scores on a Brief Opioid Overdose Knowledge measure and these effects did not differ between participants presented with the information and those required to respond until demonstrating mastery. The design was also acceptable with 92.9% reporting they would recommend the intervention to a family member. These findings replicated those previously observed in individuals with opioid use disorder from an outpatient detoxification clinic (Dunn et al. 2017) providing further evidence for the validity of online data collection. Collectively, these studies offer promising preliminary support for the use of crowdsourced samples to pilot novel brief interventions and evaluate mechanisms underlying established ones.

Another potential intervention receiving extensive attention on mTurk is episodic future thinking (EFT). EFT is a form of prospective thought that encourages participants

to think about episodic future events in order to increase the temporal window of thought (Atance and O'Neill 2001). The clinical utility of EFT is believed to lie in this increase in the temporal window and subsequent enhanced value of delayed outcomes and encouragement of choices for long-term rather than short-term benefit. Laboratory measurement has demonstrated effective reductions in delay discounting rate (i.e., greater choice for delayed outcomes) following EFT (see review in Bickel et al. 2017). Recent research has extended these results to show that EFT may translate to an online-delivery method (Stein et al. 2017; Stein et al. 2018b; Sze et al. 2017). In each of these studies, the generation of personal narratives describing positive episodic future events reduced delay discounting rates. These studies correspond with other mTurk research demonstrating that episodic future thinking about negative events can produce the opposing effects with increases in delay discounting rates compared to neutral control narratives (Bickel et al. 2016b; Mellis et al. 2018; Sze et al. 2017). Brief EFT training delivered through an online platform may also produce functional impacts on negative health behaviors with one study demonstrating reductions in demand for fast food in overweight/obese participants (Sze et al. 2017) and another reduced demand for cigarettes in tobacco cigarette smokers (Stein et al. 2018b). Further studies will help in determining the clinical applicability of these findings as well as extensions to other drugs of abuse previously studied in the human laboratory (Bulley and Gullo 2017; Snider et al. 2016).

A particularly elegant set of studies has utilized mTurk to evaluate public opinion concerning interventions targeting substance use. These studies have examined public perspectives surrounding potentially controversial interventions, including the expansion of naloxone access (Rudski 2016), the use of financial incentives to promote drug abstinence (i.e., contingency management) (Wen et al. 2016), and the use of medication assisted therapy (i.e., agonist replacement) (Huhn et al. 2017). One study also included

an experimental manipulation evaluating the effects of educational materials on participants' opinions concerning the use of financial incentives to promote smoking cessation during pregnancy (Wen et al. 2016). Individuals in the intervention group showed increased acceptance of contingency management for pregnant smokers with 90.3% agreeing that "paying pregnant women who smoke to quit smoking is a good idea" compared to only 69.4% in a control group. Low-cost, high impact interventions such as these serve as a simple demonstration of the ability to use online sources to pilot interventions concerning substance use and misuse.

Limitations of mTurk Research

Although mTurk presents numerous strengths for addiction science, there are several limitations to the approach that deserve discussion. One common criticism of research with mTurk samples it that these samples may systematically differ from the populations to which the results ideally would generalize. Several demographic differences have been documented, such as participants being more liberal, younger, and educated in mTurk samples as well as reporting lower rates of employment, marriage, and racial diversity (e.g., Berinsky et al. 2012; Huff and Tingley 2015; Paolacci and Chandler 2014). Some of these deviations are similar to those observed in other forms of convenience sampling (e.g., college student samples or those from a single clinic site) and carry with them the typical concerns related to non-probability convenience sampling. For example, it would be ill advised to suggest that point estimates of interest (e.g., percentage of individuals reporting a specific type of substance use or engaging in a particular form of behavior) observed on mTurk reflect true population estimates. However, it is important to note that generalization is a problem inherent to all non-probability sampling methods and not one that is unique to mTurk research (Landers and Behrend 2015). In fact, the ability to conduct research through many alternative platforms (e.g., in-person community samples, mTurk-

generated samples) could provide more robust support for an experimental or clinical finding of interest than a focus on any one form of sampling alone. Therefore, although researchers should be aware of this generalizability concern when using mTurk samples, these concerns should ultimately be weighed against those of other viable sampling approaches and the benefits that could result from data collection through varied sampling format.

Another concern when conducting research on mTurk is the experience of research participants and potential prior exposure to experimental tasks and procedures (see discussion in Chandler and Shapiro 2016). For example, one study found that the number of prior HITs completed was correlated with performance on standard, but not novel, version of a common cognitive task suggesting that individuals may have been exposed to and potentially learned the correct responses to the standard task variants (Chandler et al. 2014). Similarly, reductions in the effect sizes of certain psychological phenomenon (e.g., the anchoring effect) were observed when the same participants were retested at a time point a day, week, or month from first assessment (Chandler et al. 2015). These findings suggest that exposure to the same or similar version of a task can influence future behavior with that or related procedures. The exact implications for research conducted in addiction science have not yet been explored. However, one way to index an individual's potential familiarity with an experimental protocol is to ask them if they had completed the task or similar variants previously. Alternatively, recording the number of prior HITs completed can provide a general index of a participant's potential familiarity with research protocols and provide an important covariate for planned analyses.

A primary limitation of mTurk for research in addiction science is the inability to biologically verify substance use or to deliver pharmacological manipulations. Although self-report often provides a reliable and valid measure of drug-taking behavior (Elman et

al. 2000; Kokkevi et al. 1997; Napper et al. 2010), it is possible that individuals on mTurk could engage in disingenuous behavior regarding their substance use history for a variety of reasons. For example, exaggeration of current or past substance use may occur so as to qualify for studies one may not otherwise qualify for. Alternatively, individuals may be wary to share health information online and underreport substance use. The use of internal attention and validity checks can help to identify these and other forms of problematic responding (for more details see *Attention and Validity Checks* below). Kim and Hodgins (2017) also recently reported that individuals with a history of alcohol use, cannabis use, or gambling behavior found it easier to answer sensitive questions on mTurk as compared to an in-person or phone interview. We observed similar results in a sample of alcohol-using individuals, with approximately three-quarters of participants reporting that they were more comfortable reporting sensitive material through mTurk than through other sources (Strickland and Stoops 2018a). These results are consistent with other reports indicating that online data collection can help reduce underreporting biases that may occur with stigmatized behavior, such as substance use (Harrison and Hughes 1997; Turner et al. 1998).

Methodological Considerations when Conducting mTurk Research

The practical implementation of mTurk has been discussed elsewhere and we suggest that readers interested in incorporating these techniques in their research also read these peer-reviewed manuscripts (Litman et al. 2017; Mason and Suri 2012; Woods et al. 2015) and other online documentation (e.g., blogs and “how-to” guides). In the following section, we briefly review considerations that are worthy of emphasis and/or those that are particularly relevant for research in addiction science. These sections include general discussion points that researchers will likely need to consider when conducting mTurk research as well as existing empirical research addressing

these topics. We have avoided making specific recommendations given that individual research questions and agendas will likely necessitate varied approaches.

Screening Questionnaires and Qualification Restrictions

Screeners are often necessary to determine study eligibility. These questionnaires typically include questions relevant to substance use quantity, frequency, or severity to determine eligibility based on current or prior behaviors (e.g., drinks alcohol once per week or more, meets criteria for cannabis use disorder). Alternatively, these screeners may identify individuals with specific health or behavioral histories (e.g., HIV+ individuals, those reporting past or recent injection drug use). Screening questionnaires are often designed to be easy and quick to complete as well as to include questions unrelated to substance use to mask the specific items relevant for qualification. This latter aspect is particularly important given the possibility of participants engaging in dishonest behavior in order to qualify for a study. Researchers may elect to pay participants a nominal fee for these short screeners ($0.05/screener) or use them as unpaid questionnaires prior to access to a larger study. Many survey platforms (e.g., REDCap, Qualtrics) have built in features that will prevent repeated survey completion by participants and the worker ID system on mTurk provides an additional mechanism for preventing repeat participation. An example Qualtrics-designed screening questionnaire can be found at (doi:10.17605/OSF.IO/BKDTV).

Also relevant to a discussion of screening questionnaires is the “in-house” qualification system available on mTurk (see Figure 2.3A for example interface). mTurk includes a number of built-in screening methods that allow researchers to filter who can and cannot view the study. These qualifications included filtering by the number of past mTurk HITs completed, the percentage of prior HITs accepted, and geographic location

(state or country). Additional qualifications known as “Premium Qualifications” allow for more specific demographic restrictions determined from the worker profile (e.g., age, employment type, language fluency) and may be used for an additional fee per HIT. Many studies conducted in the United States will elect to restrict to individuals from the United States given between country variations in English language abilities, sociodemographic characteristics, and socio-cultural norms concerning substance use. However, participants may also be sampled from other countries to allow for cross-national comparisons or global generalizability.

The decision to use a HIT number or acceptance criteria is an important discussion point in mTurk research. Specifically, restricting participants based on approval ratings or number of completed tasks is common in mTurk research (e.g., Kaplan et al. 2017; McKerchar and Mazur 2016; Morean et al. 2017; Reed et al. 2016; Strickland and Stoops 2015). The use of restrictions can improve data quality and reduce undesired patterns of responding. Specifically, one study comparing individuals based on number of HITs or acceptance rates found that those of higher “reputation” (i.e., above 95% approval rating) or “productivity” (completed more than 100 HITs) failed fewer attention checks, had higher reliability scores for previously developed measures, and showed decreases in problematic responding (e.g., social desirability or central tendency biases) (Peer et al. 2014). However, the use of restrictions may generate a sample that is systematically different from the expected population. For example, screening restrictions could result in a sample with greater conscientiousness or responsibility that may be systematically related to measures of decision-making or impulsivity typically used in addiction science. Similarly, sampling bias could result in a sample that systematically differs in demographic and substance use variables. A recent study in participants with past year cocaine use, regular cigarette smoking, or no history of cigarette or illicit drug use found few differences between those sampled using no

restrictions and those restricted by a 95% approval on 100 or more HITs criteria (Strickland and Stoops 2018a). These results were consistent with the aforementioned study in which in the age or gender distribution of participants did not differ based on completion or approval rates (Peer et al. 2014). Future research will ultimately benefit from empirical work systematically evaluating the influence of restriction criteria on findings relevant to addiction science.

Attention and Validity Checks

Conducting research on mTurk necessitates some loss of experimental control over the effort and attention provided by participants. A common approach to address this loss of control is using attention and validity checks. These checks are designed to verify that individuals provide due attention throughout participation as well as the fidelity of responding. Simple attention checks may include selecting a particular response for an item (e.g., select “strongly agree”; enter the text “I’m paying attention”) or including a number or phrase midway through a survey that will be required at the end (e.g., remember the number 3, you will be asked for it later). Although these checks are easy to implement and may identify participants engaging in problematic responding, some research suggests that they may be ineffective given their repeated use on the mTurk platform (Hauser and Schwarz 2015; Hauser and Schwarz 2016). mTurk workers outperform college samples on traditional attention checks (Hauser and Schwarz 2016) and other evidence suggests that the overuse of attention checks can alter participants response patterns resulting in more deliberate decision-making on later questions (Hauser and Schwarz 2015).

An alternative to traditional attention checks is the use of unobtrusive means of determining response fidelity. Asking participants to enter demographic information, such as age or sex, in two or more separate locations of the survey can help identify individuals responding dishonestly. The use of repeated questions may be particularly

helpful for the purposes of addiction science in order to verify that the substance use history reported remains consistent (e.g., someone who reports cocaine use on a screener also reports cocaine use in the body of the survey).

Payment

Researchers are required to set the compensation rate for participants prior to recruitment and data collection. However, bonus payments can be made through the platform allowing for task or response contingent payments (see Figure 2.3B for the payment and approval interface). The appropriate rate for payment remains a controversial issue in crowdsourced research (see recent discussions in Chandler and Shapiro 2016; Gleibs 2017; Goodman and Paolacci 2017) and no strict guidelines have been broadly accepted, to date. Determining an ethically appropriate wage is difficult given the necessary balance of providing a fair wage and avoiding undue influence or disingenuous responding due to high compensation relative to community standards. No empirical research to our knowledge has systematically evaluated the impact and empirical ethics of different payment schedules in crowdsourced research. Evaluating such topics will be important for research on mTurk as well as future work in other crowdsourcing platforms. Ultimately, however, it is important that researchers remain transparent with participants and ensure that all expectations (e.g., expected time of completion; expected effort for a task) and incentives (e.g., payment, time to payment) are clearly articulated.

Tools to Supplement mTurk Research

Researchers may elect to supplement research conducted on mTurk with third-party tools or programs. One common example of this approach is using third-party survey programs, such as Qualtrics or REDCap, to develop and administer the study survey. Although mTurk does contain its own survey program, these alternative platforms allow for more complex survey functions and designs while utilizing programs traditionally

applied in research settings. Other mTurk supplements have been developed for integration with mTurk to facilitate research-related activities. The R package MTurkR, for example, provides functions that allow researchers to write simple code to complete a large number of operations, such as bulk creation of HITs or emailing a large number of participants. Turk Prime is a similar resource that has been developed to automate and streamline such tasks through an easy-to-use online tool (Litman et al. 2017). These tools can provide an indispensible resource when completing studies with a large number of participants or when repeated contact is needed (e.g., longitudinal designs).

Other Crowdsourcing Platforms

Although this review has focused on mTurk's application to addiction science, other crowdsourcing platforms exist that can and have been leveraged for research purposes. Several research groups have utilized Facebook and its advertising system to recruit participants for online studies and intervention delivery research (Lee et al. 2016; Ramo and Prochaska 2012; Ramo et al. 2018). The Facebook advertising platform provides a particularly versatile system allowing for targeted delivery based on an individual's demographic characteristics and interests (for a review of using Facebook in online research see Borodovsky et al. 2018). Other forms of social media have been similarly used to sample substance-using individuals and target specific interest groups, for example through postings on relevant online chat rooms and forums (e.g., Reddit, Bluelight) (Carbonaro et al. 2016; Johnson et al. 2017a). Other crowdsourcing opportunities exist within more traditional public opinion and marketing research resources (e.g., Growth from Knowledge (GfK); Qualtrics Panels). Several new platforms, including Prolific (formally Prolific Academic) and Prime Panel, have been developed by members of the research community. Features of these newer resources are directly promoted as responses to limitations presented by mTurk (e.g., decreasing

participant non-naiveté). Such developments suggest an enduring interest in the application of crowdsourcing within the research community as well as efforts towards developing new and improved crowdsourcing systems.

Future Directions for mTurk (and Crowdsourced) Research

Crowdsourcing has witnessed a dramatic growth over the past decade. Researchers in psychological and addiction science have efficiently and effectively used crowdsourcing to sample research participants. Addiction science is particularly well positioned to benefit from crowdsourced sampling given the ability to recruit populations with specific behavioral or health histories. Existing research has supported the reliability and validity of data gathered using crowdsourced samples. Promising research relevant to substance use and misuse has also been conducted, including studies with cross-sectional designs and for measure development as well as more recent studies using longitudinal methods and for interventions development.

Nevertheless, the mTurk literature is still one in its infancy. Additional studies specifically designed to evaluate the reliability and validity of mTurk for addiction science research are needed. Further methodological studies will also provide important information about the constraints of research conducted through online platforms for addressing questions relevant to substance use and misuse. The majority of research conducted to date has used simple cross-sectional or test-retest investigations. Some studies have demonstrated the feasibility, acceptability, and validity of more elaborate designs, such as intensive longitudinal approaches or the implementation of interventions through the platform. Additional demonstrations, whether successful or not, will help to determine the degree to which mTurk may be applied for these variety of research purposes. Future research will also benefit from the conduct of systematic reviews of the mTurk literature both within addiction science and throughout the variety of psychological sciences. Such reviews could evaluate many of the methodological

considerations posed above, such as how variations in payment schedules, attention checks, and qualification restrictions may impact study results.

It is unlikely that mTurk will remain the exclusive source for crowdsourced research (see Other Crowdsourcing Platforms section for discussion of existing and emerging alternative platforms). The focus on mTurk provided here is largely due to the current prominence of this source. Many of the benefits of crowdsourced work and applications to the study of substance use and misuse discussed in the context of mTurk will likely translate to new platforms developed in coming years. Our hope is that this review highlights the ways in which mTurk has ushered in a new methodological approach for researchers interested in health behaviors, broadly, and addiction science, specifically. Ultimately, the ability to complement those existing methods used in human laboratory, clinical trial, community intervention, and epidemiological work with the participant recruitment and testing afforded by crowdsourcing should help improve the rigor, reproducibility, and overall possibilities of research conducted in addiction science.

Table 2.1. Common Language Used in Crowdsourcing and Associated Research Terms

Term	Definition	Related Research Terms
Batch	A block of HITs available for completion by eligible participants. mTurk studies typically post multiple, smaller batches (e.g., 50 HITs) in order to survey individuals at varying times of day and days of the week.	Cohort, Wave
Crowdsourcing	The use of the Internet to complete a task through a flexible, open call to a large number of people.	Recruitment, Sampling
Human Intelligence Task (HIT)	A research study listing on mTurk. Typically includes a brief description of the study as well as compensation rate and expected time of completion.	Experiment, Study, Survey
mTurk ID	A unique identifier assigned to all worker and requester accounts on mTurk.	Participant ID
Qualification	A specific criterion used to determine if a worker can view the study HIT.	Inclusion/Exclusion Criteria
Requester	The individual (or group) posting tasks on mTurk.	Principal Investigator, Researcher
Worker	The individual completing a task on mTurk.	Participant, Subject

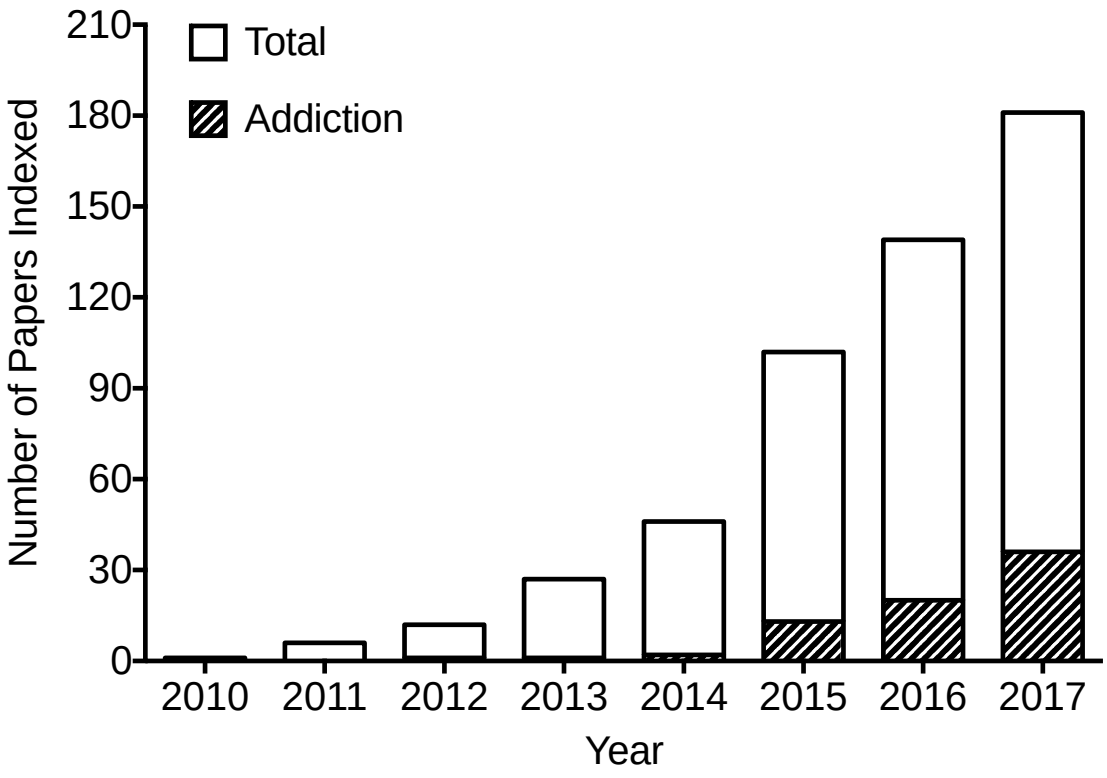

Figure 2.1. Number of manuscripts indexed in PsycINFO by the terms “Mechanical Turk” or “mTurk”. Plotted are the total (full bar) and addiction-related (crossed region) peer-reviewed manuscripts indexed in PsycINFO from 2010-2017. Addiction-related papers were identified by search terms related to addiction science (e.g., “substance use”, addiction, alcohol, cocaine).

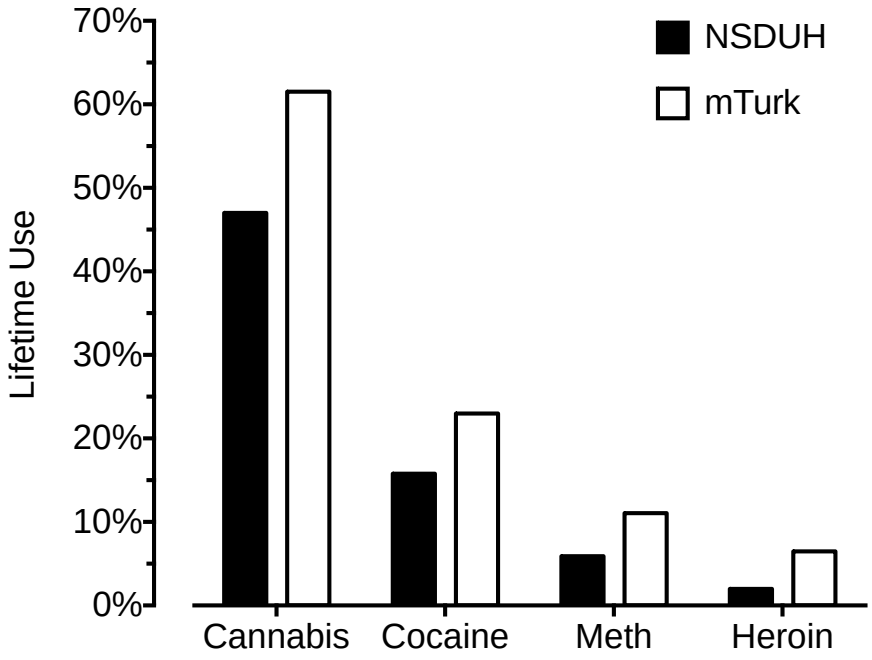

Figure 2.2. Percentage of adults reporting lifetime illicit substance use. Plotted are data from a nationally representative sample from the 2016 National Survey on Drug Use and Health (Black Bars [NSDUH]; Center for Behavioral Health Statistics, 2017) and data collected on mTurk as a part of study screeners within studies conducted in 2016 by our research team. mTurk data are estimated from 5269 unique participants.

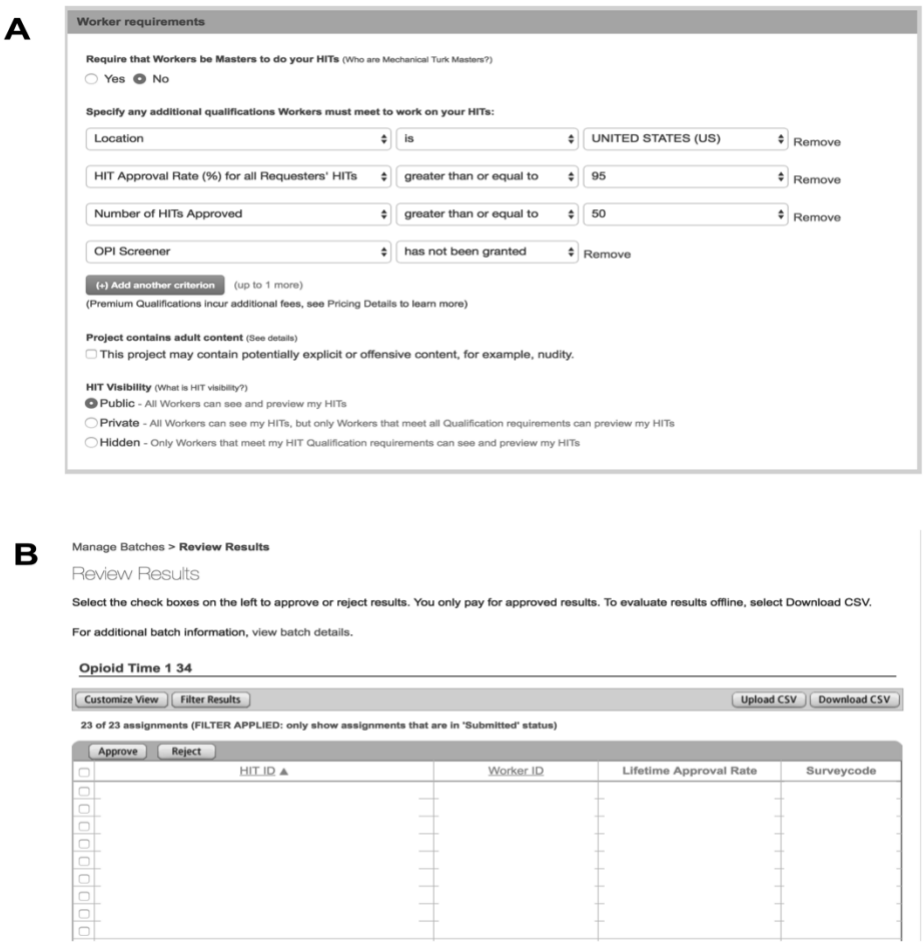

Figure 2.3. Example of the mTurk interface. Panel A depicts the qualification page in which a requester can set specific qualifications for research participants (e.g., location is United States). Requesters can also develop their own qualification to ensure that individuals do not complete the study more than once (e.g., OPI Screener). Panel B depicts the task approval page. This page allows the requester to view the mTurk workers who have completed a HIT, lifetime approval rating, and a survey code that the requester can include in the task to verify that participants have completed the survey. Bonus payments can also be delivered from this page. Note that identifying information has been removed from Panel B.

Chapter 3

UNIQUE PREDICTION OF CANNABIS USE SEVERITY AND BEHAVIORS BY DELAY DISCOUNTING AND BEHAVIORAL ECONOMIC DEMAND

(Experiment 1; Strickland et al. 2017b)

Introduction

The marriage of behavioral economics with substance use research over the last three decades has advanced an understanding of the etiology and treatment of drug-taking behavior (Bickel et al. 2000; Chivers and Higgins 2012; Hursh 1984; MacKillop 2016). Two of the most common applications of behavioral economics to drug use are delay discounting and behavioral economic demand. Delay discounting is the systematic reduction in value of a reinforcer as a function of the delay to its delivery (Green and Myerson 2004; Odum 2011; Rachlin and Green 1972). Excessive delay discounting is thought central to substance use disorders and may represent a trans-disease process relating drug use to other maladaptive health behaviors (Bickel et al. 2012; Koffarnus et al. 2013). Several meta-analyses support this assertion by demonstrating a robust relationship between delayed reward discounting and drug use severity, dependence, and quantity-frequency variables (Amlung et al. 2017b; MacKillop et al. 2011). This literature also provides evidence for an association between delay discounting and treatment response across diverse clinical populations (e.g., Krishnan-Sarin et al. 2007; MacKillop and Kahler 2009; Washio et al. 2011; Yoon et al. 2007).

Substance use researchers have also situated drug use within a commodity purchase framework and used behavioral economic demand models to describe drug-taking behavior (Hursh 1984, 1991; Johnson and Bickel 2006). A recent popular extension of these methods is the commodity purchase task in which participants report hypothetical or realized commodity consumption across a range of prices per unit of the commodity (Jacobs and Bickel 1999; MacKillop et al. 2008; Murphy et al. 2009). To date,

purchase tasks have been successfully applied to a variety of drugs and drug classes, including alcohol, cannabis, cigarettes, cocaine, opioids, and synthetic cathinones (Amlung and MacKillop 2015; Aston et al. 2015; Aston et al. 2016; Bruner and Johnson 2014; Collins et al. 2014; Johnson and Johnson 2014; MacKillop et al. 2008; Murphy and MacKillop 2006; Pickover et al. 2016). These studies have demonstrated that commonly used and misused substances follow the same prototypic patterns of consumption as other goods, including decreases in consumption with increases in price, and price ranges at which consumption is sensitive (i.e., elastic) or insensitive (i.e., inelastic) to price change. Such research has also helped reveal behavioral mechanisms by which putative interventions may decrease drug consumption (Bujarski et al. 2012; McClure et al. 2013b) and prognostic variables predicting treatment success (MacKillop and Murphy 2007; Murphy et al. 2015).

Despite being the most widely used illicit substance in the United States, cannabis has received comparatively little attention in the delay discounting and demand literatures. Further information on cannabis delay discounting and demand could be useful given the increasing number of states proposing or that have passed legalized recreational use, and the growing prevalence of cannabis use reported in the United States over the last decade. Although many first-time or recreational users will not continue to regular use, others will progress to problematic usage patterns and seek treatment for cannabis use disorder. The application of behavioral economic theory could help identify behavioral mechanisms contributing to maladaptive use and hasten the design of preventative and therapeutic interventions.

The association between delay discounting and cannabis use is more variable than for other drugs. For example, several studies have failed to show significant differences in delay discounting rates between cannabis users and controls (Johnson et al. 2010) or a relationship between delay discounting and treatment outcomes (Heinz et al. 2013;

Peters et al. 2013; but see Stanger et al. 2012). However, others have demonstrated a role for delay discounting in aspects of cannabis use with significant associations observed between delay discounting and cannabis use initiation and severity (Aston et al. 2016; Bidwell et al. 2013; Heinz et al. 2013; Kollins 2003). These differences in experimental outcomes could stem from the variations in the questions posed, namely delay discounting in cannabis users compared to controls as opposed to correlations between delay discounting and features of cannabis use (e.g., use severity). This limited literature also displays heterogeneity with respect to sample characteristics (e.g., college students versus treatment samples), delay discounting measures, and the commodity discounted. Such heterogeneity underscores the need for additional research to evaluate the association between delay discounting and aspects of cannabis, and to compare delay discounting rates between cannabis users and controls.

Cannabis demand has received even less attention than delay discounting, representing a research literature in its infancy. Existing studies have revealed outcomes consistent with previous work with other drugs, such as a sensitivity of cannabis demand to increased cost and an expected relationship between cannabis demand and measures of cannabis use severity and frequency (Aston et al. 2015; Collins et al. 2014). These studies have also demonstrated changes in cannabis demand following manipulations affecting state disposition, with increases observed in cannabis demand after cannabis -cue presentation (Metrik et al. 2016).

Recent attempts have been made to unify the ideas of delay discounting and demand under a broader “reinforcer pathology” conceptualization of substance use (Bickel et al. 2011a). This approach posits that substance use disorders are characterized by an extreme preference for immediate consumption of a drug reinforcer (i.e., delay discounting) combined with high valuation for that reinforcer (i.e., behavioral economic demand). Relatively little research, however, has simultaneously evaluated

demand and delay discounting metrics despite their mutual importance for this and other theoretical models. In fact, only one study has done so in the context of cannabis use (Aston et al. 2016). Data in that study were combined from participants completing a delay discounting task for money following placebo smoked cannabis administration and a purchase task for hypothetical cannabis completed during a baseline screening session. Delay discounting and behavioral economic demand functioned as independent predictors, with monetary delay discounting uniquely predicting cannabis dependence (CD) symptom count and cannabis demand uniquely predicting frequency of cannabis use. However, this previous study only evaluated monetary delay discounting and cannabis demand. The inclusion of delay discounting and demand measures for multiple drug commodities (e.g., cannabis, alcohol) would help to demonstrate the specificity of cannabis-relevant outcomes for predicting use behaviors.

The purpose of the present study was to replicate and extend those findings relating delay discounting and demand to cannabis use behaviors. To this end, Amazon Mechanical Turk (mTurk) was used to sample individuals reporting recent cannabis use. This study sought to extend previous research by 1) using an alternative sampling method allowing for greater demographic and drug use variability, 2) evaluating multiple commodities for delay discounting and demand variables, and 3) including a measure of use quantity (i.e., grams used per week) to further describe cannabis use patterns. Cannabis delay discounting was expected to uniquely predict CD symptom count and cannabis demand was expected to uniquely predict frequency and quantity of cannabis use. Alcohol delay discounting and demand were not expected to relate to cannabis use behaviors. As a secondary analysis, measures of delay discounting and behavioral economic demand were compared between cannabis users and non-cannabis using controls to add to the limited literature evaluating these outcomes between these groups. Significant differences between cannabis users and controls were not expected.

Methods

Participants and Procedures

Participants were sampled using mTurk where tasks are advertised as Human Intelligence Tasks (HITs). Participants were required to have a 95% or higher approval rating on all previously submitted mTurk HITs, over 100 approved HITs, and current residence in the United States to view the study HIT. The accuracy of these inclusion criteria was verified by the mTurk platform. Participants reviewed an informed consent document describing the study procedures, compensation, and the fact that anonymity would be retained throughout the study. All respondents indicated by electronic confirmation that they understood this document and agreed to participate. The University of Kentucky Institutional Review Board approved all protocols, including the consent process, and the protocol was carried out in accordance with the Declaration of Helsinki.

Participants completed the study tasks as a part of larger study on choice and decision-making. A short screening questionnaire was used to determine if participants qualified for this study. Participants were only able to complete the screener once. Eligible participants were individuals reporting cannabis use during the past two weeks and 50 or more lifetime uses (n = 78). An additional control group was used to compare behavioral economic demand and delay discounting outcomes and included participants who did not report cannabis use in the past two weeks and five or fewer lifetime uses (n = 86). All participants were 18 years of age or older. Four attention checks were used to identify non-systematic, inattentive, or inconsistent participant data: 1) comparison of age and sex responses at the start and end of the survey, 2) recall at the end of the survey of a single digit number presented halfway through that participants were instructed to remember, 3) an item that instructed participants to select a specific response (i.e., “Select ‘A Little Bit’”), and 4) an item that asked participants if they had

been attentive and thought their data should be included. Participants failing one or more checks were removed from data analysis, which resulted in a final sample size of 136 participants (cannabis users = 64; controls = 72). Demographic and drug use variables for cannabis users and controls are presented in Table 3.1.

Measures

5-Trial Adjusting Delay Task

A 5-trial adjusting delay task was used to evaluate delay discounting rates (for details see Koffarnus and Bickel 2014). This task has been previously validated against traditional adjusting amount delay discounting tasks (Cox and Dallery 2016; Koffarnus and Bickel 2014). Participants were asked to select between some amount of a delayed commodity and half that amount available immediately. The delayed and immediate amounts remained constant while the delay to the larger amount was adjusted after each choice. The first choice was always at a three-week delay, which then adjusted up (longer delay following delayed choice) or down (shorter delay following immediate choice) based on decisions. The ED_{50}, the inverse of the delay discounting rate or *k*, was determined following five choices and included 32 potential values between 1 hour and 25 years. The benefits of this 5-trial task include rapid assessment of delay discounting rates and minimal computing requirements. These advantages are particularly important given the online research context wherein time is limited and data collection is constrained by the participant's computer equipment.

Participants completed up to three different versions of the task. All participants completed a traditional monetary delay task, with $1000 available delayed versus $500 available now. All cannabis-using participants also completed a cannabis delay discounting task, with 1 ounce of typical quality cannabis available delayed versus ½ ounce available now. Any participant endorsing current alcohol use also completed an alcohol delay task, with 24 US standard drinks available delayed versus 12 drinks

available now (standard drink: one 12 oz. beer, 5 oz. wine, or 1.5 oz. shot/mixed drink). These commodity amounts were within ranges used in previous alcohol and cannabis delay discounting studies (Johnson et al. 2010; Petry 2001). All delay discounting tasks included a statement emphasizing that consumption was not constrained by time and that the hypothetical goods could be kept.

Commodity Purchase Tasks

Commodity purchase tasks were used to evaluate behavioral economic demand for cannabis and alcohol. The same scenario was used in each task and all purchasing situations were framed as hypothetical in the present tense. Participants were asked to imagine a typical day over the last month when they used the commodity. They were told that they could only get the commodity from this source, had no commodity saved or kept from previous days, could not stockpile, and would have to consume all purchases in a single day. Participants were then asked how many drinks (alcohol) or hits (cannabis) they would purchase at 13 monetary increments ranging from $0.00 [free] to $11/unit, presented sequentially. This price range was selected due to its similarity to price ranges used in recent commodity purchase task studies (Amlung et al. 2015a; Aston et al. 2015; Aston et al. 2016; Murphy et al. 2015). Alcohol drinks were described as one US standard drink. Cannabis hits were quantified as 0.09 g of average quality cannabis (i.e., 10 hits = 1 joint or 0.9 g or 1/32nd of an ounce) consistent with previous literature (Aston et al. 2015). All cannabis-using participants completed the commodity purchase task for cannabis (Marijuana Purchase Task [MPT]; note the term Marijuana Purchase Task [MPT] is used to avoid confusion with the commonly used Cigarette Purchase Task [CPT]) and all participants endorsing current alcohol use the task for alcohol (Alcohol Purchase Task; APT).

Drug Use Variables

A written version of the Mini-International Neuropsychiatric Interview (MINI) was used to evaluate cannabis dependence (CD) through endorsement of statements indicative of *DSM-IV* criteria (Sheehan et al., 1998). A Cannabis History and Smoking Questionnaire included questions about age of use onset, cannabis use patterns, and routes of administration (Aston et al. 2015; Metrik et al. 2009). Cannabis use variables included use severity (number of DSM-IV CD symptoms endorsed), frequency (percentage past month use days), and quantity (grams used per week), consistent with previous delay and demand literature (Aston et al. 2016). Other relevant drug use variables and demographics, such as age and the Fagerström Test for Nicotine Dependence (FTND), were also measured.

Data Analysis

Delay discounting rates (*k* values) were calculated by taking the inverse of ED_{50} values derived from the 5-trial adjusting delay task (Koffarnus and Bickel 2014). Delay discounting rates were log-transformed prior to analysis to obtain normality. Demand curves were first evaluated for inattentive data or non-systematic curves using standard criteria (Stein et al. 2015). Briefly, curves were evaluated for increased consumption with increased price, frequent price-to-price consumption increases, or reversals from zero consumption as well as extreme consumption (i.e., greater than 100 drinks or hits [9 grams of cannabis] in one day). Nine cannabis curves (14.1%) and 16 alcohol curves (15.7%) were determined non-systematic/inattentive and removed from demand analysis. Price elasticity and intensity were generated using the exponentiated demand equation:

$$Q = Q_0 * 10^{k(e^{(-\alpha * Q0 * C)} - 1)}$$

where Q = consumption; Q_0 = derived intensity of demand (consumption at zero price); k = a constant that denotes consumption range in log units (*a priori* set to 2); C = the price of the commodity; and α = derived elasticity of demand. The exponentiated model is a recently developed and validated equation that allows for the inclusion of zero consumption values (Koffarnus et al. 2015; Strickland et al. 2016b). Area under the curve (AUC) values were also generated as described previously (Amlung et al. 2015b; Aston et al. 2016). Briefly, the total area was operationalized as the AUC value when the maximum consumption value across the sample was inputted for each price (100 for both the MPT and APT). Proportionate AUC values were then generated by dividing each participants raw AUC by this total AUC (range = 0.0 to 1.0). Recent reports have proposed AUC as a single demand metric that is useful and valid measure to minimize repeated testing with multiple demand metrics and to allow for convergence with other behavioral economic measures (Amlung et al. 2015b; Aston et al. 2016). All demand metrics were log-transformed to correct for skew.

Demographic and drug use variables for cannabis users and controls were compared using independent sample *t*-tests. Independent sample *t*-tests were also used to compare monetary delay discounting as well as alcohol delay discounting and demand between groups. The relationship between cannabis use and demographic, demand, and delay discounting variables was first described using bivariate correlations in the cannabis-using group. The independent contribution of demand and delay discounting for predicting cannabis use was then determined using multiple regression models. AUC was first used to quantify demand in these analyses consistent with previous studies (Aston et al. 2016). Follow up analyses were then conducted using demand intensity and elasticity in place of AUC. Additional tests were also conducted controlling for age, sex, income, and cigarette use (given the close association between cigarette use and delay discounting) (Bickel et al. 1999; Johnson et al. 2007). Only

cannabis-using participants with all demand and delay discounting variables were included in regression analysis (n = 46 with complete data). All tests were conducted using SPSS 24 and GraphPad Prism 6.0 with a type I error rate of .05.

Results

Between-Group Comparisons

Demographics and Drug Use

Individuals reporting cannabis use were younger and more likely to report recent cigarette and alcohol use (i.e., in the past two weeks). Cannabis users also reported more drinking days per week and alcoholic drinks per week. Other demographic variables (e.g., race, income) did not differ between groups (Table 3.1).

Delay Discounting

Raw monetary, cannabis, and alcohol delay discounting rates are presented in Table 3.2. Monetary delay discounting did not differ between cannabis users and control (t_{134} = 0.52, p = .60). Similarly, alcohol delay discounting did not differ between these groups (t_{100} = 0.38, p = .70). No differences in the magnitude or significance of these findings were observed after controlling for relevant covariates (e.g., age, cigarette use, alcohol use).

Drug Demand

Demand curves for the MPT and APT showed prototypical decreases in consumption with increases in price for both groups (Figure 3.1). The exponentiated demand equation provided an excellent fit to group data (MPT: R^2 = .99; APT: Cannabis Users R^2 = .98, Controls R^2 = .98) as well as individual curves (MPT: median R^2 = .95, IQR = .91 to .97; APT: median R^2 = .87, IQR = .79 to .93). No between-group differences were observed for alcohol AUC (t_{85} = 0.39, p = .70). Similarly, no differences were observed for alcohol demand intensity (t_{85} = 0.50, p = .62) or elasticity (t_{85} = 0.51, p =

.61). No changes in the magnitude or significance of these comparisons between cannabis users and controls were observed in covariate analyses.

Drug Use Prediction in Cannabis Users

Bivariate Relationships

Bivariate relationships among drug demand, delay discounting variables, and self-reported drug use in cannabis users are presented in Table 3.3. Cannabis AUC was significantly and positively related to grams of cannabis used per week (r = .45, p = .001), percentage past month use days (r = .31, p = .02), and number of CD symptoms endorsed (r = .42, p = .001). Cannabis delay discounting was related to grams of cannabis used per week (r = .38, p = .002) and number of CD symptoms endorsed (r = .44, p < .001).

Intercorrelations between delay discounting and AUC revealed three significant associations all involving cannabis delay discounting. Cannabis delay discounting was significantly and positively related to cannabis demand (r = .34, p = .01), monetary delay discounting (r = .42, p = .001), and alcohol delay discounting (r = .33, p = .01).

Regression Models

Results from three regression models evaluating the unique prediction of cannabis use by AUC and delay discounting variables are presented in Table 3.4. Cannabis delay discounting was uniquely associated with the number of CD symptoms endorsed (sr^2 = .16). The model predicting percentage of past month use days indicated a significant independent effect of cannabis AUC (sr^2 = .09), but not alcohol AUC or delay discounting variables. A significant unique effect of cannabis AUC was also observed in the model predicting grams of cannabis used per week (sr^2 = .12). Models including additional covariates (e.g., age, cigarette use) revealed outcomes similar in magnitude and significance. Cannabis AUC by delay discounting interactions were not significant when tested.

Additional models were used to determine if the association between cannabis use behaviors and behavioral economic demand was related to demand intensity (Q_0) and/or elasticity (α). Models including demand intensity and elasticity revealed a similar pattern of effects as the AUC analysis, with cannabis demand intensity uniquely contributing to cannabis use frequency and quantity and cannabis delay discounting uniquely contributing to use severity (Table 3.5). Specifically, cannabis demand intensity was significantly and positively related to use quantity (sr^2 = .29) and frequency (sr^2 = .19). Alcohol demand intensity also showed an inverse relationship with grams of cannabis used per week (sr^2 = .15).

Discussion

The primary finding of the present study was that cannabis delay discounting uniquely predicted use severity (i.e., CD symptom count), whereas cannabis demand uniquely predicted use frequency (i.e., past month use days) and quantity (i.e. grams used per week) in regression models. Follow-up analyses indicated that the primary behavioral mechanism contributing to the relationship between behavioral economic demand and cannabis use frequency and quantity was demand intensity. These findings are consistent with the only other study to evaluate the unique contribution of cannabis delay discounting and demand to cannabis use behaviors (Aston et al. 2016). Several methodological concerns potentially limited the generalizability of this aforementioned study, including the use of an exclusively white sample, low prevalence of CD symptoms, and data collection following placebo cannabis self-administration that may have influenced reported delay discounting outcomes (see Metrik et al. 2009; Metrik et al. 2012, for discussion of cannabis expectancies). Our findings suggest that these experimental parameters did not contribute to the observed relationships and that the unique contribution of delay discounting and demand to cannabis use outcomes likely generalize to diverse experimental settings and populations. Replication studies such as

this one are particularly important given recent challenges the psychological and behavioral sciences have faced regarding reproducibility (Nosek et al. 2015).

A consistency across online and in-laboratory samples also strengthens the case for the use of online sampling techniques to evaluate substance-using populations and behavioral mechanisms related to drug use (Koffarnus et al. 2015; Rass et al. 2015a; Strickland et al. 2016a). The use of crowdsourcing comes with many benefits including increased geographic and demographic variability, targeted recruitment of hard-to-reach populations, and a relative cost and time efficiency. A requisite step for using crowdsourcing, however, is the validation of findings across Internet and laboratory settings. This is particularly important for addiction research given that drug use cannot be biologically verified in participants online. The current study adds to the extant literature demonstrating similar outcomes across in-laboratory and online samples (Johnson et al. 2015; Strickland et al. 2016a), thereby supporting the validity of the crowdsourcing approach. These findings are particularly exciting as they offer a method by which findings can be replicated across diverse samples using sampling methods that provide complementary benefits and limitations.

In addition to replicating associations involving cannabis demand and delay discounting in this novel sample, this study extends previous findings in at least three ways. First, cannabis and alcohol demand and delay discounting measures were used to determine if the observed relationships were specific to cannabis. Alcohol was selected given the extensive literature validating the alcohol purchase task, the common use of alcohol in the general population, and the expectation that alcohol use should not uniquely associate with elevated cannabis use frequency, quantity, or severity. Cannabis demand and delay discounting specifically contributed to these cannabis use variables, even after controlling for the contribution of another common drug commodity (i.e., alcohol). Alcohol demand intensity was related to cannabis use quantity when testing

demand intensity and elasticity measures; however, this relationship was not observed at the bivariate level and was smaller than the relationship involving cannabis demand intensity in the multiple regression model. Although monetary delay discounting was related to CD symptom count at the bivariate level, consistent with previous findings (Aston et al. 2016), the association was not unique or significant when controlling for the cannabis commodity relationship. These outcomes indicate that these behavioral economic relationships do not likely represent a general propensity to respond in a non-specific manner to the task requirements. Instead, they suggest that these relationships with drug use outcomes are specific to drug of interest, in this case cannabis.

Second, a quantity measure was included in addition to the frequency and severity measures previously evaluated. Similar outcomes as the frequency measure were observed, wherein cannabis demand uniquely predicted the quantity of cannabis use above and beyond other demand and delay discounting measures. This distinction is important given that cannabis use frequency and quantity represent unique dimensions of use patterns and can provide unique prediction of cannabis-related problems (Zeisser et al. 2012). Third, the recently developed 5-trial adjusting delay task was used, which allowed for rapid (~ 1 minute) generation of delay discounting rates. Prior research has validated this task by revealing a close relationship between this rapid task and traditional adjusting delay tasks (Cox and Dallery 2016; Koffarnus and Bickel 2014). To our knowledge, this is the first use of drug commodities, namely alcohol and cannabis, with this rapid delay discounting task. That the relationship between monetary, alcohol, and cannabis delay discounting rates and drug use outcomes were generally significant and in the expected direction provides further support for use of this rapid assessment task in substance use research. The more complex relationships involving delay discounting (i.e., multiple regression analyses) were also consistent with previous

findings (Aston et al. 2016), further supporting the validity of this delay discounting task and its future use.

Secondary analyses that focused on differences in delay discounting and demand variables between cannabis-users and non-using controls were also conducted given the limited research comparing these groups on these measures. Significant differences in monetary delay discounting as a function of cannabis use status were not observed, although this outcome is generally consistent with previous research (Johnson et al. 2010). Differences in alcohol delay discounting or demand were also not detected among cannabis users and controls. These outcomes remained after controlling for potentially important covariates, such as age or cigarette use, suggesting that the failure to detect differences was not due to low power or imbalances in relevant demographics. Taken together, these findings suggest that alterations in delay discounting and demand may not capture between-group differences in the likelihood of reporting current cannabis use. Instead, these measures may reflect differences in the propensity to use cannabis in a maladaptive or clinically relevant manner.

These findings should be considered within the context of their limitations. First, a large proportion of subjects were excluded due to inattentive or non-systematic data. This rate is generally consistent with previous mTurk research (Johnson et al. 2015), but is a limitation of conducting online research, and underscores the need to verify response fidelity due to decreased experimental control. Second, all participants completed the demand measures first followed by the delay discounting measures and completed monetary delay discounting prior to drug tasks. Although this non-random completion order did not likely alter study outcomes, additional research would be needed for confirmation. Third, as noted above, cannabis use in the cannabis-using group was not biologically verified. However, the use of rigorous internal control measures and the consistency between this study and the previous laboratory one

(Aston et al. 2016) supports the integrity of the data collected and outcomes reported. Fourth, participants were not asked about how much money they typically spent on cannabis. Because participants reported 50 or more lifetime uses, they were likely at least familiar with cannabis pricing. Fifth, the 5-trial delay discounting task did not permit examination of underlying orderliness in the raw data like traditional delay discounting tasks. However, the use of rigorous manipulation checks to evaluate participant attention and response fidelity helped to offset this limitation. Despite these potential limitations, this study replicates and extends previous research uniquely relating delay discounting and behavioral economic demand variables with cannabis use. This study also demonstrates the ability of online crowdsourcing to complement standard human laboratory procedures, thereby enhancing the overall rigor of research methods used to examine behavioral mechanisms of drug use disorders.

Table 3.1. Demographics and Drug Use Variables.

	Cannabis User (n = 64)	Control (n = 72)	
	Mean (SD)/%	Mean (SD)/%	*p*
Demographics			
Age	30.2 (7.3)	33.7 (9.9)	.02*
Male	48.4%	44.4%	.73
White	74.2%	75.0%	.99
College Education	50.0%	65.3%	.08
Household Income	$41094 ($28122)	$43889 ($28756)	.57
Cannabis Use			
Ever Use	100.0%	34.7%	-
Lifetime Use	5810 (21166)	0.8 (1.7)	-
% Month Use Days	63.7% (34.4%)	0%	-
Grams/Week	6.4 (7.1)	-	-
# CD Symptoms	2.0 (2.1)	-	-
Other Drug Use			
CIG Use	60.9%	15.3%	.001***
FTND[a]	3.9 (2.6)	4.0 (3.0)	.88
ALC Use	85.9%	59.7%	.001**
Drinks/Week[b]	9.2 (14.4)	3.7 (7.8)	.03*
Drinking Days/Week[b]	2.5 (2.0)	1.6 (1.4)	.01*
Heavy Use Days[b]	72.7%	41.9%	.003**

Note. CIG Use = cigarette use in the past two weeks determined from study screener; ALC Use = alcohol use in the past two weeks determined from study screener; # CD Symptoms = number of cannabis use dependence symptoms endorsed; FTND = Fagerström Test for Nicotine Dependence; Heavy Use Days = number of past month heavy alcohol use days (5 or more [males] or 4 or more [females] drinks in a single day).

[a] Variable only for participants reporting cigarette use in the past two weeks (n = 39 Cannabis Users; n = 11 Controls)
[b] Variable only for participants reporting alcohol use in the past two weeks (n = 55 Cannabis Users; n = 43 Controls)

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

Table 3.2. Drug Behavioral Economic Demand and Delay Discounting in Cannabis Users and Controls.

	Cannabis User (N = 64)		Control (N = 72)		
Variable	Mean (SD)	n	Mean (SD)	n	*p*
Demand (AUC)					
Cannabis	.04 (.08)	55	-	-	-
Alcohol	.02 (.02)	50	.02 (.01)	37	.70
Demand (Q_0)					
Cannabis	35.6 (32.5)	55	-	-	-
Alcohol	7.0 (8.7)	50	7.7 (16.8)	37	.62
Demand (α)					
Cannabis	.028 (.047)	55	-	-	-
Alcohol	.061 (.138)	50	.027 (.047)	37	.61
Delay Discounting (*k*)					
Money	.02 (.06)	64	.02 (.05)	72	.60
Cannabis	.97 (3.79)	64	-	-	-
Alcohol	3.83 (7.97)	59	3.90 (8.59)	43	.70

Note. AUC = area under the curve from purchase task data; Q_0= demand intensity from the exponentiated demand equation; α = demand elasticity from the exponentiated demand equation; *k* = delay discounting rate from 5-trial adjusting delay discounting task. Raw values are presented for descriptive purposes; all data were log-transformed for normality prior to analysis.

Table 3.3. Correlations Among Behavioral Economic Demand, Delay Discounting, and Drug Use Variables.

			Cannabis Use			Demand		Delay Discounting		
	Age	Male	CD Count	Grams/ Week	Month Use	CAN	ALC	MON	CAN	ALC
Demand										
CAN (n = 55)	.02	-.01	**.42****	**.45****	**.31***	-	-	-	-	-
ALC (n = 50)	.07	-.14	.02	.00	.11	.25	-	-	-	-
Delay Discounting										
MON (n = 64)	-.16	-.11	**.27***	.24	.10	.20	.15	-	-	-
CAN (n = 64)	.01	.06	**.44****	**.38****	.04	**.34***	.22	**.42****	-	-
ALC (n = 59)	.04	-.01	.12	.07	-.10	-.08	.01	.18	**.33***	-

Note. CAN = Cannabis; ALC = alcohol; MON = money; AUC = area under the curve from purchase task data; *k* = delay discounting rate from 5-trial adjusting delay discounting task; Alcohol Heavy = past month heavy alcohol use days (5 or more [males] or 4 or more [females] drinks in a single day): CD Count = number of cannabis use dependence symptoms endorsed; Month Use = percentage past month cannabis use days. n = 64 cannabis users, sample sizes included in the table vary for each measure due to self-reported use or rates of systematic data. Bold = significant at $p < .05$
* $p < .05$; ** $p < .01$

Table 3.4. Area Under the Demand Curve and Delay Discounting as Predictors of Cannabis Use Variables.

	# CD Symptoms			Grams Cannabis/Week			% Month Use Days		
Predictor	*b*	β	*P*	*b*	β	*p*	*b*	β	*p*
Money *k*	.50	.19	.18	.81	.12	.46	.03	.06	.69
Cannabis AUC	.77	.21	.15	**3.70**	**.39**	**.02***	**.23**	**.35**	**.05***
Cannabis *k*	**.93**	**.48**	**.003****	.18	.04	.84	-.08	-.22	.22
Alcohol AUC	-.52	-.17	.21	-.98	-.12	.42	.06	.10	.51
Alcohol *k*	-.13	-.10	.50	.32	.09	.58	.02	.08	.66

Note. AUC = area under the demand curve from purchase task data; *k* = discounting rate from 5-trial adjusting delay discounting task; % Month Use Days = percentage past month cannabis use days; # CD Symptoms = number of cannabis use dependence symptoms endorsed. n = 46 cannabis users. Bold = significant at $p < .05$

* $p < .05$; ** $p < .01$

Table 3.5. Demand Intensity, Demand Elasticity, and Delay Discounting as Predictors of Cannabis Use Variables

	# CD Symptoms			Grams Cannabis/Week			% Month Use Days		
Predictor	*b*	β	*p*	*b*	β	*p*	*b*	β	*p*
Money-*k*	.48	.18	.21	-.10	-.02	.91	-.02	-.04	.82
Cannabis-Q_0	1.14	.24	.12	**8.03**	**.66**	**.001*****	**.46**	**.53**	**.003****
Cannabis-α	-.35	-.09	.53	-1.57	-.16	.24	-.09	-.14	.40
Cannabis-*k*	**.87**	**.44**	**.01****	-.20	-.04	.78	-.10	-.28	.10
Alcohol-Q_0	.12	.03	.88	**-6.64**	**-.52**	**.002****	-.30	-.33	.09
Alcohol-α	.55	.19	.26	-1.82	-.24	.14	-.18	-.33	.07
Alcohol-*k*	-.13	-.10	.49	.18	.05	.70	.01	.04	.78

Note. Q_0= demand intensity from the exponentiated demand equation; α = demand elasticity from the exponentiated demand equation; *k* = discounting rate from 5-trial adjusting delay discounting task; % Month Use Days = percentage past month cannabis use days; # CD Symptoms = number of cannabis use dependence symptoms endorsed. n = 46 cannabis users. Bold = significant at $p < .05$

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

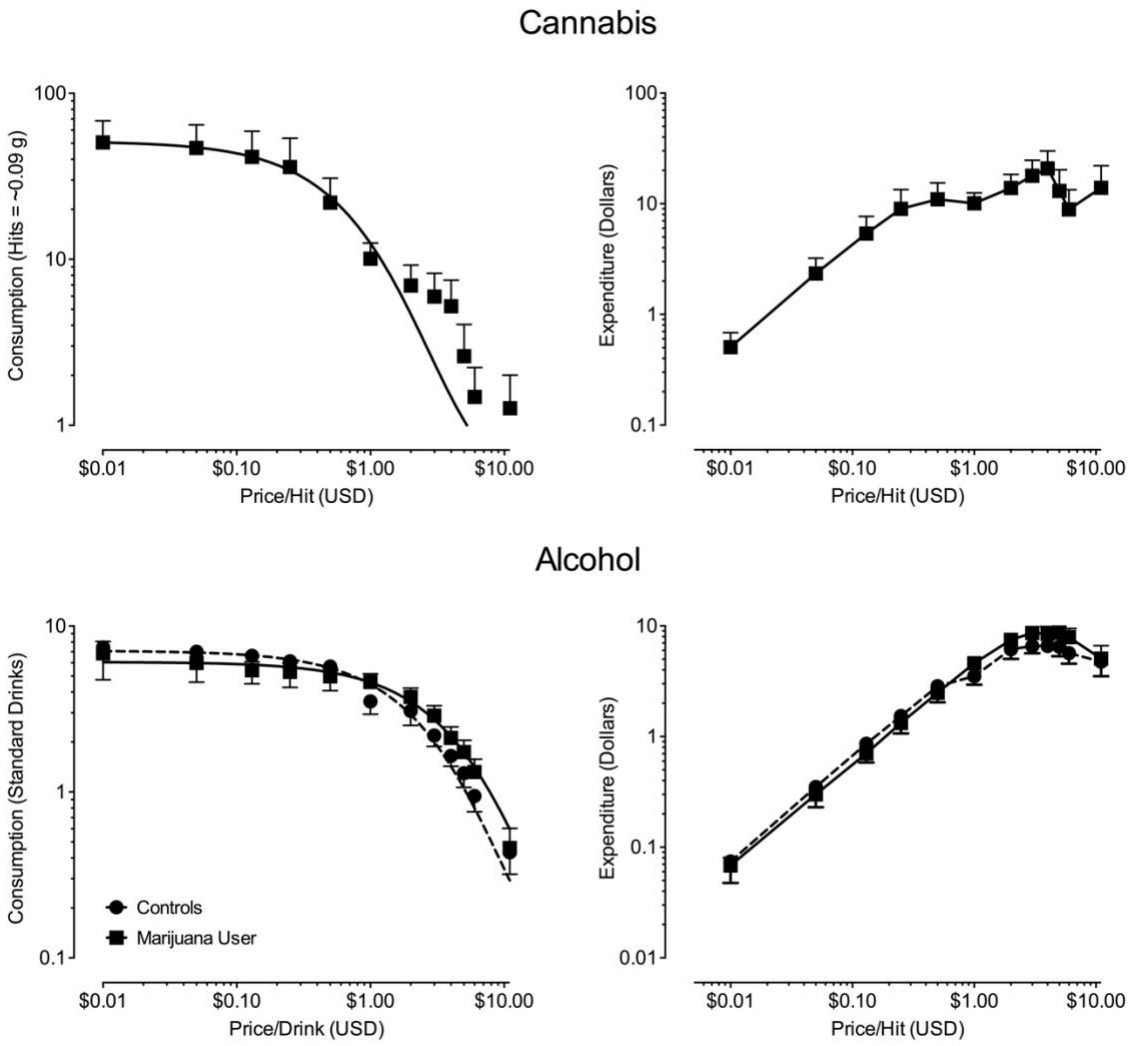

Figure 3.1. Behavioral economic demand (left panels) and expenditure (right panels) for cannabis (top panels) and alcohol (bottom panels). Participants completed commodity purchase tasks in which hypothetical cannabis (one hit quantified as 0.09 g) or alcohol (one US standard drink) were available. Price varied in United States dollars (USD). Plotted are mean (SEM) group data on a log-log axis fit using the exponentiated model. Squares and solid lines represent cannabis users (n = 55 for cannabis demand and 50 for alcohol demand) and circles and the dotted lines represent controls (n = 37 for alcohol demand).

Chapter 4

STIMULUS-SELECTIVITY OF DRUG PURCHASE TASKS: EVALUATING ALCOHOL AND CIGARETTE DEMAND

(Experiment 2; Strickland and Stoops 2017)

Introduction

The merger of theoretical perspectives and methodologies from behavioral economics and operant theory has resulted in numerous advances in addiction science (Bickel et al. 2014; Bickel et al. 2000; Hursh 1984). One prominent example of this interdisciplinary approach is the application of consumer demand theory to drug-taking behavior. Demand curves allow researchers to graphically represent drug consumption across variations in price and are used to generate metrics thought to underlie drug use and reinforcement (Hursh and Roma 2013). A widely used method for evaluating economic demand in humans is the hypothetical purchase task. Demand curves are generated with these purchase tasks by asking participants to report hypothetical consumption of a good (e.g., alcohol) across a range of prices (e.g., $0.01, $1.00, $10.00/drink). This methodology is particularly appealing because of its temporal reliability (Few et al. 2012; Murphy et al. 2009), cost and time efficiency, and adaptability for populations with whom drug self-administration or other typical measures of drug use are not ethically or practically feasible (e.g., patients in residential treatment; participants with contraindications to drug administration).

Alcohol and cigarettes are the most commonly studied commodities in drug purchase task research, likely due to their legal status, widespread use, and relevance for other substance use and mental health conditions (Degenhardt et al. 2001; Grant and Harford 1995; McKay et al. 1999). Alcohol and cigarette purchase tasks have been largely successful, with consistent relationships observed between demand metrics and measures of drug use and use disorder (see reviews in Bickel et al. 2014; MacKillop

2016). These studies have also demonstrated that alcohol and cigarette purchase tasks are sensitive to state-level changes in drug demand, such as those following stress-induction, withdrawal, or cue presentation (Amlung and MacKillop 2014; MacKillop et al. 2012; Owens et al. 2015). Although the clinical relevance of drug demand is still under investigation, preliminary evidence suggests that demand metrics may help identify behavioral mechanisms underlying effective interventions (Bujarski et al. 2012; McClure et al. 2013b; but see Schlienz et al. 2014) or function as prognostic variables predicting treatment success (MacKillop and Murphy 2007; Madden and Kalman 2010; Murphy et al. 2015).

The use of purchase tasks in human behavioral pharmacology and addiction research has grown in recent years given these promising clinical findings and the numerous benefits that purchase tasks may offer. As applied research utilizing purchase tasks has proliferated, however, so has the continued need for methodological and parametric evaluation of these procedures. Certainty in capturing the essential aspects of demand that purchase tasks are purported to measure relies on such research concerning measurement reliability, validity, and fidelity.

Several studies have demonstrated the psychometric properties of purchase tasks, including their test-retest reliability, construct validity, and incremental validity (Few et al. 2012; MacKillop et al. 2008; Murphy et al. 2009; Murphy et al. 2011). One area that has received less attention is the systematic study of stimulus-selectivity. Stimulus-selectivity for this purpose is broadly defined as a condition under which a specific stimulus input or target (e.g., alcohol, cigarette) is the primary determinant of behavior (e.g., demand) (Powell et al. 2016). In the context of cognitive-behavioral research, stimulus-selectivity implies that the stimulus presented during a task determines behavior as opposed to a general propensity to respond without respect to specific contextual determinants. Purchase tasks, as typically utilized, are thought to determine commodity specific

demand (e.g., cigarette valuation in the cigarette purchase task). If behavior is stimulus-selective then responses should reflect only the value of or demand for that commodity under study. However, it is possible that responses could represent an overall valuation for reinforcers without regard to the commodity under study. Although domain-general outcomes and a related hypo- or hyper-valuation of reinforcement may be important for understanding reinforcer sensitivity as it relates to drug use, this generalized responding weakens the fidelity of purchase tasks for specifically measuring demand for particular drug commodities.

Little research has focused on and systematically evaluated the stimulus-selectivity of purchase task metrics. A recent study included purchase tasks for six common non-drug commodities (e.g., toilet paper, vacation packages) across a range of price densities (Roma et al. 2016). Differences in and the rank order of demand metrics across and within commodity manipulations were generally consistent with the commodity under purchase, supporting the notion that the commodity was the primary determinant of purchasing behavior (i.e., that the task was stimulus selective). To our knowledge, only one study has simultaneously examined demand for a drug (i.e., cigarettes) and non-drug (i.e., chocolate) commodity to establish this selectivity within the context of behavioral pharmacology and addiction research (Chase et al. 2013). Chocolate demand in that study was not associated with nicotine dependence, thereby providing preliminary support for the stimulus-selectivity of the purchase task metrics. However, the relationship between cigarette demand and chocolate use was not measured, preventing the reciprocal interpretation of stimulus-selectivity.

The overall purpose of the present study was to provide a preliminary evaluation of the stimulus-selectivity of drug purchase tasks. Participants either completed alcohol and soda purchase tasks (Experiment 2a) or cigarette and chocolate purchase tasks (Experiment 2b) and demand metrics were compared to self-reported use behaviors.

Demand was predicted to closely associate with commodity-similar variables (e.g., alcohol demand to weekly alcohol use), but not with commodity-dissimilar ones (e.g., alcohol demand to weekly soda use). Such commodity-similar associations would support stimulus-selectivity by demonstrating that the commodity under study is the primary determinant of choice and behavior.

Experiment 2a Methods

Participants and Procedures

Participants were recruited from Amazon.com's Mechanical Turk (mTurk), a crowdsourcing platform that provides cost-effective and efficient sampling of diverse populations. All surveys were completed on the Qualtrics (Provo, UT) platform. Data were collected as a part of a larger study on choice and drug-related cues. Participants were required to have an approval rating of 95% or higher on at least 100 mTurk tasks, currently reside in the United States, and be 18 years of age or older to view the parent studies. Previous research in substance-using populations has documented a close correspondence between laboratory and online crowdsourced outcomes, supporting the validity of the approach (e.g., Johnson et al. 2015; Strickland et al. 2016a). Participants were compensated $0.05 for completion of a screener survey and up to a $2.50 bonus for completion of the full survey. Bonus amounts varied in the parent study depending on the number of tasks completed; however, participants were not informed of total payment until the end of the survey to ensure that differential payment did not influence experimental outcomes. All participants provided informed consent via electronic confirmation. The University of Kentucky Institutional Review Board approved all procedures, including the consent process.

Participants qualified if they endorsed current alcohol and current soda use (n = 166; no time period of consumption other than "current" was specified). Several attention checks were used to identify inattentive or non-systematic participant data. These

checks included: 1) comparison of age and sex responses at the start and end of the survey, 2) recall of a single digit number presented halfway through the survey that participants were instructed to remember and enter at the end of the survey, 3) an item that instructed participants to select a specific response (i.e., "Select 'A Little Bit'"), and 4) an item asking participants if they had been attentive and thought their data should be included. Nineteen participants were removed for failing one or more attention checks included to ensure participant engagement and response fidelity. Eight additional participants were removed due to non-systematic demand data (see Purchase Tasks below). This resulted in a final analyzed sample of 139 participants. See Table 4.1 for demographic and alcohol/soda use variables.

Measures

Commodity Purchase Tasks

An alcohol purchase task (Murphy et al. 2009) and a novel soda purchase task were used to evaluate demand. Participants were asked to imagine a typical day over the last month when they would drink alcohol (or soda) and to indicate the hypothetical number of alcoholic drinks (i.e., one preferred brand US standard drink) or sodas (i.e., one preferred brand 12 oz serving of soda) they would purchase at 16 monetary increments ranging from $0.00 to $140/drink, presented sequentially (full range: $0.00 [free], $0.01, $0.05, $0.13, $0.25, $0.50, $1.00, $2.00, $3.00, $4.00, $5.00, $6.00, $11.00, $35.00, $70.00, $140.00/unit). This price range was selected to accommodate the elastic and inelastic portion of the demand curves for a wide range of commodities. This range was also within those used in other purchase task literature, including studies conducted with alcohol (Bujarski et al. 2012; MacKillop et al. 2010b) and cigarettes (MacKillop et al. 2008; Wilson et al. 2016). Participants were instructed that they could only get drinks from this source, could not stockpile them, and would have to consume all purchases in a single day (see Appendix for example instructions). All choices were hypothetical and

participants completed the tasks in a fixed order of the alcohol purchase task before soda purchase task.

Alcohol and Soda Use Behaviors

Participants completed questions evaluating drug use and other health behaviors (e.g., “How many alcoholic drinks do you typically have in a week”, “How many days per week do you typically drink soda?”). Corresponding measures were evaluated or computed for alcohol and soda use. Quantity-frequency measures included: 1) number of drinks per week (one US standard alcohol or one 12 oz serving of soda) and 2) number of drinking days per week. Three severity measures were also calculated based on Substance Abuse and Mental Health Services Administration (SAMHSA) and National Institute on Alcohol Abuse and Alcoholism (NIAAA) guidelines (National Institute on Alcohol Abuse Alcoholism 2007): 1) endorsement of a past month heavy use day (i.e., 5/4 or more drinks in a single day for men/women), 2) “heavy” drinking (i.e., 5 or more heavy drinking days/month), and 3) “at risk” drinking (i.e., more than 14/7 drinks/week or 5/4 or more drinks/typical occasion for men/women). All severity measures were dichotomously coded. Although these guidelines were developed for alcohol use and may not directly reflect heavy soda drinking criteria or at-risk soda consumption, corresponding variables were computed for soda variables to provide analogous comparisons and decrease the likelihood that the observed pattern of results was due to systematic differences in the measures used for each commodity.

Data Analysis

Non-systematic curves were identified according to standardized criteria (Stein et al. 2015). Specifically, demand curves were examined for frequent price-to-price consumption increases, reversals from zero consumption, and increased consumption with increased price as well as for extreme consumption (i.e., greater than 100 drinks in

a single day). Price elasticity and intensity were generated using the exponentiated demand equation:

$$Q = Q_0 * 10^{k*(e^{(-\alpha*Q_0*C)}-1)}$$

where Q = consumption; Q_0 = derived intensity of demand (consumption at zero price); k = a constant that denotes log consumption range (*a priori* set to 2); C = the price of the commodity; and α = derived elasticity of demand. The exponentiated model is a recently developed and validated equation that provides superior modeling for zero consumption values (Koffarnus et al. 2015; Strickland et al. 2016b). Model adequacy was evaluated by R^2 values and the relationship between derived intensity and reported "free" consumption. We focused our analyses on derived intensity and elasticity metrics to reduce type I error due to repeated testing and given that the latent structure of alcohol and cigarette demand is fully captured by demand intensity and elasticity (Bidwell et al. 2012; Mackillop et al. 2009). However, one derived measure (i.e., breakpoint or the price at which consumption dropped to zero) was also included. Breakpoint may intuitively differ from intensity and elasticity and its inclusion allowed for comparison between the selectivity of derived and observed values. Demand variables showed skew that was corrected by log-transformation prior to analysis. Pearson bivariate correlations were used to explore the relationship between alcohol and soda demand and use measures. The relationship between individual difference variables (i.e., age, sex, race, college education, and body mass index [BMI]) and commodity demand was also evaluated using bivariate correlations. A secondary analysis by mixed drink preferences was conducted by dividing participants into mixed drink favoring (i.e., rated Quite a Bit or Very Much on a mixed drink likability scale; n = 61) and non-favoring (rated Not at All, A Little Bit, or Moderately on a mixed drink likability scale; n = 78) groups. Demand curves

were generated using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA). All other analyses were conducted in SPSS Statistics 22 (IBM; Armonk, NY) with $\alpha = 0.05$.

Experiment 2a Results

Response Topography and Model Fit

Figure 4.1 depicts alcohol and soda demand fit to mean (SEM) values using the exponentiated model. Demand was characterized by prototypic decreases in consumption with increases in unit price. The exponentiated model provided an excellent fit to mean alcohol and soda demand as well as to individual data (see Figure 4.1). Model derived and observed intensities were also closely associated for alcohol (r = .95) and soda (r = .96) demand providing further support for model adequacy.

Individual Differences in Alcohol and Soda Demand

Correlations between demand variables and age, sex, race, and BMI were not statistically significant (r values = -.16 to .16). Having a college education was modestly associated with lower soda demand intensity (r = -.27, p = .001) and higher alcohol breakpoints (r = .19; p = .03).

Association Between Alcohol and Soda Demand

Correlations between alcohol and soda demand intensity (r = .21, p = .01), elasticity (r = .42, p < .001), and breakpoint (r = .49, p < .001) were all statistically significant.

Association Between Alcohol and Soda Consumption Measures

Only the cross-commodity relationship between endorsement of “more than 14/7 drinks/week or 5/4 or more drinks per typical occasion” was significant (r = .20; p = .02). All other cross-commodity consumption variables were not significantly related (r values = .02 to .12).

Alcohol and Soda Demand in Relation to Use Behavior

Table 4.2 contains correlations between demand metrics and use measures. Correlations between alcohol demand and alcohol use variables were generally

statistically significant and medium-to-large in effect size. For example, greater alcohol demand intensity was associated with more alcoholic drinks per week and days drinking per week as well as endorsement of severity measures (e.g., 5/4 or more drinks in a single day for men/women). The exception to this trend was alcohol breakpoint, which showed less robust and one non-significant association with alcohol use variables. A similar pattern of statistically significant associations was observed for soda demand and soda use variables.

Alcohol and soda demand showed excellent selectivity to the stimulus-related use variables, with no significant associations observed between alcohol demand and soda use and only one significant association between soda demand and alcohol use (soda breakpoint and alcoholic drinks per week; r = .22).

Analysis by mixed drink favorability group revealed a more robust cross-commodity correlation for demand intensity in the mixed drink non-favoring group (Favoring: Intensity r = .07; Elasticity r = .39; Breakpoint r = .52; Non-Favoring: Intensity r = .31; Elasticity r = .46; Breakpoint r = .46). Commodity-similar consumption correlations were generally similar between the two groups, with the exception of alcohol demand elasticity. Alcohol elasticity was not correlated with any alcohol consumption variables in the mixed drink favoring group. Importantly, no systematic differences for commodity-different correlations were observed, with a similar pattern of small and generally non-significant associations detected in both subgroups (only four significant correlations were observed, three of which involved the breakpoint measure; significant r values < .27).

Experiment 2a Summary

The primary aim of Experiment 2a was to demonstrate the stimulus-selectivity of alcohol and soda purchase tasks for measuring alcohol and soda demand, respectively. Modest correlations were observed for corresponding cross-commodity demand metrics

(e.g., demand elasticity for soda and alcohol) suggesting that some overlap does exist in purchasing tendencies. This similarity in demand is consistent with the idea that reinforcer sensitivity may reflect shared neurobiological and environmental risk factors related to alcohol and soda use (e.g., both may be associated with chronic stress or elevated discounting) (Bickel et al. 2012; Sinha 2008; Spillman 1990). However, metrics from each task showed a consistent and robust association with commodity-similar use variables (e.g., alcohol demand elasticity and weekly alcohol use), but not with commodity-different ones (e.g., alcohol demand elasticity and weekly soda use). Derived demand measures (i.e., demand intensity and elasticity) generally showed a more robust and selective relationship with consumption measures than the observed variable studied here (i.e., breakpoint; see General Discussion for more details). Taken together, these discriminating associations support stimulus-selectivity by showing that the stimulus or commodity under question was the primary determinant of behavior.

We observed a mostly consistent pattern of effects when participants were divided by mixed drink preferences. The exception to this trend was the lack of significant associations between alcohol elasticity and alcohol use variables in the mixed drink favoring group. Previous research has demonstrated an association between alcohol demand and combined alcohol and caffeine use as well as the unique contribution of this alcohol combination to alcohol misuse (Amlung et al. 2013). Such findings highlight the need for further study of this potentially important individual difference for alcohol use behaviors. It is important to note that we used an indirect measure of mixed drink usage (i.e., ratings of likability for mixed drinks), and therefore recommend that future research use prospective designs to evaluate the potential contribution of mixed drink use to economic demand and related variables.

In Experiment 2b, a sample of individuals reporting daily cigarette use was evaluated. The aim of Experiment 2b was to replicate previous findings showing no

relationship between chocolate demand and nicotine dependence variables (Chase et al. 2013). We also wanted to extend these findings by using an alternative sampling method (i.e., in-laboratory screening versus online crowdsourcing) as well as by evaluating the reciprocal relationship between cigarette demand and a chocolate use behavior.

Experiment 2b Methods

Participants and Procedures

Experimental procedures were identical to those reported for Experiment 2a. Briefly, participants were sampled from mTurk and required to report daily cigarette use and any chocolate use (no time period specified) to qualify for this analysis (n = 66). Although data were collected as a part of a series of parent studies on choice and drug-related cues, no participants evaluated in Experiment 2a were also included in Experiment 2b. Seven participants were removed for failing one or more attention and/or fidelity checks and 13 additional participants were removed due to non-systematic demand data, as described in Experiment 1. This resulted in a final sample size of 46 participants. See Table 4.3 for demographics and cigarette/chocolate use variables for Experiment 2.

Measures

Commodity Purchase Tasks

Cigarette and chocolate purchase tasks instructions and price range/densities were identical to those described in Experiment 2a. Hypothetical cigarettes were quantified as one preferred brand cigarette (Chase et al. 2013; MacKillop et al. 2008). Hypothetical chocolate was quantified as one Hershey Kiss size chocolate candy. This commodity size was selected given its similarity to the commodity used in a previous chocolate purchase task (Chase et al. 2013, ; Cadbury Dairy Milk Chocolate Bars) and its relevance for a United States sample. Participants completed the purchase tasks in the fixed order of cigarette purchase task before chocolate purchase task.

Cigarette and Chocolate Use Variables

Cigarette and chocolate use variables were collected as a part of a health and drug use history questionnaire. Cigarette use variables included cigarettes smoked per day and the Fagerström test for Nicotine Dependence (FTND) (Heatherton et al. 1991). The only chocolate use variable collected was typically chocolate consumed per occasion, operationalized as the number of Hershey Kiss size chocolate candies.

Data Analysis

Data analysis and evaluation of demand curves was identical to Experiment 2a. All analyses were conducted using GraphPad Prism (GraphPad Software, Inc., La Jolla, CA) and SPSS Statistics 22 (IBM; Armonk, NY) with $\alpha = 0.05$.

Experiment 2b Results

Response Topography and Model Fit

Figure 4.2 depicts cigarette and chocolate demand fit to mean (SEM) values using the exponentiated model. Demand was characterized by prototypic decreases in consumption with increases in unit price. The exponentiated model provided an excellent fit to mean cigarette and chocolate demand as well as to individual data (see Figure 4.2). Model derived and observed intensities were also closely associated for cigarette (r = .96) and chocolate (r = .93) demand providing further support for model adequacy.

Individual Differences in Cigarette and Chocolate Demand

Correlations between cigarette and chocolate demand variables and age, race, education, and BMI were not statistically significant (r values = -.27 to .21). Cigarette breakpoints were higher for men (r = .35), but no sex differences were observed for chocolate breakpoints or other demand intensity or elasticity values.

Association Between Cigarette and Chocolate Demand

Correlations between cigarette and chocolate demand intensity (r = .35; p = .02), elasticity (r = .40; p = .01), and breakpoint (r = .43; p = .003) were all statistically significant.

Association Between Cigarette and Chocolate Consumption Measures

Chocolate use was not significantly related to usual cigarettes per day (r = -.06) or FTND scores (r = .01).

Cigarette and Chocolate Demand in Relation to Use Behavior

Table 4.4 contains correlations between demand metrics and cigarette and chocolate use behaviors. Correlations between cigarette demand intensity and usual cigarettes per day (r = .39) and FTND scores (r = .52) were statistically significant and medium-to-large in effect size. Cigarette demand elasticity was associated with cigarette use variables in the expected direction, but these correlations were not statistically significant. Cigarette breakpoint was not related to cigarette use variables. Chocolate demand intensity, but not elasticity or breakpoint, was significantly associated with the chocolate use variable (i.e., typical amount of chocolate eaten per occasion).

Cigarette and chocolate demand showed acceptable selectivity to the stimulus-related use variables. Specifically, chocolate demand intensity was modestly associated with cigarette use variables, but these relationships were not statistically significant. Cigarette demand values were not associated with chocolate use.

Experiment 2b Summary

The primary aim of Experiment 2b was to replicate and extend previous research evaluating the stimulus-selectivity of cigarette and chocolate purchase tasks. Similar to Experiment 2a, moderate correlations were observed for corresponding cross-commodity demand metrics (e.g., demand elasticity for cigarette and chocolate). Satisfactory stimulus-selectivity was obtained, with significant associations observed

between some commodity-similar variables and non-significant associations observed between commodity-different variables. However, the selectivity of these relationships was not as consistent as those observed for alcohol and soda demand. For example, the relationship between cigarette demand elasticity and cigarette use frequency and severity was not statistically significant (but see Bidwell et al. 2012; MacKillop et al. 2008; Strickland et al. 2016b, for similar results). The correlations between chocolate demand intensity and cigarette use variables, although not statistically significant, were also modest in size (*r* values of .23 to .28).

It is unclear why selectivity for these cigarette and chocolate purchase tasks was less robust than for the alcohol and soda tasks in Experiment 2a, but several explanations are plausible. First, the chocolate purchase task described a very specific commodity (i.e., one Hershey Kiss size candy). Participants were instructed that they could substitute this with an alternative, but similarly sized, chocolate. However, the exactness of this commodity may have made it difficult for participants to adequately imagine their typical purchasing behavior. This potential problem with the task parameters may also explain why we observed a relatively high proportion of non-systematic data in Experiment 2b (although note that comparable exclusion rates were described in previous research) (Chase et al. 2013). Cigarettes and chocolate are also not directly comparable with respect to cost or time to consume. We used chocolate as the non-drug commodity in Experiment 2b to facilitate comparisons with previous research (Chase et al. 2013). Cigarettes and chocolate also share many of the same hedonic and purchasing qualities (e.g., typically purchased as a larger “pack” and consumed as distinct units) that should have helped improve the equivalence between these items. Second, the sample was relatively small especially compared to Experiment 2a. Observations obtained from a larger sample may have provided better estimation of the association between demand and use outcomes. We should note that the magnitude

of the relationships observed here are similar to those reported in other studies in the demand literature, including in one of the original validation studies of the cigarette purchase task (MacKillop et al. 2008). Nevertheless, the small sample size makes the results from Experiment 2 preliminary and in need of replication in additional studies. Third, we only evaluated a single, coarse measure of chocolate use and did not have a battery of frequency and severity measures as in Experiment 2a. Future research including alternative measures of chocolate use would help determine if additional measures could help clarify this discrepancy. Fourth, it is possible that the relative decrement in stimulus-selectivity observed in Experiment 2b could be due to demographic differences. Comparisons of demographics between Experiments 2a and 2b's participants did not reveal statistically significant differences; however, there was trend towards a greater percentage of participants with a college education in Experiment 2a (p = .06; all other comparisons p values > .13). These differences reflect, in part, the populations typically studied using alcohol and cigarette purchase tasks. Specifically, Experiment 2a included a sample reporting a range of alcohol use behaviors (from light to heavy use), whereas Experiment 2b was a sample more narrowly defined as daily cigarette users. Future research could focus on other cigarette-using populations (e.g., non-daily "chippers" or social cigarette users) to evaluate if sampling a range of cigarette use behaviors helps reveal improved stimulus-selectivity. These possibilities withstanding, the observation that stronger and more consistent relationships were observed between commodity-similar than dissimilar items provides modest support for the stimulus-selectivity of the cigarette and chocolate purchase tasks as described here.

General Discussion

The overall purpose of this study was to evaluate the stimulus-selectivity of drug purchase tasks. To this end, participants completed purchase tasks for drug (i.e., alcohol

or cigarettes) and non-drug comparators (i.e., soda or chocolate). Stimulus-selectivity was defined as consistent relationships between commodity-similar and not commodity-different variables. This stimulus-selectivity was examined in a double-dissociative manner by measuring demand and use behaviors for both drug and non-drug commodities. We observed robust selectivity for alcohol and soda purchase tasks and modest selectivity for cigarette and chocolate purchase tasks. These findings indicate that demand metrics likely reflect the value of or demand for only the commodity under study. Taken together, our results reinforce the fidelity of drug purchase tasks for specifically evaluating valuation of the commodity under study and support their continued use in behavioral pharmacology and addiction research.

Stimulus-selectivity was generally more consistent and robust for the equation derived (i.e., demand intensity and elasticity) than graphically observed (i.e., breakpoint) measures. This outcome suggests that model derived variables may provide a more stimulus-selective measure of demand, potentially because these metrics are generated using data encompassing the entire curve rather than from a single point (e.g., the breakpoint location). However, we must note that we did not make specific *a priori* hypotheses about observed and derived variables so these differences should be taken as preliminary and future research conducted to test this observation.

Although some discrepancies were observed, our findings are generally consistent with the outcomes reported by Chase and colleagues (2013) for cigarette and chocolate demand and extend them in at least three ways. First, we collected data using a soda purchase task and compared those metrics to data from an alcohol purchase task. Alcohol purchase tasks are one of the most widely used in the research literature making this generalization an important one (MacKillop 2016). Alcohol is also commonly evaluated in the context of other substance use and mental health disorders given its association with drug use relapse and psychiatric comorbidities (e.g., Degenhardt et al.

2001; McKay et al. 1999), highlighting the importance of its study for a variety of health behaviors.

Second, we provided explicit evidence for stimulus-selectivity by comparing demand in a reciprocal and comprehensive manner (i.e., drug demand to non-drug consumption and vice versa). These comparisons also supported the construct validity of the novel soda purchase task used in Experiment 2a. Future studies in addiction science and other health fields (e.g., nutrition) could utilize this soda purchase task to investigate soda demand as it relates to other health-related outcomes (e.g., obesity and diet). The chocolate purchase task could prove equally useful in health psychology and related fields, although further research is needed to refine and validate this task (see Experiment 2b Summary).

Finally, we collected data using online crowdsourcing as opposed to sampling methods typically used in the university laboratory setting (e.g., Chase et al. 2013; Murphy et al. 2009; but see Koffarnus et al. 2015). The use of this novel sampling method supports the generalizability of stimulus-selectivity across diverse experimental settings and populations. Importantly, alcohol and cigarette demand generally correlated with consumption variables in a way that was similar to previous studies using in-person, laboratory techniques (MacKillop et al. 2008; Murphy et al. 2009). These finding adds to the growing literature demonstrating a close correspondence between data obtained using laboratory and online methods (e.g., Johnson et al. 2015; Strickland et al. 2016a). This demonstration is important because the use of complementary in-laboratory and online studies provides an effective and efficient opportunity for the replication of experimental findings across diverse settings and samples.

Several limitations must be considered. First, these analyses were conducted as a secondary evaluation of data collected in a parent series of studies. The variables available for studying commodity use frequency and severity were therefore limited in

breadth and depth. This was a particular concern for chocolate use for which only one use variable was available. Second, a consistent price density and range was used for each purchase task. Although this range was consistent with those used in other purchase task studies (Jacobs and Bickel 1999; MacKillop et al. 2010a), more recently researchers have elected to remove extreme prices from the price range (Murphy et al. 2015). Similarly, although the specific instructions used in these tasks were similar to those used elsewhere, they did differ in some respects from some studies evaluating the psychometric properties of alcohol and cigarette demand (e.g., framing the event as a weekend party versus as a "typical day" here) (Murphy et al. 2009). Nevertheless, the high density of prices in the initial portion of the range likely provided sufficient coverage across the elastic and inelastic portions of the demand curve and allowed for accurate estimation of demand intensity and elasticity.

Third, the order of completion was not randomized and all participants completed drug purchase tasks prior to non-drug purchase tasks. Few studies have evaluated demand across multiple commodities, and those that exist either have not clearly indicated if counterbalancing was used or, if it was, if an order effect was observed (Chase et al. 2013; Jacobs and Bickel 1999; Pickover et al. 2016; Strickland et al. 2016b). One of these studies was completed by our research laboratory and included both cigarette and alcohol purchase tasks. Analysis of these data for possible order effects indicated that order of completion (i.e., alcohol before cigarette purchase task or vice versa) did not influence the magnitude of alcohol or cigarette demand intensity or elasticity observed in that study (data not reported in the original report) (Strickland et al. 2016b). The use of repeated and specific instructions prefacing each purchase task could have also lessened the potential for order effects. Namely, participants were provided a detailed overview of the commodity available prior to completion in each task to ensure awareness of the operational parameters. Nevertheless, future studies should

include a randomized order to test if order of completion influences the stimulus-selectivity of purchase tasks.

Fourth, soda and chocolate were chosen as the non-drug comparators for alcohol and cigarettes given general similarities in use topography, qualitative appearance, and typical serving size. Our focus was on unhealthy commodities given that these items were expected to show the closest relationship with drug demand and provide a more rigorous test of stimulus-selectivity than healthier consumables (e.g., fruit). We attempted to equate all commodities in some respect by allowing participants to purchase their "preferred brands". However, differences in the type (e.g., gin, beer, regular, diet), container (e.g., glass, can), and brand (e.g., Coca Cola®, Pepsi®) used may have influenced decision-making. Nevertheless, such variation is inherent to the stimulus qualities and selectivity of commodity purchase tasks to the item under question and as such should not be considered problematic for the present study. We also did not consider the status of soda and chocolate as economic substitutes or complements for alcohol or cigarettes, respectively. A recent study suggests that fast food items are not economic substitutes for cigarettes, whereas cigarettes are a modest complement for food (Murphy et al. 2016). It is unlikely that substitutes or complements affected the pattern of results reported here given that all purchase tasks were completed as independent commodities without reference to other drug or non-drug items. However, these economic mechanisms are a critical area for future research given their importance for the allocation of behavior away from undesired drug use to desired alternatives activities. Fifth, drug use could not be biologically verified and experimental control was not guaranteed in the online setting. We used several techniques to help increase data quality (e.g., attention checks) and, as noted above, demand and consumption correlations were generally consistent with the previous literature. Finally, we must emphasize that these analyses represent a preliminary study of the stimulus-

selectivity of drug purchase tasks given the limited scope and small sample size in Experiment 2. Future research is needed to replicate these and other experimental findings to support the validity of drug purchase tasks across a variety of experimental conditions (e.g., study setting; drug and non-drug commodity types) and populations (e.g., recreational users; treatment-seeking participants).

Despite these limitations, the current study provides preliminary evidence supporting the stimulus-selectivity of commonly used drug purchase tasks. As the use of drug purchase tasks in behavioral research proliferates, it is critical that research continue to address the reliability, validity, and fidelity of these procedures. Such methodological and parametric studies will help reinforce the capacity of purchase tasks and econometric analyses for revealing behavioral mechanisms underlying drug-taking behavior and help encourage the use of best practice methods in health and addiction science.

Table 4.1. Experiment 2a Participant Demographics and Alcohol/Soda Use Behaviors.

	Median/%	IQR
Demographics		
Age	31	26–39
Male	45.3%	
White	74.8%	
College Education	64.0%	
BMI	26.1	23.0–32.7
Alcohol Use		
Drinks/Week	4	1–10
Days/Week	2	1–3
Past Month Day with ≥5/4 Drinks	59.0%	
≥5 Past Month Days with ≥5/4 Drinks	20.1%	
>14/7 Drinks/Week or ≥5/4 Drinks/Usual Occasion	40.3%	
Soda Use		
Drinks/Week	3	1–10
Days/Week	2	1–7
Past Month Day with ≥5/4 Drinks	23.7%	
≥5 Past Month Days with ≥5/4 Drinks	10.8%	
>14/7 Drinks/Week or ≥5/4 Drinks/Usual Occasion	23.7%	

Note. IQR = interquartile range; BMI = body mass index; all divided criteria (e.g., 5/4) refer to separate criteria for men/women, respectively

Table. 4.2 Association Between Demand and Alcohol and Soda Use Measures.

	Drinks/ Week	Days/ Week	Past Month Day with ≥5/4 Drinks	≥5 Past Month Days with ≥5/4 Drinks	>14/7 Drinks/ Week or ≥5/4 Drinks/ Usual Occasion
Demand			Alcohol Outcomes		
Alcohol					
Q_0	**.48**	**.39**	**.52**	**.44**	**.48**
α	**-.28**	**-.31**	**-.29**	**-.21**	**-.32**
BP	**.20**	**.18**	**.17**	.10	**.17**
Soda					
Q_0	.04	-.01	.01	-.05	<.01
α	-.09	-.07	-.03	.05	-.06
BP	**.22**	.08	.10	.02	.12
			Soda Outcomes		
Alcohol					
Q_0	<.01	.06	.05	.06	.08
α	.04	.04	.03	.02	.03
BP	-.09	-.09	-.05	-.07	-.12
Soda					
Q_0	**.52**	**.45**	**.57**	**.43**	**.50**
α	**-.43**	**-.39**	**-.39**	**-.34**	**-.43**
BP	**.30**	**.30**	**.24**	**.17**	**.30**

Note. Q_0 = demand intensity from the exponentiated demand equation; α = demand elasticity from the exponentiated demand equation; BP = breakpoint; all divided criteria (e.g., 5/4) refer to separate criteria for men/women, respectively. **Bold** = statistically significant correlation.

Table 4.3. Experiment 2b Participant Demographics and Cigarette/Chocolate Use Behaviors.

	Median/%	IQR
Demographics		
Age	34	28-42
Male	54.3%	
White	80.4%	
College Education	47.8%	
BMI	27.7	23.8-34.2
Cigarette Use		
CPD	10	6-19
FTND	4	1-6
Chocolate Use		
Chocolate/Occasion	4	3-6

Note. IQR = interquartile range; BMI = body mass index; CPD = cigarettes/day; FTND = Fagerström Test for Nicotine Dependence.

Table 4.4. Association Between Demand and Cigarette and Chocolate Use Measures.

	Cigarettes		Chocolate
	CPD	FTND	Chocolate/ Occasion
Cigarettes			
Q_0	**.52**	**.39**	.01
α	-.17	-.21	.05
BP	.01	.06	-.02
Chocolate			
Q_0	.23	.28	**.32**
α	.08	-.01	-.17
BP	-.06	<.01	-.01

Note. Q_0 = demand intensity from the exponentiated demand equation; α = demand elasticity from the exponentiated demand equation; BP = breakpoint; CPD = cigarettes/day; FTND = Fagerström Test for Nicotine Dependence. **Bold** = statistically significant correlation.

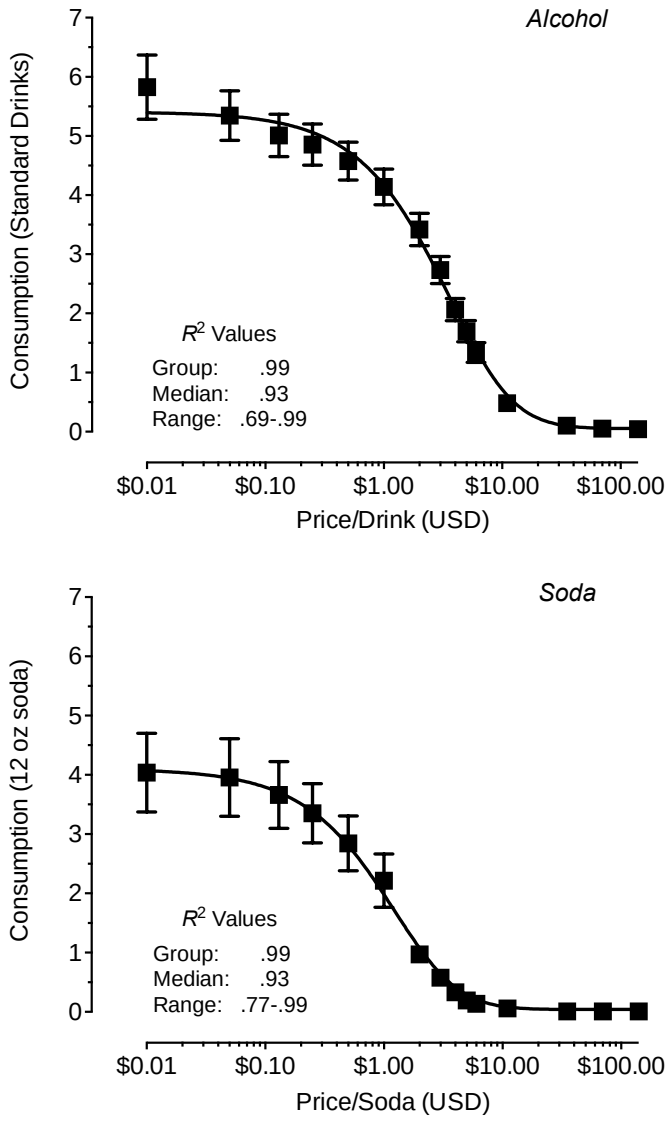

Figure 4.1. Economic demand for alcohol (top panel) and soda (bottom panel). Participants (n = 139) completed commodity purchase tasks in which hypothetical alcohol (one US standard drink) or soda (one 12 oz soda) were available. Price varied in United States dollars (USD). Plotted are mean (SEM) group data on a log-linear axis fit using the exponentied model. Also included are group R^2 values for model fit as well as median and ranges for individual data.

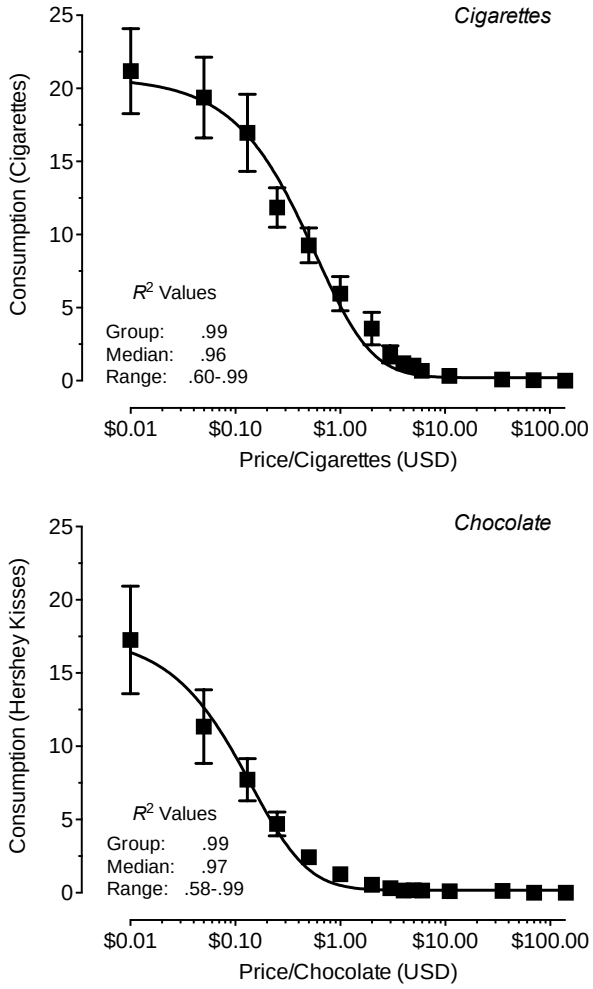

Figure 4.2. Economic demand for cigarettes (top panel) and chocolate (bottom panel). Participants (n = 46) completed commodity purchase tasks in which hypothetical cigarettes (one preferred brand cigarette) or chocolate (one Hershey Kiss size chocolate) were available. Price varied in United States dollars (USD). Plotted are mean (SEM) group data on a log-linear axis fit using the exponentied model. Also included are group R^2 values for model fit as well as median and ranges for individual data.

Chapter 5

EVALUATING NON-MEDICAL PRESCRIPTION OPIOID DEMAND USING COMMODITY PURCHASE TASKS: TEST-RETEST RELIABILITY AND INCREMENTAL VALIDITY

(Experiment 3; Strickland et al., under review)

Introduction

The non-medical use of prescription opioids and opioid use disorder (OUD) present a significant and growing public health concern in the United States. Over 2 million people reported initiation of non-medical prescription opioid use in 2017 and over 11 million reported past year use (Center for Behavioral Health Statistics 2018). A steady rise in the rates of overdose fatalities attributable to prescription opioids has also occurred, with a four-fold increase since 1999 (Hedegaard et al. 2017). Improvements in monitoring systems and pill reformulations have shown some promise for deterring use, but in many at-risk populations (e.g., rural Appalachian regions) rates of opioid overdose and use-related burden remain high (e.g., Brown et al. 2018; Mack et al. 2017; Schranz et al. 2018; Van Handel et al. 2016). One research priority then is to identify behavioral mechanisms underlying OUD and this persistence of use. Such research will ultimately aid the development of novel and improved prevention and intervention approaches.

The merger of theoretical perspectives from behavioral economics and operant theory has resulted in numerous advances for psychological science, broadly (Hursh and Roma 2013), and addiction science, specifically (Bickel et al. 2017; MacKillop 2016). These theoretical accounts broadly propose that systematic choice and decision-making processes described by behavioral economic theory may help to reveal the behavioral mechanisms contributing to the development and persistence of substance use disorders. For example, a reinforcer pathology approach posits that substance use

disorder is characterized by high reinforcer valuation combined with an extreme preference for immediate reinforcers (Bickel et al. 2017).

Behavioral economic demand (i.e., the relationship between commodity price and purchase) has received particular attention under a behavioral economic framework as a measure of reinforcer valuation. Demand analysis presents several advantages over traditional measures of relative reinforcer value, including accounting for the multi-dimensional nature of reinforcement rather than treating reinforcement as a homogenous construct (Johnson and Bickel 2006; Hursh and Silberberg 2008). The study of demand has been facilitated, in part, by the development of the commodity purchase task procedure (Jacobs and Bickel 1999; see reviews by Kaplan et al. 2018; MacKillop 2016). Participants are asked to report hypothetical consumption of a good across varying prices in this procedure in order to effectively and efficiently generate demand curves for analysis. That these simulated procedures can be completed in the absence of active drug administration also affords the opportunity to work with populations that cannot be evaluated using drug self-administration procedures (e.g., treatment-seeking individuals, individuals with compromised health). A growing body of literature has supported the clinical relevance of demand as measured by the purchase task procedure by using demand to understand mechanisms by which interventions are clinically effective or as a prognostic variable predicting reductions in substance use following intervention delivery (e.g., Bujarski et al. 2012; Murphy et al. 2015; MacKillop and Murphy 2007).

To date, the majority of research on behavioral economic demand in the human laboratory or clinic has studied alcohol and tobacco cigarette use. A smaller body of research has examined prescription opioid use within this emerging framework (Jacobs and Bickel 1999; Pickover et al. 2016). Those studies that exist have focused on samples drawn from more narrowly-defined populations and have evaluated demand in

the absence of other behavioral economic measures. For example, Jacobs and Bickel (1999) evaluated heroin demand in individuals from an outpatient opioid clinic and found that hypothetical heroin consumption was well described by quantitative demand models. Pickover and colleagues (2016) measured demand for non-medical prescription drugs among college students and found that demand was predictive of opioid use frequency and OUD diagnosis. Both of these studies emphasized opioid demand, however, which precludes the determination of a unique and commodity-specific contribution of demand to an understanding of the behavioral mechanisms underlying patterns of illicit prescription opioid use.

The overall purpose of the current study was to replicate (Experiment 3a) and extend (Experiment 3b) prior work on the use of the purchase task procedure to evaluate behavioral economic demand for prescription opioids (Jacobs and Bickel 1999; Pickover et al. 2016). Experiment 3a was designed to replicate prior work by demonstrating the utility of the purchase task procedure to describe prescription opioid demand. To this end, adult participants reporting past year non-medical prescription opioid use completed a purchase task for prescription opioids. We hypothesized that prescription opioid demand would be well described by quantitative models of demand and would be related to OUD consistent with prior work (Jacobs and Bickel 1999; Pickover et al. 2016).

Experiment 3a Methods

Participants and Screening

Participants were recruited using the crowdsourcing site Amazon Mechanical Turk (mTurk) as a part of a larger study on reinforcement learning. Crowdsourcing utilizes the Internet to sample a large number of individuals from varied geographic regions and with varied health histories. Prior research has demonstrated the validity of using crowdsourcing to sample participants in psychological and addiction science (see

reviews by Chandler and Shapiro 2016; Strickland and Stoops 2019; see further description in the General Discussion).

Inclusion criteria were: 1) past year non-medical prescription opioid use, 2) 20 or more lifetime prescription opioid uses, and 3) age 18 or older. Inclusion criteria were verified using a short screening questionnaire. Access to the screening survey was limited to individuals with at least 50 completed mTurk tasks, a ≥ 95% approval rating on prior tasks, and United States residence (see similar qualifications in Cunningham et al. 2017; Strickland and Stoops 2015). Participants meeting inclusion criteria were then directed to the full survey containing the opioid purchase task and opioid use measures. The University of Kentucky Medical Institution Review Board approved all procedures and participants reviewed an informed consent document prior to participation.

Behavioral Economic Demand

Behavioral economic demand for prescription opioids was evaluated using a commodity purchase task (Pickover et al. 2016). A standard instructional vignette was provided in which participants were instructed to consume all purchases in a single day, could not stockpile, could only get the commodity from this source, and had no commodity available from previous days. Understanding of these instructions was verified by a required correct response to two questions prior to advancing. The commodity available was "the standard dose that you use when you use these pills" consistent with prior research (Pickover et al. 2016). Purchases were evaluated across 13 monetary increments ranging from $0.00 [free] to $11/pill, presented sequentially (full range: $0.00 [free], $0.01, $0.05, $0.13, $0.25, $0.50, $1, $2, $3, $4, $5, $6, $11).

Data from commodity purchase tasks were analyzed using the exponentiated demand equation (Koffarnus et al. 2015):

$$Q = Q_0 * 10^{k(e^{(-\alpha * Q0 * C)} - 1)}$$

where Q = consumption; Q_0 = derived demand intensity; k = a constant related to consumption range (*a priori* set to 2); C = commodity price; and α = derived demand elasticity. Demand intensity refers to a theoretical consumption of a commodity at a unit price of zero or near-zero. Demand elasticity refers to the sensitivity of consumption to changes in price. We focused our analyses on intensity and elasticity given that prior factor analytic studies have demonstrated improved stimulus-selectivity when using derived measures (Strickland and Stoops 2017) and that these two measures reflect the two-factor structure underlying purchase task data (Aston et al. 2017; Bidwell et al. 2012; Epstein et al. 2018; Mackillop et al. 2009). Intensity and elasticity were log-transformed to achieve normality.

Data Analysis

Fifty-one participants met the above inclusion criteria and completed the study measures. Six failed one or more attention or validity checks and were removed from data analysis. An additional five provided non-systematic purchase task data according to standardized criteria (Stein et al. 2015). This resulted in a final sample for analysis of 40 participants.

Bivariate associations were evaluated between demand outcomes and opioid use and demographic variables. Pearson correlations were used in most cases, however negative binomial regression was used for past month use days given the observation of zero-inflation in this variable. SPSS Statistics (IBM; Armonk, NY) was used for analyses. All inferential tests were two tailed and used an alpha rate of .05.

Experiment 3a Results

Demographics and Opioid Use

Table 5.1 contains demographic and opioid use information for participants in Experiment 3a. A majority of participants were white and female with an average age of

35.5. Half endorsed statements indicative of *DSM-IV* criteria for opioid dependence and a fifth reported a preference for a risky route of opioid administration (i.e., intranasal, smoked, or injected versus oral).

Opioid Behavioral Economic Demand

Opioid demand was well characterized by decreases in consumption with increases in price (Figure 5.1). Good model fits were also observed with the exponentiated demand equation for individual participant data (median R^2 = .84; IQR = .80 to .94).

Bivariate Relationships

Bivariate relationships between opioid demand and demographic and opioid use variables are presented in Table 5.2. Opioid dependence was significantly associated with higher opioid demand intensity, r = .43, p = .006. Figure 5.2 plots this relationship involving opioid dependence for group mean demand curves. More inelastic opioid demand was also significantly associated with more past month opioid use days, RR = 0.51, p = .03.

Experiment 3a Discussion and Experiment 3b Aims

Experiment 3a found that behavioral economic demand for prescription opioids was well described by quantitative models of demand and was systematically related to OUD. Specifically, individuals with OUD showed more intense demand as well as a trend towards more inelastic demand. These findings are consistent with work previously conducted in college students (Pickover et al. 2016) and patients from an outpatient opioid clinic (Jacobs and Bickel 1999) thereby demonstrating that the utility of the purchase task procedure generalizes across research and clinical contexts.

Experiment 3b was designed to advance this prior work in several ways. First, a battery of behavioral economic measures, including commodity purchase tasks and delay discounting tasks, were included to evaluate the unique and stimulus-selective contribution of opioid demand to OUD. Measures of alcohol demand under varying

environmental contexts (i.e., drink price specials) were also included to further establish the validity of data collection (i.e., to replicate previous associations described in other clinical populations). We hypothesized that prescription opioid demand would be related to opioid use measures in a unique and commodity-selective manner above and beyond other measures of demand and delay discounting. Second, opioid demand measures were collected at two time points separated by approximately one month to establish test-retest reliability and temporal stability. We hypothesized that opioid demand would show acceptable test-retest reliability consistent with purchase task procedures for other substances (Acuff and Murphy 2017; Few et al. 2012; Murphy et al. 2009)

Third, cross-commodity tasks were included to determine the behavioral economic relationship between opioid and cannabis use. Cross-commodity or cross-price demand refers to the responsiveness of the quantity demand for a good to changes in price of another good. Commodities may function as a substitute meaning that as the price increases for the price-manipulated good that consumption increases for the alternative (i.e., positive cross-commodity elasticity). Commodities may alternatively function as a complement meaning that as the price increases for one good that consumption decreases for the alternative (i.e., negative cross-commodity elasticity). The relationship between opioid and cannabis was evaluated given suggested similarities in the behavioral response and neurobiological pathways associated with pain, which has led to a proposed substitution of cannabis for prescription opioids in the medical management of chronic pain (see discussion of this issue in Choo et al. 2016; Hill 2015; Lucas 2012). We hypothesized that cannabis and opioids would function as economic substitutes given this putative clinical relationship.

Experiment 3b Methods

Participants and Screening

General recruitment procedures were similar to Experiment 3a. Inclusion criteria for this study were 1) past year non-medical prescription opioid use, 2) 30 or more lifetime prescription opioid uses, and 3) age 18 or older. Participants were also asked to complete a follow up survey approximately one month after the initial survey, which contained the same purchase task and delay discounting measures. No participants were repeated from Experiment 3a.

Measures

Single-Commodity Purchase Tasks

Behavioral economic demand for prescription opioids and cannabis was evaluated using commodity purchase tasks. The instructional set was identical to Experiment 3a. Cannabis hits were quantified as 10 hits/joint with 1 joint equal to 0.9 g cannabis (~0.09 g/hit) (Aston et al. 2015; Strickland et al. 2017b). The price range used in Experiment 3b was expanded and included 17 prices from $0.00 [free] to $20/unit, presented sequentially (full range: $0.00 [free], $0.25, $0.50, $1, $1.50, $2, $2.50, $3, $4, $5, $6, $7, $8, $9, $10, $15, $20).

Commodity purchase tasks were also used to evaluate the effect of drink price specials on alcohol demand (Kaplan and Reed 2018). Participants completed one of two purchase tasks for alcohol in which drinks were either regularly priced (no special) or under a buy one get one free special (BOGO). All drinks referred to one standard drink (one 12 oz beer, 5 oz wine, or 1.5 oz shot of liquor alone or in a mixed drink). Only one purchase task was completed at each time point and was randomized for each participant. Consumption was converted to standard drinks for comparative analysis purposes (Kaplan and Reed 2018).

All commodity purchase task data were analyzed using the exponentiated demand equation as described above. Demand intensity and elasticity were log-transformed to achieve normality.

Cross Commodity Demand

Cross-commodity purchase tasks were used to evaluate the behavioral economic relationship between prescription opioid and cannabis price. Cross-commodity tasks were developed from previous work (e.g., Amlung et al. 2019; Johnson et al. 2017b; Peters et al. 2017). Tasks were generally identical to single-commodity demand tasks expect that 1) two commodities were concurrently available in each task (i.e., prescription opioid pills and cannabis hits) and 2) two questions were presented (i.e., consumption for each commodity). One commodity was price-fixed and the other price-manipulated for each task. The price-manipulated commodity followed the same price sequence as single-commodity tasks. Price-fixed opioids were set at $3.00/pill and price-fixed cannabis was set at $0.50/hit. These prices were selected because they represented similar areas of transition from inelastic to elastic demand for each commodity for most participants in previous studies (Experiment 3a; Strickland et al. 2017b). Tasks were presented in a randomized order.

Analysis of cross-commodity demand was conducted in two ways. First, linear regression was performed between log consumption data and log price to estimate cross-price elasticity for the fixed-price commodity over the entire demand curve (Johnson et al. 2017b; Peters et al. 2017; Quisenberry et al. 2017; Stein et al. 2018a). Cross-price elasticity was also evaluated at each price change along the curve using the formula (Allison 1983; Petry and Bickel 1998):

$$E_{Cross} = \left[\log\left(Q_{fixed2}\right) - \log\left(Q_{fixed1}\right)\right] / \left[\log\left(Price_{manipulated2} - \log(Price_{manipulated1}\right)\right]$$

Second, cross-commodity demand curves were fit using the cross-price elasticity equation provided by Hursh and Roma (2013):

$$Q_B = \log(Q_{alone}) + Ie^{-\beta Price_A}$$

Where Q_{alone} equals consumption of the fixed price commodity (B) at infinite price of the varying price commodity A, *I* is an interaction constant, β is the sensitivity of commodity B consumption to price of commodity A, and P_a is the price of the price-varying commodity A. The interaction term (*I*) reflects the relationship between A and B with negative terms indicating a substitute good and positive terms reflecting a complement good. Demand curves for the price-manipulated commodity in these cross-commodity tasks were also compared to the single-commodity task for evidence of changes in purchasing behavior in the presence of a concurrent commodity.

5-Trial Adjusting Delay Task

Delay discounting rates for money, cannabis, and opioids were determined using a 5-trial adjusting delay task (for task details see Koffarnus and Bickel 2014). Participants made five choices between an immediate, smaller reinforcer ($500/$500 of opioids/$500 of cannabis now) and a delayed, larger reinforcer ($1000/$1000 of opioids/$1000 of cannabis delayed) at delays that titrated up or down based on prior selections. This task was selected for its prior utility in an online setting (e.g., Stein et al. 2017; Strickland et al. 2017b) and validation against traditional longer test forms (Cox and Dallery 2016; Koffarnus and Bickel 2014). Delay discounting rates were log-transformed to achieve normality.

Brief DSM-5 Substance Use Disorder Diagnostic Assessment

DSM-5 substance use disorder was evaluated using an adapted version of the Brief *DSM-5* Diagnostic Assessment (Hagman 2017). This questionnaire evaluated each of the 11 *DSM-5* criteria for alcohol, cannabis, and opioid use disorders. Prior research has demonstrated the internal reliability and validity of this assessment for alcohol use disorder (Hagman 2017). Diagnostic categories were determined using *DSM-5* criteria (i.e., 2-3 = mild; 4-5 = moderate; 6+ = severe substance use disorder).

Brief Pain Inventory

A modified version of the brief pain inventory was used to evaluate chronic and current pain (Mendoza et al. 2006). Participants were asked to indicate taking all pain into account the 1) average past week pain levels (0-10 scale), 2) average interference from pain across common daily activities (0-10 scale), and 3) typical relief from pain when using prescription opioids (0-100 scale).

Data Analysis

One hundred and five participants met the above inclusion criteria and completed the time 1 survey. Six failed one or more attention or validity checks and were removed from data analysis. An additional 16 provided non-systematic opioid purchase task data according to standardized criteria (Stein et al. 2015). This resulted in a final primary sample for analysis of 83 participants. Sixty-five of these participants completed the time 2 assessment (78.3%). Measures involving cannabis use (i.e., cannabis purchase task, cannabis discounting, and cross-commodity tasks) were only completed by individuals reporting past year cannabis use (n = 76; 91.6%).

Bivariate relationships were evaluated as in Experiment 3a. Significant outcomes were then followed up with multivariable models evaluating the incremental and unique association for opioid behavioral economic variables controlling for demographic variables (i.e., age, sex, education, and income) and opioid use frequency (i.e., past month opioid use). Test-retest reliability was determined using bivariate correlations comparing time 1 and time 2 values. Temporal stability for demand, discounting, and *DSM-5* substance use disorder values were evaluated using dependent-samples *t*-tests or McNemar tests for paired nominal data. Alcohol demand under drink special conditions and across time was evaluated using linear mixed effect models. Finally, cross-commodity variables were determined as described above (see *Cross Commodity*

Demand). Group mean cross-commodity demand was evaluated to address quantitative issues concerning zero consumption values (e.g., requiring value replacement for log transformation) (see similar approaches in Amlung et al. 2019; Quisenberry et al. 2017). SPSS Statistics (IBM; Armonk, NY) and R statistical analysis (R Core Team, 2018) were used for analyses. All inferential tests were two tailed and used an alpha rate of .05.

Experiment 3b Results

Demographics and Opioid Use

Table 5.3 contains demographic and opioid use information for participants in Experiment 3b. A similar demographic composition was observed as in Experiment 3a with a majority of participants being white and female with an average age of 34.0. Current chronic pain was endorsed by three-quarters of participants (74.7%) with an average pain level of 4.6 on a 0-10 scale.

Two-thirds of participants endorsed statements indicative of *DSM-5* criteria for opioid use disorder and a quarter reported a preference for a risky route of opioid administration (i.e., intranasal, smoked, or injection versus oral). The majority of participants also reported past year cannabis use (91.6%) and 41% met criteria for cannabis use disorder.

Behavioral Economic Demand and Delay Discounting

Opioid and cannabis demand were well characterized by decreases in consumption with increases in price (Figure 2). Good model fits were also observed with the exponentiated demand equation for individual opioid (median R^2 = .86; IQR = .79 to .93) and cannabis (median R^2 = .90; IQR = .84 to .97) data. Money delay discounting rates were shallower than for opioids, $t_{82} = 6.33$, $p < .001$, $d_z = 0.69$, or alcohol, $t_{75} = 4.16$, $p < .001$, $d_z = 0.48$, which did not significantly differ from each other, $t_{75} = 0.64$, $p = .52$, $d_z = 0.07$.

Table 4 contains bivariate associations and significance for Experiment 2. A preference for risky opioid routes was associated with less elastic opioid demand. Cannabis demand was also significantly associated with cannabis use frequency and quantity variables. OUD was significantly associated with more intense and less elastic opioid demand (see Figure 2 for group mean plots). Similarly, cannabis use disorder was associated with more intense cannabis demand. Neither opioid nor cannabis demand were associated with the other substance use disorders (e.g., alcohol or opioid use disorder for cannabis demand). Steeper opioid discounting rates were also associated with OUD.

Average pain and the typical impact of pain on everyday life were both associated with more intense opioid demand and steeper monetary and opioid discounting rates. In contrast, typical pain relief from opioids was not related to any demand or discounting variables.

Multivariable Models

Multivariable models including opioid demand and monetary and opioid discounting rates were conducted to test incremental and unique associations with opioid use disorder, risky opioid route preference, and average pain (i.e., variables with significant bivariate associations) controlling for demographic variables and opioid use frequency. Higher opioid intensity (OR = 31.30, p =.004) and higher opioid discounting rates (OR = 7.46, p = .018) were each significant and independent predictors of OUD in multivariable models. More inelastic opioid demand was significantly associated with risky opioid route preferences (OR = 0.07, p = .003) and greater opioid demand intensity was significantly associated with higher average pain levels (β = .31, p = .027). Other behavioral economic variables were not significant in multivariable models.

Test-Retest Reliability

Good test-retest reliabilities were observed for opioid demand (Q_0 r_{xx} = .75; α r_{xx} = .63) and cannabis demand (Q_0 r_{xx} = .53; α r_{xx} = .58) (Figure 3). Temporal reliabilities were also acceptable and significant for discounting rates with lower reliability for money compared to commodity discounting (money r_{xx} = .42, opioid r_{xx} = . 58, cannabis r_{xx} = .61). These values were temporally stable with no significant changes in demand or discounting values from time 1 to time 2, *p* values > .05.

Temporal reliability was also good for scores on the Brief *DSM-5* Substance Use Disorder Diagnostic Assessment for OUD, r_{xx} = . 76, as well as for cannabis and alcohol use disorder, r_{xx} = .63 and .77, respectively. Substance use disorder classifications were stable over the one-month period as indicated by non-significant McNemar tests for paired nominal data.

Alcohol Demand and Sensitivity to Environmental Influences

Group mean curves for alcohol demand across time and by BOGO special are plotted in Figure 5.4. Linear mixed effect models indicated a significant effect of BOGO special for demand intensity, *b* = 0.10, *p* = .007, reflecting higher intensity with BOGO specials. No effect of BOGO special was observed for elasticity, *b* = 0.07, *p* = .23. No effects of time were observed reflecting no change in overall alcohol demand from time 1 to time 2, *b* values < 0.01, *p* values > .80.

Cross-Commodity Demand

Figure 5.5 contains group mean cross-commodity demand curves for opioid and cannabis consumption as well as group mean curves for own-price demand alone and in the presence of the concurrent commodity.

Changes in Own-Price Demand

Analysis of individual demand curves indicated decreased demand intensity, t_{73} = 2.61, *p* = .011, d_z = .30 and increased demand elasticity, t_{73} = 2.40, *p* = .019, dz = .28,

for opioids when cannabis was concurrently available. Similarly, decreased demand intensity, t_{62} = 4.17, p < .001, d_z = .53, and increased demand elasticity, t_{62} = 2.64, p =.01, d_z = .33, for cannabis was observed when opioids were concurrently available.

Cross-Price Elasticity

Cross-price elasticity was first evaluated by determining the regression slopes of log mean demand on log other product price. This analysis revealed a cross-price elasticity of -0.11 (95% CI [-0.13, -0.09]) for opioids and 0.01 (95% CI [-0.02, 0.03]) for cannabis. Inspection of 95% confidence intervals showed that the opioid estimate was statistically significant indicating that opioids functioned as a weak complement for cannabis. The cannabis cross-price elasticity was not different from zero reflected by a zero-slope line and indicating that cannabis was an independent commodity. Cross-price elasticity values over each price change along the demand curve are presented in Table 5.5. Visual inspection of these values was consistent with the above analyses. Specifically, cannabis did not show a consistent substitute or complement relationship with opioids across the curve. Prescription opioids functioned as a weak complement over the entire cross-elasticity curve reflected by an overall negative slope.

Cross-price elasticity was then evaluated using the cross-price elasticity equation proposed by Hursh and Roma (2013). This equation provided an excellent fit to mean cross-price opioid data (R^2 = .97) and indicated a positive interaction term whose 95% confidence interval did not overlap zero (I = 0.21, 95% CI [0.19-0.24]). Mean cross-price cannabis data were not well described by the cross-price elasticity equation (R^2 = .29), likely due to fluctuating values around a zero slope, and a zero-value interaction term was observed (I = 0.03 [-0.22, 0.27]). These results were consistent with the interpretation of the log-log slope analyses above.

General Discussion

The overall purpose of this study was to evaluate the utility of the purchase task

procedure for describing non-medical prescription opioid use. Participants across two independent experiments reported prescription opioid demand that was systematically associated with OUD whether measured using *DSM-IV* or *DSM-5* criteria. That elevated opioid demand was related to diagnostically relevant opioid use is consistent with prior work conducted in college student samples (Pickover et al. 2016) and indicates that this relationship is replicable and generalizes to a general adult population. This association was also selective to opioids in that opioid demand variables were associated with OUD and not cannabis or alcohol use disorders. A similar selectivity was observed for cannabis wherein cannabis consumption and cannabis use disorder, but not other substance use variables, were associated with cannabis demand. These findings contribute to a growing body of literature demonstrating the stimulus selectivity of the purchase task procedure for indexing valuation that is specific to the substance of interest (Chase et al. 2013; Strickland and Stoops 2017). More broadly, these findings indicate that more intense and inelastic demand could be behavioral mechanisms underlying the progression to diagnostically relevant use among a broader population of individuals reporting non-medical prescription opioid use. Future longitudinal work will be important for establishing the causal relationship between variations in opioid demand and the development of OUD.

Opioid Demand and Pain

Average self-reported pain and level of interference in daily function were associated with increased opioid demand intensity even after controlling for other relevant demographic and behavioral economic variables. To our knowledge, this is the first study to describe a relationship between drug demand and pain. This relationship between pain and the relative intensity of non-medical prescription opioid use is consistent with the notion that self-medication of un- or under-managed chronic pain could contribute to problematic opioid use, but contrasts with the observation that

perceived pain relief from prescription opioids was not related to opioid demand. The discrepancy between these two outcomes could signify a decoupling between the level of opioid intake and strength of opioid relief due to processes such as pharmacological tolerance or an ineffective targeting by opioids of underlying causes of chronic pain (e.g., Arner and Meyerson 1988; Ashburn and Staats 1999).

Unique Prediction by Behavioral Economic Variables

Multivariable models indicated that behavioral economic demand provided unique and incremental information about OUD above and beyond delay discounting rates and frequency of opioid use. These models specified that higher opioid demand intensity and steeper opioid delay discounting rates each significantly and uniquely predicted the presence of OUD. This finding that demand accounted for unique information about OUD provides evidence for distinct behavioral mechanisms that could underlie clinically relevant non-medical opioid use. This evidence is consistent with previous work demonstrating the relationship of demand (Pickover et al. 2016) and discounting (e.g., Kirby et al. 1999; Kirby and Petry 2004) with heroin and prescription opioid use when measured alone. These findings are also in accordance with the predictions of reinforcer pathology theory, which posits that substance use disorder is associated with high reinforcer valuation (i.e., demand) combined with an extreme preference for immediate reinforcers (i.e., discounting) (Bickel et al. 2017). A uniquely predictive relationship involving opioid demand also supports incremental validity insofar as relevant information about OUD was offered above and beyond another significant and relevant behavioral economic variables (i.e., delay discounting) and the frequency of opioid use.

Adjusted models also revealed that opioid demand elasticity was a significant predictor of a preference for risky routes of administration and was unique among behavioral economic variables in this regard. Intranasal, smoked, and intravenous routes of opioid administration are associated with increased health risks, such as STI

transmission and overdose (Conrad et al. 2015; Strathdee and Beyrer 2015). The transition from oral to non-oral routes of administration also represents an important risk factor for the initiation of heroin and other illicit substance use (Carlson et al. 2016; Compton et al. 2016; Young and Havens 2012). Continued use in the face of these putative health consequences is consistent with the association reported here in which a preference for non-oral routes was related to more inelastic opioid demand. Such a relationship suggests that these preferences may be mechanistically related to a decreased sensitivity to the costs and consequences of substance use as reflected by less sensitive changes in use to increases in unit price (i.e., more inelastic demand).

Test-Retest Reliability of Opioid Use Behavioral Measures

Good support for the reliability of opioid demand intensity (r_{xx} = .75) and elasticity (r_{xx} = .63) were observed over one month of testing. These reliabilities are similar to those for alcohol demand when measured over a one-month period in college students (intensity r_{xx} = .67, elasticity r_{xx} = .71) (Acuff and Murphy 2017). Reliabilities were also generally acceptable for cannabis demand (Q_0 r_{xx} = .53; α r_{xx} = .58), albeit lower than those for prescription opioids. This is the first study to evaluate the temporal reliability of purchase tasks for substances other than alcohol or cigarettes. The temporally stability of these tasks supports a continued use in repeated measure or longitudinal designs of laboratory and clinic research.

Clinical classifications based on the Brief *DSM-5* Substance Use Disorder Diagnostic Assessment were also temporally reliable and stable. This finding is important for at least two reasons. First, the test-retest reliability of this brief assessment has not been previously established. Prior research has demonstrated strong internal consistency reliability and construct validity for the assessment when evaluating alcohol use disorder in college students (Hagman 2017). The current study extends this research by showing that this measure can be easily adapted for other substance use disorders and that

these classifications show good stability over at least a month period. Second, the stability of these clinically classifications supports the validity of self-reported substance use behaviors in this crowdsourced sample. This outcome is particularly important given the inability to biologically verify substance use over the mTurk platform. Offsetting this limitation are previous studies indicating that crowdsourced samples do not engage in higher rates of problematic responding, such as socially desirable bias, and that these samples report feeling more comfortable sharing sensitive materials, such as substance use, over an online platform than in person (Kim and Hodgins 2017; Necka et al. 2016; Strickland and Stoops 2018b).

Cross-Commodity Demand

Cross-commodity demand tasks indicated that prescription opioids functioned as a weak complement for cannabis and that cannabis functioned as an independent commodity for opioids. These results were observed across varied analytic approaches and are generally consistent with a previous study conducted in patients recruited from an outpatient heroin clinic (Petry and Bickel 1998). Participants in that study completed an alternative version of the commodity purchase task in which hypothetical drug commodities are purchased following allocation of an experimental income. Cannabis was an independent commodity for heroin in that study and showed a similar pattern across heroin unit prices. Demand for both prescription opioids and cannabis was also reduced in a concurrent setting. This finding is in accordance with other studies in which the availability of a concurrent commodity has resulted in the reallocation of behavior even when those commodities did not function as strong complements or substitutes (for example see Johnson et al. 2017b). Taken together, these findings suggest that prescription opioids and cannabis do not function as strong complement or substitute goods and are thus generally insensitive to price changes in the alternative good.

Limitations

This study should be considered in the context of its limitations. First, the use of crowdsourcing methods does present potential concerns related to sampling bias and generalization. A substantive body of literature has documented the ways in which crowdsourced sampling may differ from nationally representative sources. These studies have found that individuals recruited from crowdsourcing platforms tend to be younger, more educated, and less likely to be employed, married, or a racial minority compared to nationally representative sources (Berinsky et al. 2012; Huff and Tingley 2015; Paolacci and Chandler 2014). Higher rates of alcohol and illicit substance use has also been observed in crowdsourced samples (Shapiro et al. 2013; Strickland and Stoops 2019) (but see Caulkins et al. 2015b, for information on the under-estimation of substance use in nationally representative sources). Other research, however, has provided good evidence for the validity of data collected via crowdsourced methods by demonstrating a correspondence between outcomes observed in laboratory, clinic, and online settings (see reviews in Chandler and Shapiro 2016; Strickland and Stoops 2019). The current study similarly replicated findings reported elsewhere both related to and independent of opioid use. Relationships between opioid demand and OUD were consistent with prior research conducted in college samples and were replicated in two independent samples using variations in task design (e.g., price structure). Discounting rates were shallower for money than for opioids or cannabis, which also replicates a canonical finding that consumable goods are more steeply discounted than money (e.g., Baker et al. 2003; Bickel et al. 2011b; Charlton and Fantino 2008; Johnson et al. 2007). Although limitations associated with the convenience nature of crowdsourced sampling should be considered, ultimately the combination of research from laboratory, clinical, and crowdsourced sources should benefit the rigor, reproducibility, and scope of research conducted in addiction science.

Second, we only evaluated cross-commodity demand at a single price for the fixed-price commodity. It is possible that selection of an alternative price could have produced a quantitatively and qualitatively different relationship between these commodities. The fixed-price values were selected to approximate realistic prices for those goods and to provide measurement at the intersection of the inelastic and elastic portion of the demand curve to avoid restrictions in range and maximize variability in responding. Third, prescription opioids were defined generally in the purchase task procedure as "the standard dose that you use when you use these pills". This approach has been successfully used previously and likely helps provide for a more general task accounting for the heterogeneity of prescription opioid medications. However, alternative procedures, such as defining specific opioid types and/or doses should be explored (for similar problems in defining quantities in e-cigarette purchase tasks see (for similar problems in defining quantities with an e-cigarette purchase task see Cassidy et al. 2017).

Conclusions

The primary finding of this study was that the commodity purchase task provided an incrementally valid and temporally reliable measure of opioid demand. These findings are consistent other research indexing valuation for alcohol and cigarettes using the purchase task procedure. Coupled with the present data, this body of work demonstrates that the purchase task procedure provides a clinically useful measure of drug valuation that is sensitive to individual difference variables relevant to drug-taking behavior. These studies also provide clear evidence for the utility of demand in providing relevant information about the behavioral mechanisms underlying the relative reinforcing effects of drugs that can be used to inform prevention and treatment efforts targeting substance use disorders.

Table 5.1. Demographic and Opioid Use Variables Experiment 3a

	Mean	SD	IQR
Age	35.5	9.2	28-41
Female	67.5%		
White	87.5%		
College	47.5%		
Income (USD)	46000	31000	20 to 70k
AUDIT-C	4.9	3.6	2 to 8
Opioid Use			
Past Month Opioid Use	4.9	7.7	1 to 6
DMS-IV Opioid Dependence	50.0%		
Risky Route	20.0%		
Opioid Demand			
Intensity (Q_0) [log]	1.04	0.45	0.72 to 1.33
Elasticity (α) [log]	-2.26	0.62	-2.79 to -1.86

Note. Risky Route = preferred use of intranasal, smoked, or injection administration.

Table 5.2. Bivariate Associations for Experiment 3a

	Opioid Q_0	Opioid α
Age	.27	.02
Female	.08	-.08
White	-.15	.19
College	.03	-.06
Income (USD)	.30	-.28
AUDIT-C	.10	.07
Opioid Use		
Past Month Opioid Use[a]	0.64	**0.51***
DMS-IV Opioid Dependence	**.43****	-.26
Risky Route	.17	-.15

Note. AUDIT-C = Alcohol Use Disorder Identification Test-Consumption. Risky Route = preferred use of intranasal, smoked, or injection administration.

[a]These variables were evaluated using negative binomial regression given the observation of zero-inflation. Values represent rate ratios.

* $p < .05$; ** $p < .01$ (**bold** = statistically significant)

Table 5.3. Demographics and Substance Use Variables Experiment 3b

	Mean	SD	IQR
Age	34.0	8.0	29 to 37
Female	63.9%	0.5	
White	89.2%	0.3	
College	51.8%	0.5	
Income	43000	28000	20k to 70k
Substance Use			
Month OPI	7.6	9.5	0 to 12
Risky Route	26.5%	0.4	
Month CAN	15.0	12.6	1 to 30
Grams/Week	7.4	9.5	2 to 10
***DSM-5* SUD**			
OUD	67.5%		
Number OUD	4.9	4.0	1 to 8
CUD	41.0%		
Number CUD	1.7	2.1	0 to 2
AUD	59.0%		
Number AUD	4.0	3.9	0 to 7
Pain			
Average Pain	4.6	2.4	3 to 6
Pain Effect	4.1	2.8	1.6 to 6.3
Opioid Relief	55.1	29.6	30 to 79
BE Variables			
Opioid Q_0 [log]	1.14	0.55	0.70 to 1.49
Opioid α [log]	-2.33	0.58	-2.72 to -1.95
Cannabis Q_0 [log]	1.57	0.57	1.16 to 2.05
Cannabis α [log]	-2.14	0.52	-2.47 to -1.79
Money *k* [log]	-2.08	0.69	-2.48 to -1.63
Opioid *k* [log]	-1.65	0.79	-2.17 to -1.08
Cannabis *k* [log]	-1.69	0.98	-2.43 to -1.23

Note. OPI = opioid; CAN = cannabis; SUD= substance use disorder; OUD = opioid use disorder; CUD = cannabis use disorder; AUD = alcohol use disorder; BE = behavioral economic; Risky Route = preferred use of intranasal, smoked, or injection administration.

Table 5.4. Bivariate Associations for Experiment 3b

	Q_0 OPI	α OPI	Q_0 CAN	α CAN	*k* USD	*k* OPI	*k* CAN
Age	.06	.13	-.13	.17	-.04	-.01	-.05
Female	.07	.06	-.12	**.26***	.15	.00	.05
White	-.02	-.16	.09	**-.30****	.16	.12	.12
College	-.09	.11	.02	-.04	**-.26***	**-.25***	-.20
Income	**-.40*****	.15	-.14	-.07	-.13	.01	-.08
Substance Use							
Month OPI[a]	1.44	0.62	1.18	0.47	1.13	1.23	1.11
Risky Route	.10	**-.38*****	.19	-.21	-.06	.05	.03
Month CAN[a]	1.31	0.61	**2.08****	0.69	0.95	1.07	1.05
Grams/Week	.19	-.20	**.47*****	**-.28***	.08	.14	.05
DSM-5 SUD							
OUD	**.52*****	**-.28***	.15	-.05	.11	**.30****	.20
Number OUD	**.40*****	**-.31****	.10	-.06	.13	**.29****	**.23***
CUD	-.08	.15	**.25***	-.01	-.02	-.06	-.01
Number CUD	-.01	.09	**.27***	-.08	.03	.01	-.01
AUD	.05	.00	.03	-.04	-.09	.00	.14
Number AUD	.11	-.04	-.01	-.04	-.12	-.01	.05
Pain							
Average Pain	**.29****	-.06	-.06	.07	**.27***	**.22***	.16
Pain Effect	**.32****	-.11	.03	.02	**.30****	**.25***	.14
Opioid Relief	-.06	.09	-.01	.16	.18	.08	.09

Note. OPI = opioid; CAN = cannabis; OUD = opioid use disorder; CUD = cannabis use disorder; AUD = alcohol use disorder; BE = behavioral economic; Risky Route = preferred use of intranasal, smoked, or injection administration.

[a]These variables were evaluated using negative binomial regression given the observation of zero-inflation. Values represent rate ratios.

* $p < .05$; ** $p < .01$; *** $p < .001$ (**bold** = statistically significant)

Table 5.5. Individual Price Change Cross-Price Elasticity

Price Change	Cannabis	Opioids
$0.25	NA	NA
$0.50	-0.145	0.032
$1	-0.030	-0.289
$1.50	-0.058	-0.126
$2	-0.124	-0.109
$2.50	0.024	-0.081
$3	-0.010	-0.302
$4	0.085	-0.029
$5	0.107	-0.041
$6	0.092	-0.059
$7	0.080	-0.192
$8	0.110	-0.036
$9	0.136	-0.013
$10	0.053	-0.137
$15	0.048	-0.016
$20	0.027	-0.023
Overall	0.005	-0.110

Note. Values represent cross-price elasticity generated across each individual price along the cross-price demand curve for the indicated price-fixed commodity.

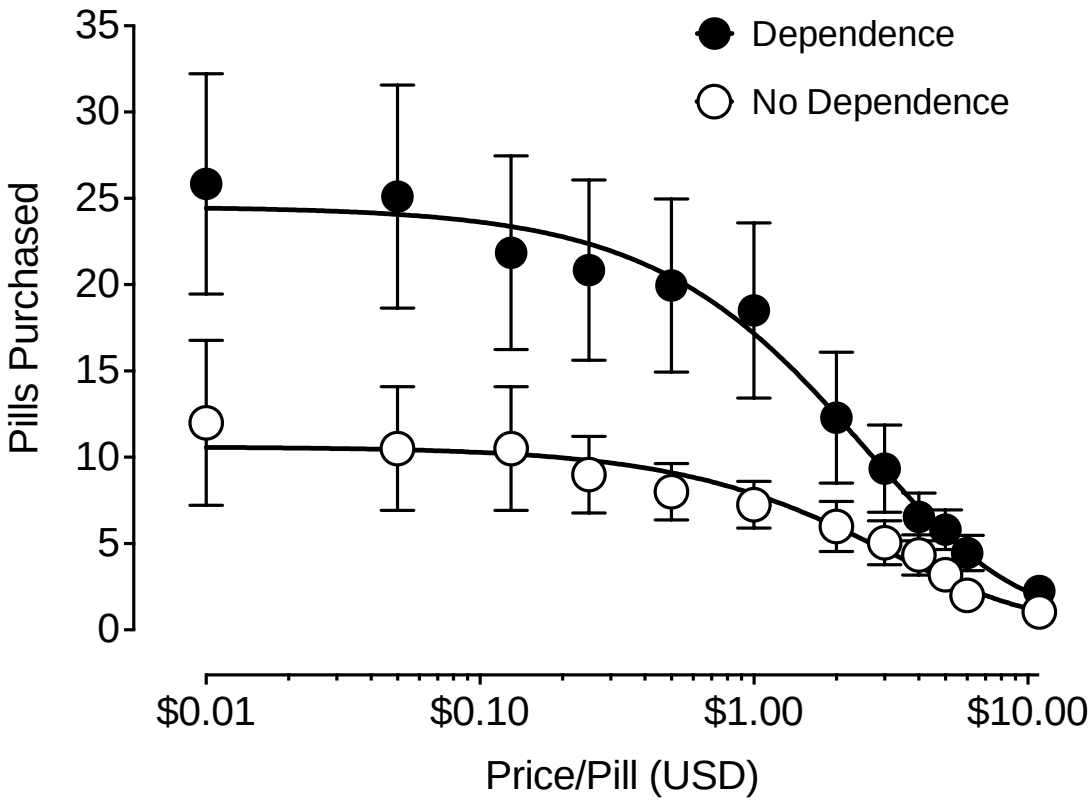

Figure 5.1. Behavioral economic demand for prescription opioids. Participants completed a commodity purchase tasks in which prescription opioids were available. Price varied in United States dollars (USD). Plotted are mean (SEM) group data fit using the exponentiated model for individuals with (closed circles) and without (open circles) *DSM-IV* opioid dependence.

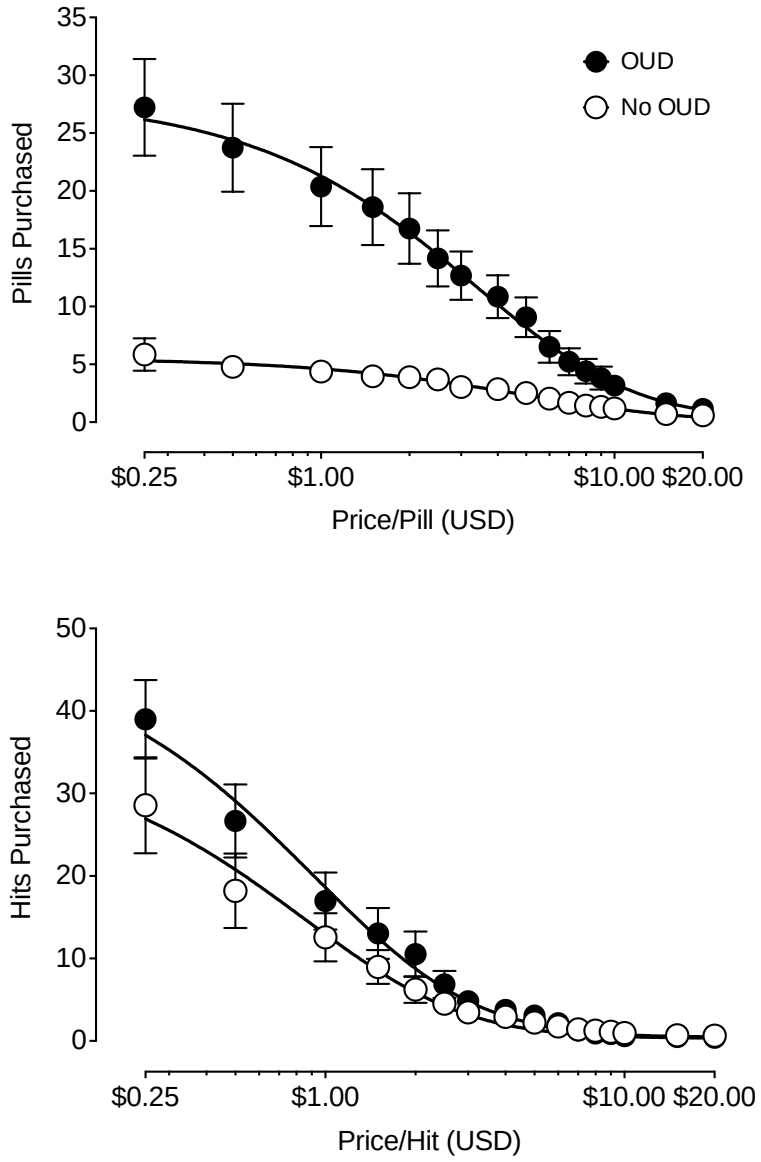

Figure 5.2. Behavioral economic demand for prescription opioids and cannabis. Participants completed commodity purchase tasks in which prescription opioids (top) or cannabis (bottom) were available. Price varied in United States dollars (USD). Plotted are mean (SEM) group data fit using the exponentiated model for individuals with (closed circles) and without (open circles) *DSM-5* opioid use disorder (OUD).

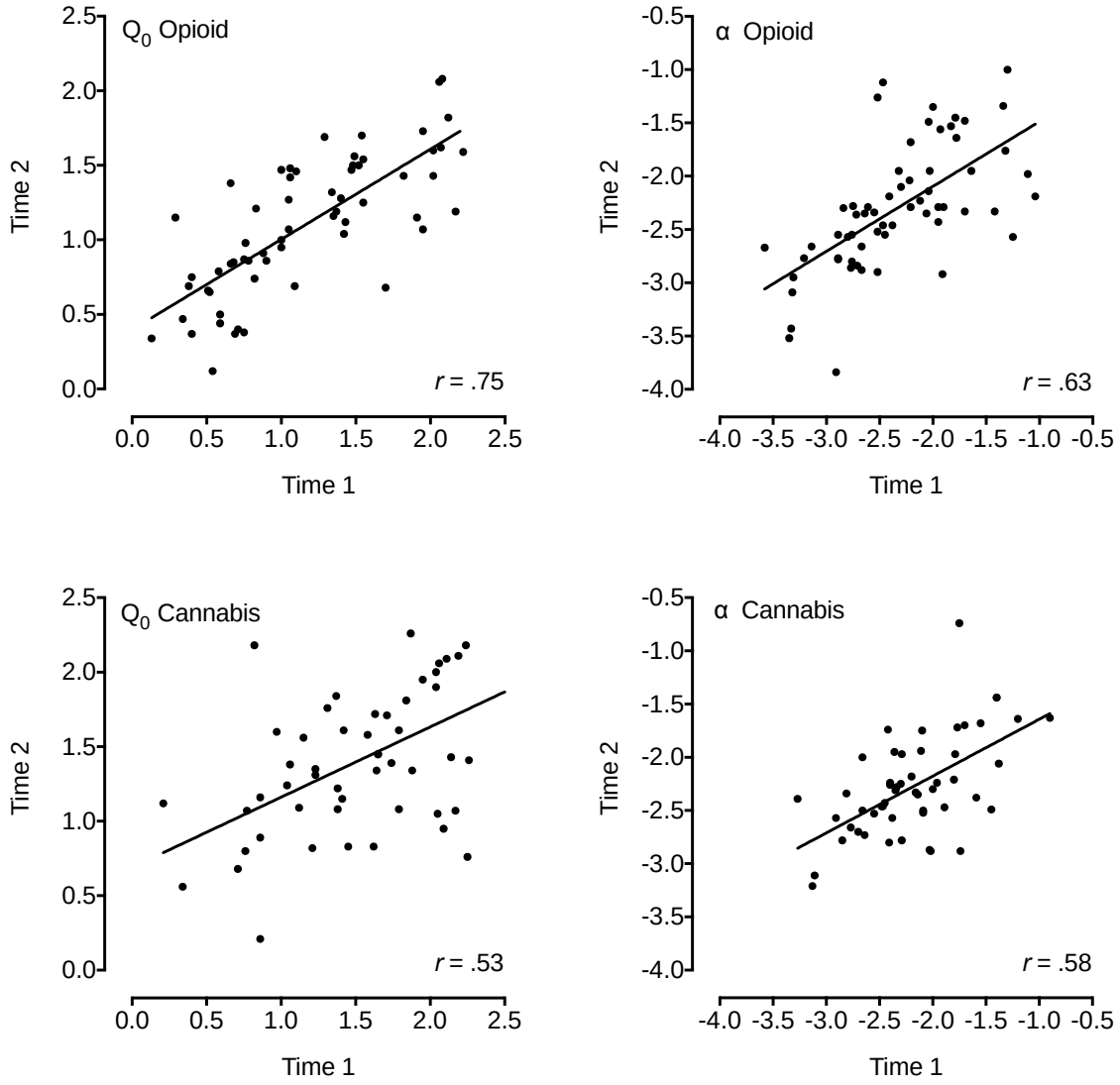

Figure 5.3. Test-retest reliability of behavioral economic demand. Plotted are values for opioid (top) and cannabis (bottom) demand at time 1 and time 2 separated by a one month period. Test-retest reliabilities are located in the bottom-left corner of each panel and all were statistically significant.

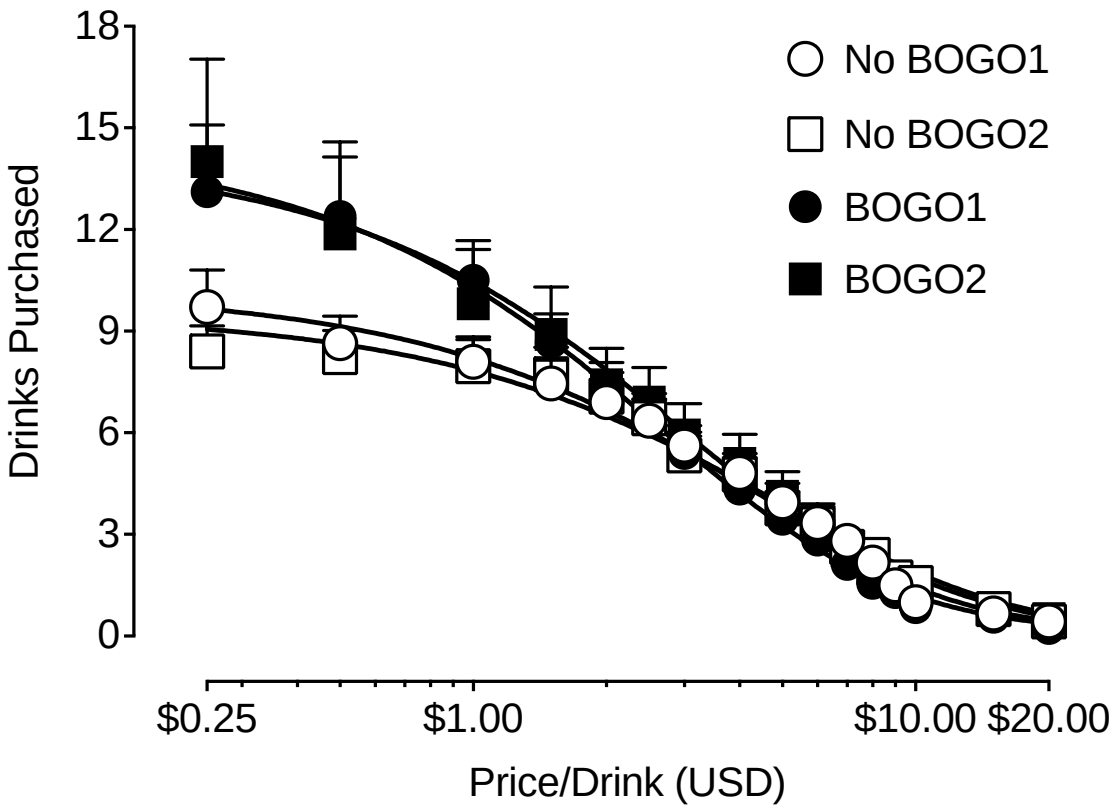

Figure 5.4. Behavioral economic demand for alcohol under varying drink specials. Participants completed commodity purchase tasks in which alcohol was available under a buy-one get-one (BOGO; closed shapes) special or no-special (open shapes). Values were collected at two time points separated by one month (Time 1 = circles; Time 2 = squares). Price varied in United States dollars (USD). Plotted are mean (SEM) group data fit using the exponentiated model.

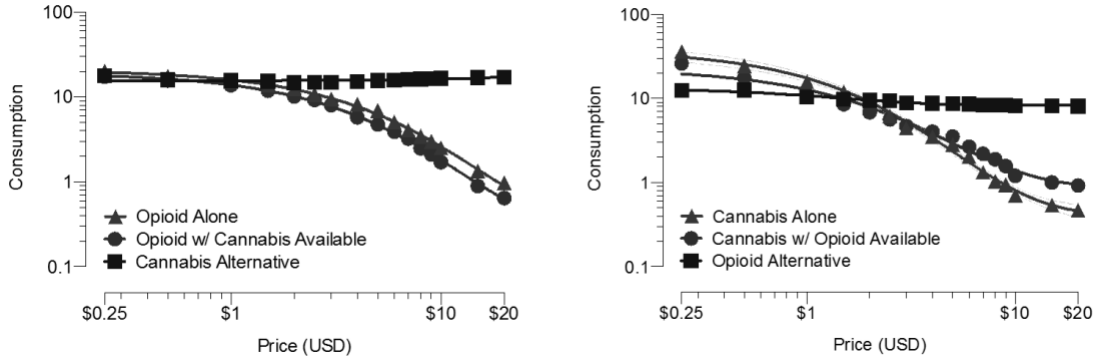

Figure 5.5. Cross-commodity demand for opioids and cannabis. Plotted are cross-commodity demand curves fit using the Hursh and Roma (2013) cross-commodity formula. Mean values are presented for the price-varying commodity alone (triangles), the price-varying commodity with the concurrent commodity available (circles), and for the price-fixed commodity (squares).

Chapter 6

FEASIBILITY, ACCEPTABILITY, AND VALIDITY OF CROWDSOURCING FOR COLLECTING LONGITUDINAL ALCOHOL USE DATA

(Experiment 4a; Strickland and Stoops 2018b)

Introduction

Longitudinal research methods provide numerous benefits for the study of human health and behavior. The inclusion of repeated data collection time points from the same individual allows for the elegant evaluation of both between- and within-person change processes (Bolger and Laurenceau 2013; Singer and Willett 2003). In the case of the behavioral and addiction sciences, the prospective study of drug-taking behavior and related time-varying covariates can provide an improved understanding of disease progression over time and the environmental influences controlling these trajectories. These benefits of longitudinal methods are nevertheless tempered by financial, time, and geographic constraints. Costly incentive schedules and extensive staffing are often needed to ensure adequate follow up rates, exponentially increasing the budgets of longitudinal projects compared to cross-sectional ones. These problems may be compounded when tracking transient individuals as is common with substance-using populations. Recruitment and logistic burden can also result in samples limited to select geographic regions, which can diminish the generalizability and external validity of subsequent findings.

An emerging sampling method positioned to address these concerns is crowdsourcing. Generally speaking, crowdsourcing refers to the use of the Internet to outsource work through an open call to solve a specific problem. One of the most prominent crowdsourcing platforms is Amazon Mechanical Turk (mTurk; also commonly abbreviated as AMT, MTurk, or MTURK). mTurk belongs to a class of online workforce markets in which individuals may sign up to complete varied tasks for financial

compensation. Over the past decade, behavioral and social scientists have recognized the practical benefits afforded by mTurk for generating convenience samples (see reviews by Chandler and Shapiro 2016; Keith et al. 2017; Woods et al. 2015). The large and readily accessible population of participants available on mTurk is analogous to undergraduate psychology participant pools that are often used for convenience sampling in these disciplines. Unlike participant pools, however, mTurk provides a sampling pool that is geographically diverse and not limited to young-adult college students.

Existing research on mTurk has supported the reliability and validity of cross-sectional data collection in behavioral science, broadly, and addiction science, specifically. Several studies have demonstrated scale reliabilities and factor structures for common research scales (e.g., the Big Five Inventory) that are consistent with those observed using traditional sampling methods (e.g., Behrend et al. 2011; Buhrmester et al. 2011; Shapiro et al. 2013). Replications of common behavioral phenomenon (e.g., the Stroop effect) via the mTurk platform have also been demonstrated (e.g., Crump et al. 2013; Enochson and Culbertson 2015). Although more recent, researchers in addiction science have also begun to utilize mTurk. These studies have spanned a diverse range of theoretical perspectives and methods including, but not limited to, behavioral economics (e.g., Bickel et al. 2012; Johnson et al. 2015; Kaplan et al. 2017; Morris et al. 2017; Peters et al. 2017), tobacco control policy (e.g., Lazard et al. 2017; Pearson et al. 2016; Shi et al. 2017), behavioral addictions (e.g., Bock et al. 2016; Gearhardt et al. 2016), public opinion related to addiction-related policy (e.g., Huhn et al. 2017; Rudski 2016; Wen et al. 2016), and measure development (e.g., Dunn et al. 2016a; Dunn et al. 2016b; Lac and Berger 2013). Existing evidence supports the reliability and validity of common substance use scales when used on mTurk (e.g., the Alcohol Use Disorder Identification Test [AUDIT]) (Kim and Hodgins 2017). This

emerging literature has also replicated common effects described in the addiction laboratory (e.g., higher discounting rates among tobacco cigarette smokers relative to non-smokers) or demonstrated correspondence in behaviors between in-person and online samples, further supporting the validity of the approach (e.g., Jarmolowicz et al. 2012; Johnson et al. 2015; Morris et al. 2017; Strickland et al. 2016a; Strickland et al. 2017b).

One practical benefit of mTurk is the unique identifiers assigned to participants that allow for easy and efficient repeated research contact in test-retest and other longitudinal designs. For example, Daly and Nataraajan (2015) observed response rates of 75% at two months, 56% at four months, and 47% at 13 months across three independent surveys. In substance-using populations specifically, Kim and Hodgins (2017) observed high one-week follow up rates (> 87%) when evaluating alcohol-using, cannabis-using, and gambling participants. Recent studies have extended this longitudinal approach to intensive methods (Boynton and Richman 2014; Hartsell and Neupert 2017; Lanaj et al. 2014). Intensive methods involve the use of frequent measurements to characterize rapid fluctuations in behavior and include designs such as daily diary, experience sampling, and ecological momentary assessment (Bolger and Laurenceau 2013). These methods are particularly appealing because they provide a precise temporal design for measuring day-to-day or moment-to-moment changes, which may help to detect subtle changes in psychological or behavioral processes.

To our knowledge, the only study to use intensive longitudinal methods on mTurk with a substance-using population evaluated alcohol consumption over a 14-day period using a daily diary design (Boynton and Richman 2014). Participants completed an average of 8.5 daily measurements (60.7%) providing preliminary support for feasibility of the approach. Effects consistent with the extant literature were also observed, such as more frequent and severe drinking by individuals with positive CAGE scores. However,

the short testing window (i.e., 14 days) and lack of acceptability measures limits the conclusions that may be made about the generalizability of this approach to longer-term protocols and for alternative applications.

The purpose of the present study was to extend these earlier findings by comprehensively evaluating the feasibility, acceptability, and validity of using mTurk for collecting intensive longitudinal data in addiction and behavioral science. Participants recruited through mTurk completed an 18-week intensive longitudinal design in which alcohol use was recorded. Soda use was also recorded as a non-drug comparator to evaluate the specificity of observed findings to the reporting of drug-related rather than general consummatory behavior. Acceptability measures were collected to characterize participants experience with the study protocol and likelihood of future research participation. The guiding hypothesis was that long-term, intensive data collection (i.e., methods with dense measurement in design) would be feasible, acceptable, and valid.

Method

Participants and Screening

Participants were recruited from mTurk. In order to view the study participants were required to have an approval rating of 95% or higher on previous tasks, have completed at least 100 mTurk tasks, and currently reside in the United States, consistent with other research (e.g., Cunningham et al. 2017; Reed et al. 2016; Strickland and Stoops 2015). A short (~1 minute) screening survey including demographic and alcohol use questions was used to determine study eligibility. Inclusion criteria were: 1) age 21 or older, 2) self-reported alcohol use in the week prior to screening, 3) AUDIT score of 1 or higher, and 4) willingness to complete an 18-week study. All surveys were hosted on Qualtrics (Provo, UT, USA). Participants received $0.05 for completing the screening survey.

General Procedures

Qualifying participants first completed a baseline survey that included demographics, substance use history, and other cognitive-behavioral measures. The purpose of this analysis is to evaluate the feasibility, acceptability, and validity of the longitudinal methodology and therefore focuses on these design-related questions. Participants were paid $1.00 for completing the baseline survey.

The longitudinal phase consisted of 18 continuous weeks of surveys. These weekly surveys asked participants to record past week alcohol and soda use behaviors. Participants received weekly emails through the mTurk platform indicating that the weekly survey was available from 900AM Monday to 900AM Wednesday (EST). All participants completed the study in a contemporaneous set of 18 weeks (July 3 2017 to November 5 2017). Payment for each survey was $0.40. Active participation was incentivized by entry into a raffle for one of five $50 bonuses if participants completed 14 or more weekly surveys. No limits were placed on the number of data collection periods that could be missed (i.e., participants were not excluded from further participation if missing one or more study weeks). Participants were also asked to complete a post-study survey that included acceptability measures. This follow-up survey was completed one week after the longitudinal phase and was compensated with $0.75.

Compensation rates were initially designed to approximate United States minimum wage ($7.25/hour at the time of the study). Actual compensation rates were determined by calculating the median time per survey and computing compensation rates/hour for each study phase (Baseline = 21.38 minutes [$2.81/hour]; Weekly Follow Up = 1.92 minutes [$12.53/hour]; Post-Study Follow Up = 21.07 minutes [$2.14/hour]). Calculating overall compensation rate based on money earned versus median time spent across the entire study (assuming an individual completed all assessments) the hourly compensation rate was $6.98/hour.

Study Measures

Feasibility

The primary feasibility measure was weekly response rates. Response rates were computed as a percent expressing the number of weekly surveys completed as a function of eligible participants (i.e., those that completed the baseline survey and passed all data quality checks described below). The number of weeks completed by each participant and percent participants completing all surveys was also computed.

Acceptability

The primary acceptability measure was a modified version of the Treatment Acceptability Questionnaire (Hunsley 1992; Raiff et al. 2013). Modifications reflected the non-treatment nature of the study. Participants rated their response to six statements about the weekly surveys (i.e., ease of completion, helpful instructions, enjoyability, convenient timing, fair compensation, and overall satisfaction) on a 100-point visual analog scale (0 = Low; 100 = High). Additional questions were used to evaluate overall satisfaction, future participation, participation motives, and experience with mTurk. Participants were explicitly instructed to respond honestly and that their choices would not affect any subsequent research payments. Acceptability measures were included in a post-study follow-up survey to further decrease potential demand characteristics.

Validity

The primary measure used to evaluate validity was self-reported behavior during the longitudinal phase. Participants were asked to report the number of standard drinks consumed by alcohol type (e.g., 12 oz. beers, 1.5 oz. shot alone, 1.5 oz. shots in mixed drink, etc.) on each day during the past week (previous Monday to Sunday). Alcohol was summed across drink types for total number of drinks/day. The number of 12 oz. servings of soda consumed was also recorded. Participants were explicitly instructed

that sodas did not include those mixed with alcohol to avoid confusion with mixed drinks containing alcohol or alcopop beverages.

Data Quality Checks

A battery of attention and validity checks was used to identify inattentive or non-systematic participant data. These checks were in the baseline survey and included: 1) comparison of sex and age responses at the start and end of the survey, 2) an item that instructed participants to select a specific response (i.e., "Select 'A Little Bit'"), 3) recall of a single digit number presented halfway through the survey that participants were instructed to remember and enter at the end of the survey, and 4) an item asking participants if they had been attentive and thought their data should be included. Data were also examined for inconsistent responding (e.g., reporting smoking on the screening survey, but not in the baseline survey). We have successfully used these or similar data quality checks in previous research on mTurk (e.g., Strickland et al. 2017b; Strickland and Stoops 2017), as have other investigators in the behavioral and addiction sciences (e.g., Chavarria et al. 2015; Donaldson et al. 2016; Johnson et al. 2015; Peters et al. 2017). Data quality checks were not included in the weekly surveys, which could be considered a limitation of the design (although also see discussion of overuse of attention checks in (although also see discussion of overuse of attention checks in Chandler and Shapiro 2016).

Demographic and Alcohol Use History

Demographic variables (e.g., age, race, education) and alcohol use history (e.g., drinks/week) were also collected as a part of the baseline survey. Alcohol use disorder was evaluated using a written version of the Mini-International Neuropsychiatric Interview (MINI) and self-reported statements indicative of DSM-IV criteria for alcohol use disorder (Sheehan et al. 1998).

Data Analysis

A total of 307 participants completed the baseline survey. Baseline data were first evaluated for attentive and systematic responding. Twelve participants failed one or more data quality checks on the baseline survey. Another 13 participants did not provide any follow up assessments and 4 did not report drinking alcohol during the follow up period and were removed for acceptability and validity testing. This resulted in a final sample size of 278 (91% of the baseline sample).

Descriptive statistics were used to express completion rates. Completion rates were computed including individuals that did not provide any follow up assessments or did not report alcohol use during the follow up so as not to artificially inflate values (N = 295). Predictors of completion rates were evaluated using Spearman correlations to account for variable skew. Descriptive statistics were also used to evaluate acceptability data. Visual inspection of individual data was also conducted to summarize general patterns of alcohol consumption. Sample plots were selected for depiction based on this visual inspection. Empirical methods (e.g., latent class analysis, cluster analysis) were not used to select plots, although this strategy could be used in future work by those interested in classifying patterns of behavior reported by mTurk participants.

Generalized linear mixed models were used to evaluate correlates of self-reported alcohol and soda use. These tests were designed to evaluate the construct validity of online data collection by 1) demonstrating relationships between variables that should be related (i.e., convergent validity) and 2) demonstrating that variables that should not be related were in fact not related (i.e., discriminant validity). Placing these tests within the context of the extant alcohol literature also provides support for the external validity of this data collection method by demonstrating the extent to which the results of research conducted on mTurk generalize to the “real world” setting.

Three outcome settings were tested: 1) alcohol use (dichotomous; yes/no), 2) drinks/day (count using a negative binomial distribution given the observation of overdispersion), and 3) heavy drinking day (dichotomous; yes/no). Heavy drinking was defined using National Institute on Alcohol Abuse and Alcoholism (NIAAA) guidelines of 5+ drinks/day for men and 4+ drinks/day for women (National Institute on Alcohol Abuse Alcoholism 2007). Drinks/day and heavy drinking day models were estimated for days with alcohol use reported so as to not conflate drinking frequency (Model 1) with quantity (Model 2) and severity (Model 3). Between-subject predictors (Level 2 predictors) included demographics (age, sex, race, and education) and AUDIT scores. Within-person predictors (Level 1 predictors) included day of week (weekend [Friday, Saturday, Sunday] versus weekday). Continuous variables were grand-mean centered prior to analysis. Unadjusted (i.e., single predictor) and adjusted (i.e., all predictors) models were evaluated for each setting. No data were missing at Level 2 and all Level 1 data were missing because of study attrition. Analyses described in the Results section suggested that attrition was not systematically related to alcohol use behavior. Accordingly, data were treated as missing at random (Singer and Willett 2003).

Additional models using the same predictor variables tested 1) soda use (dichotomous; yes/no) and 2) sodas/day on soda-drinking days (count using a negative binomial distribution). Models evaluating “heavy” soda drinking were initially examined. Heavy soda use was defined using the same guidelines as heavy alcohol use set by NIAAA. Although these guidelines were developed for alcohol use and may not directly reflect at-risk soda consumption (however see Strickland and Stoops 2017, for evidence that this measure corresponds to relevant measures of soda valuation), corresponding variables were computed so as to decrease the possibility that observed patterns were due to systematic differences in measurement for each commodity. Nevertheless, problems with model convergence were encountered due to the low rates of heavy soda

use observed across the study. Models that did converge (e.g., the fully adjusted model) did not indicate significant effects of the predictor variables tested.

All inferential testing used two-tailed tests and an alpha rate of .05. Maximum likelihood estimation using a Laplace approximation was used for generalized mixed effect models. All models were conducted *R* statistical language and the *lme4* and *glmmTMB* packages.

Results

Sample Characteristics

Table 6.1 contains demographics and substance use behaviors collected at baseline. An approximately equal distribution of men and women were sampled. A majority of participants were white and reported a college education or greater. Good variability was observed in alcohol use behaviors, with 40.7% of participants meeting *DSM-IV* criteria for an alcohol use disorder.

Feasibility

Weekly response rates (top panel) and the distribution of individual response rates (bottom panel) are plotted in Figure 6.1. High response rates were observed across the study, with the highest rate observed in Week 1 (86.8%) and lowest in Week 16 (64.1%). Nearly three-quarters (73.7%) of participants completed more than half of the assessments, 65.1% completed 14 or more assessments, and 43.1% completed all 18 assessments. Response rates was not associated with sex, race, education, soda use, or cigarette use, r values < .11, p values > .05. Older participants had higher response rates, $r = .17$, $p = .004$. Response rates were also not related to baseline alcoholic drinks per week, days drinking per week, or presence of alcohol use disorder, r values < .08, p values > .16. Higher AUDIT scores were associated with lower response rates; however, this effect was of a small effect size, $r = -.16$, $p = .008$. Qualitatively similar results were

observed when comparing fully compliant individuals (i.e., completed all surveys) to those without full compliance.

Acceptability

Generally high ratings were observed for acceptability questions (median values: Ease = 100, Helpful = 96, Enjoyable = 84, Convenient = 95, Fair Compensation = 95, Overall = 92; Table 6.2). Individual participant data indicated that a majority of participants clustered in the upper quartile of the acceptability rating scale with skew driven by the minority of participants providing low acceptability ratings (Figure 6.2).

Table 6.2 contains percentage endorsement of other acceptability measures. A majority of participants reported that they would definitely or probably participate again (98.1%) and that they were satisfied with the study procedures (93.9%). The most common motivation for participation was to make money (82.6%). A majority of participants also indicated that they found it easier to answer honestly sensitive questions on mTurk compared to in person (72.8%) and that they would be interested in participating in future research on mTurk designed to reduce problem behaviors such as alcohol use or overeating (93.0%).

Validity

A total of 27104 study days were recorded. Participants reported alcohol use on 40.1% of these days and heavy alcohol use on 13.9%. Figure 6.3 plots percent participants reporting alcohol use (solid lines and closed circles) and heavy alcohol use (dotted lines and open circles) over the study. Visual inspection revealed a clear and consistent cyclic pattern of increased drinking and heavy alcohol use that corresponded to day of the week. Visual inspection also revealed trends that corresponded with environmentally relevant events. For example, a spike in drinking and heavy alcohol drinking was observed on the first and second day of the study (Days 1 and 2), which was inconsistent with typical drinking patterns observed on Mondays and Tuesdays.

However, these two days corresponded with a major United States holiday (that Tuesday was July 4th), which is associated with increased alcohol consumption.

Table 6.3 contains effect size estimates (odds and rate ratios) for unadjusted and adjusted generalized linear mixed models predicting alcohol use behaviors. Higher AUDIT scores and weekends were associated with increased odds of drinking alcohol and heavy drinking and a greater number of drinks in unadjusted models. Participants with a college education also showed fewer drinks consumed and decreased odds of heavy drinking in unadjusted models. Unadjusted comparisons also indicated that men reported more drinks consumed and younger individuals reported fewer drinks and reduced odds of heavy drinking. Adjusted comparisons were generally consistent with unadjusted models. In particular, AUDIT scores and weekends were again closely associated with all three alcohol use outcomes such that higher AUDIT scores and weekend days were associated with an increased odds of drinking alcohol and heavy drinking and a greater number of drinks consumed on drinking days.

Visual inspection of individual participant plots revealed varied patterns of behavior that were consistent with the group-level analyses. Figure 6.4 contains three sample response patterns from participants reporting low-risk alcohol use and Figure 6.5 contains sample response patterns from participants reporting high-risk alcohol use (low/high risk based on AUDIT cutoff scores). Participants in the low-risk group displayed patterns including, but not limited to, intermittent low level consumption (top panel), stable low-to-moderate consumption (middle panel), and intermittent moderate alcohol use (bottom panel). In contrast, participants in the high-risk group displayed patterns including moderate weekday consumption with heavy weekend consumption (top panel), punctuated, but consistent, heavy binge consumption on weekends (middle panel), and heavy weekly consumption with heavier weekend drinking (bottom panel).

Soda consumption was reported by 154 participants during the longitudinal period. Among these participants, soda use was reported on 37.2% of days and heavy soda use on 6.1% of days. In contrast to the alcohol use, no clear daily fluctuations in soda use or heavy soda use were observed.

Table 6.4 contains effect size estimates (odds and rate ratios) for unadjusted and adjusted generalized linear mixed models predicting soda use behaviors. White participants had higher odds of soda consumption in the unadjusted model, but no other variables were associated with the odds of drinking soda in unadjusted or adjusted models. Unadjusted and adjusted models predicting sodas/day did not reveal any significant predictors.

Discussion

The overall purpose of this study was to evaluate the feasibility, acceptability, and validity of using mTurk to collect intensive longitudinal data (i.e., methods with dense measurement in design) in addiction and behavioral science. To this end, participants recruited from mTurk completed an 18-week study in which daily alcohol and soda use was recorded at weekly intervals. Response rates were generally high over the 18-week period, participants reported that the study procedures were acceptable and that they would participate in future research, and between and within-person variations in alcohol consumption conformed to expected relationships. Taken together, these results comprehensively demonstrate the feasibility, acceptability, and validity of utilizing crowdsourcing for collecting longitudinal data with substance-using populations and support the future use of this sampling method in other behavioral research.

Feasibility was primarily confirmed by high response rates during the 18-week period. Average response rates of 73% were observed across the study, with nearly three-quarters of participants providing data for more than half of the assessments (average number of assessments completed = 13.2). Although it is difficult to compare

response rates across the extant literature given its size and the heterogeneity in procedures and populations, comparisons to some other studies using intensive longitudinal alcohol report may help highlight the relative feasibility and success of the methods used here. These comparisons were selected to reflect studies that share some features with the current study or have other desirable aspects for future work, including one study using data from the nationally representative National Study of Daily Experiences (Almeida et al. 2002), two using a similar weekly recording design (Braitman et al. 2017; Tremblay et al. 2010), and two recruiting a clinically relevant sample of individuals enrolled in a pharmacotherapy trial (Bold et al. 2016; Kranzler et al. 2009). Response rates across these studies ranged from 49.2% to 87.5%. The highest response rates were from the National Study on Daily Experiences, in which the investigators made daily phone calls for data collection, representing a likely substantive burden for the research team (Almeida et al. 2002). Two studies that used a weekly recording design reported response rates of 49.2% (Braitman et al. 2017) and 82.3% (Tremblay et al. 2010), potentially reflecting the greater incentives used in the latter case (i.e., course credit versus money [$5/weekly survey + a $50 raffle], respectively). These brief numerical comparisons between response rates are consistent with the broader literature in which this study's rate of 73% is within those generally observed in longitudinal research, albeit on the lower range. We consider this an acceptable response rate given the use of a relatively lean incentive schedule and low-intensity contact made with participants each week in the current study.

In fact, one of the distinctive strengths of mTurk for longitudinal data collection is improved feasibility regarding financial and time constraints. The current project was completed with a participant payment and mTurk fee budget of approximately $3000 (~$10 per participant) and data collection that was coordinated and executed by a single person. Although features of this study were selected to further optimize efficiency (e.g.,

contemporaneous data collection from all participants), this minimal researcher burden stands in contrast to the typical expenses incurred while conducting longitudinal research and the resources that must be dedicated to recruiting and retaining participants. A potential criticism related to this strength is the seemingly low wages provided to participants. Financial gain was the most common motivation for participation, consistent with previous research in alcohol-using, cannabis-using, and gambling participants (Kim and Hodgins 2017). This finding suggests that participants were attending to the contingencies related to compensation and that any deviations from expected or desired compensation would likely be reflected in feedback. Participants generally indicated that the wages provided were fair (median ratings = 95/100 for fairness of wages). Related to this concern is the observation that some researchers and regulation boards have argued that compensation rates on mTurk should meet a minimum wage standard (Gleibs 2017; Goodman and Paolacci 2017). Compensation for this study was initially designed to approximate United States minimum wage. The overestimation and underestimation of expected completion times highlights the difficulties that may be experienced in setting compensation in the absence of extensive pilot testing. Nevertheless, overall compensation did not markedly differ from minimum wage (assuming a participant completed all assessments), which indicates that cost would not dramatically increase with revisions to improve rates in order to better approximate a minimum wage. The appropriate compensation for online work remains a necessary conversation in the research community (Chandler and Shapiro 2016); however, the positive response recorded suggests that the compensation provided was at least experienced by participants as acceptable.

Participants indicated that they had an overall positive response, that the study procedures were easy to complete and convenient, and that they were likely to participate in future research like this study. A majority of participants also indicated that

they would be interested in future studies designed to reduce problem behaviors such as alcohol use or overeating. This finding is consistent with recent studies that have successfully evaluated brief interventions for alcohol use on mTurk, either within a single session or when using a single follow-up (Cunningham et al. 2017; Kuerbis et al. 2016; Kuerbis et al. 2017), and suggest that future research could adapt these intensive, long-term longitudinal methods for such intervention development purposes.

Acceptability measures were collected at the end of the study to decrease demand characteristics by helping to ensure participants that payments would not be affected by their responses. However, this also meant that acceptability data were only collected from three-quarters of the total sample (i.e., they did not include participants that were no longer participating in the study). It is possible that inclusion of dropout participants would have revealed less robust endorsement for study acceptability. Response rates were not closely or systematically related to alcohol use behaviors suggesting that those participants not included did not differ in this regard. Additionally, even if missing participants were conservatively coded as the most negative response (e.g., would not participate again), good support for acceptability was still evident (e.g., 75.2% of all participants reporting interest in participating again). The overwhelming positive response received from available participants regarding the study design and expressed future interest in research participation therefore supports the acceptability of these designs for future work.

Between- and within-person predictors of longitudinal patterns in alcohol use were consistent with expected effects and supported the construct and external validity of data collection. Clear effects of alcohol use severity (i.e., AUDIT scores) and environmental features (i.e., weekends) were observed for drinking frequency, quantity, and severity. These effects were specific to alcohol use and not observed for soda consumption. Numerous studies have revealed similar associations of AUDIT scores and weekends

with alcohol use across a variety of populations (e.g., college students, non-student emerging adults; older adults) (e.g., Kushnir and Cunningham 2014; Lau-Barraco et al. 2016; Sacco et al. 2016; Tremblay et al. 2010). The clear correspondence replicated here supports the construct validity of data collection by revealing expected relationships for between-subject (i.e., AUDIT scores) and within-subject (i.e., weekends) predictors when using this novel data collection method. More broadly, these findings also support the external validity of data collection by suggesting that research conducted on mTurk generalizes to the "real world" setting, a finding consistent with other work conducted on the platform (e.g., Athamneh et al. 2017; Jarmolowicz et al. 2012; Johnson et al. 2015; Morris et al. 2017; Strickland et al. 2016a). We selected relatively simple, main effect models to provide a straightforward and clear demonstration of the validity of data collected and to establish this methodological platform for future work. The feasibility, acceptability, and validity demonstrated in this study indicate that more complex models could easily be evaluated in future studies to test novel research hypotheses (e.g., moderation or meditational analyses). Improvements in web-based technology that allow for the conduct of reaction time and other cognitive-behavioral experiments will further advance the capabilities of future longitudinal projects conducted in this online setting (e.g., De Leeuw 2015; Seithe et al. 2016; Stoet 2017).

One common criticism of mTurk is that samples generated may systematically differ from populations of interest, thereby reducing generalizability and external validity. Despite the many benefits of mTurk, the method is still a form of convenience sampling and will result in samples that deviate from nationally representative sources. In general, mTurk samples tend to be younger, more educated, less religious, and more liberal as well as less likely to be married, a racial minority, or fully employed than those in a national representative study (e.g., Berinsky et al. 2012; Huff and Tingley 2015; Paolacci and Chandler 2014). Differences in demographics from national representative data sets

are inherent to convenience samples and are still likely in other forms of in-laboratory research that use convenience sampling (e.g., community posting, college student samples). When comparing the relative deviations across these different convenience methods there is some evidence that mTurk samples may be more representative than college samples or those drawn from college towns (Berinsky et al. 2012). Self-admission of engagement in problematic responding (e.g., responding in socially acceptable rather than truthful ways) also does not systematically differ between mTurk, community, and college sources (Necka et al. 2016). Comparisons in racial composition between this sample and those used in the previously noted comparator studies (see discussion of attrition rates above) also highlight some ways in which mTurk may not markedly differ from other forms of convenience sampling. The percentage of white participants across these studies varied from 63.4% (Braitman et al. 2017) to 96.9% (Kranzler et al. 2009). Surprisingly, 90.3% of participants in the nationally representative National Study of Daily Experiences were white (Almeida et al. 2002), a percentage that exceeds that of this study (82.7%). These comparisons are not meant to diminish concerns about the demographic representativeness of mTurk or to argue that the research community should not be attentive to the generalizability of research findings. Instead, these observations are meant to demonstrate that questions of generalizability and demographic representativeness are not unique to mTurk and reflect concerns when dealing with any form of convenience sampling whether online or in a laboratory.

Other limitations of the current application of mTurk for longitudinal research provide some future directions for evaluating and utilizing this technique. First, we only collected alcohol and soda consumption and it is unclear if this methodology would translate to other substances and substance-using populations. Alcohol was selected given extensive existing longitudinal research evaluating alcohol and the clinical acceptance of alcohol use self-report as a primary outcome. A growing body of research has evaluated

illicit substance users on mTurk (e.g., cannabis, cocaine, opioids) and generally revealed findings consistent with biologically verified in-person research (e.g., Dunn et al. 2016a; Dunn et al. 2016b; Peters et al. 2017; Strickland et al. 2016a; Strickland et al. 2017b). These prior studies suggest that recruitment of other substance-using populations should not prove problematic and that the methods proposed here would effectively translate.

Second, recall bias could have influenced past week recording of alcohol and soda use. We selected weekly rather than a daily diary design to help reduce participant burden and cost while maintaining the density of data collection. This selection likely helped to increase the possible temporal window of sampling and is consistent with methods using in other diary sampling studies (Braitman et al. 2017). A weekly and prospectively collected assessment window also likely resulted in less recall bias than retrospective recall of "typical" behavior as is used in cross-sectional work. It is still possible, however, that recall bias may have differentially altered responding and introduced systematic measurement error. To address this concern, we coded on what day each participant completed weekly surveys (i.e., Monday, Tuesday, Wednesday) and evaluated if day of completion was related to any of the primary outcomes. A majority of participants completed assessments on Monday (79% of recorded assessments) compared to Tuesday (18%) or Wednesday (3%). Comparisons using generalized linear mixed models (all models unadjusted) indicated that day of completion (Monday versus not Monday) was not significantly related to number of alcoholic drinks, heavy drinking, or soda use. Significant effects were observed for reporting alcohol use and number of sodas, with a modest increase in reporting drinking for any given day in the past week (OR = 1.20, $p < .001$) and number of sodas consumed on soda use days (RR = 1.06, $p = .03$) for individuals completing the survey on Mondays. These effects were of a small effect size and importantly did not alter the

primary findings reported in this manuscript when included as covariates. Without experimentally manipulating day of completion, we also cannot be certain whether these effects were due to the time of assessment or orderly differences in individual characteristics related to day of completion. Future studies using past week recall would benefit from using a staggered day of completion to decrease any systematic bias in behavioral report.

Third, we used inclusion criteria of 95% approval rates and 100 or more previous mTurk tasks that could have increased the reliability of our group of participants and inflated some findings (e.g., response rates). Restricting participants based on approval rates and/or previous tasks completed is common in mTurk studies (e.g., Cunningham et al. 2017; Reed et al. 2016; Strickland and Stoops 2015). Previous research has shown that these restrictions can improve data quality (e.g., participants are less likely to demonstrate central-tendency biases or fail attention checks) as well as result in lower rates of socially desirable responding (Peer et al. 2014). This latter finding is particularly important for research with substance-using populations given the greater potential for socially desirable responses to questions about drug use and related health behaviors. A screening method based on prior approval and completion rates is also not unlike typical screening procedures in the human laboratory and clinical in which a participant must be sufficiently reliable to show up for one or more screening appointments prior to study enrollment. An important question in this regard is whether individuals who do and do not meet these mTurk screening criteria differ on demographic and drug use behaviors that would result in sampling bias. The aforementioned study did not observe differences in the age or gender distribution of participants based on either approval or response rate criteria (Peer et al. 2014). However, future work that systematically examines patterns of drug use as a function of these kinds of screening methods will be important for evaluating potential sampling bias.

Fourth, the online nature of mTurk sampling and testing means that drug use could not be biologically verified and cannot in future applications of this method. Related to this concern is the use of a relatively transparent screener in which the majority, although not all, of questions were related to alcohol use. The use of a more opaque screener including questions completely unrelated to alcohol or other drug use would help decrease the possibility of inadvertently revealing inclusion criteria and increasing related demand characteristics. Nevertheless, as noted above, numerous studies have revealed consistent findings between mTurk samples and in-person samples in which biological verification is possible. Approximately three-quarters of participants also indicated that they were more comfortable reporting sensitive material through mTurk than they would be in-person. This described comfort is consistent with reports that online technology can help reduce underreporting biases observed with heavily stigmatized behavior, such as substance use (Harrison & Hughes, 1997; Turner et al., 1998), and adds further support for this online approach.

The online nature of mTurk does also raise potential problems if participants experience inconsistent Internet access. This may be a particular concern for future studies in substance-using populations that may be more transient and have unreliable Internet sources. Digital divides are still evident in access to the Internet and other technology, however, some research in substance-using populations does suggest that some of these divides may be closing (e.g., Cunningham et al. 2006; McClure et al. 2013a; Strickland et al. 2015). A majority of participants also reported completing the baseline survey on a computer (93.2%) and a smaller percentage on a phone (4.0%) or tablet (2.9%). That some participants completed these surveys on their phone does indicate that this technology would be amenable to alternative platforms.

Recommendations and Areas for Future Research

This study provides initial evidence of the feasibility, acceptability, and validity of using mTurk for intensive data collection in the addiction and behavioral sciences. Future research would benefit from attention to and exploration of the methodological questions raised above in the context of the study limitations. For example, a raffle was used to help incentivize active participation. Several participants wrote positively of the raffle in qualitative data that was collected at the end of the study (data not shown). Systematic manipulations of raffles and other compensation would be of interest for future work and help to reveal the relative sensitivity of participants to the incentive structure and density of these incentive schedules.

Beyond addressing methodological questions, the longitudinal approach described here could be applied to diverse research interests in the experimental analysis of behavior beyond those related to alcohol or other substance use behaviors. Repeated measurement of individual participants through reversal and other ABAB-type designs is at the core of behavior analysis and its related experimental design. Individual participant plots in the current study highlight the rich variety of behavioral patterns that could be generated through longitudinal sampling on mTurk. Any behaviors that may be captured through self-report are feasible for mTurk data collection (e.g., health behaviors, daily social interactions). Recent advances in browser-based, open-source programming tools also mean that reaction time experiments and other behavioral tasks commonly used in behavior analysis (e.g., reinforcement learning tasks, delay discounting tasks) can be easily incorporated into longitudinal designs (e.g., De Leeuw 2015; Seithe et al. 2016; Stoet 2017). In this way, mTurk and other forms of online testing could be adapted to study how behavioral mechanisms commonly studied in the behavioral analytic literature (e.g., discounting, behavioral economic demand) translate into prospectively collected behavior in the natural ecology. Alternatively, the temporal

stability of certain behavioral tasks could be studied as well as the sensitivity of these mechanisms to self-reported environmental events (e.g., changes in daily mood or experiences of daily stress). With some creativity, along with recognition of limitations imposed by online research and the loss of tight experimental control sometimes experienced in the laboratory, we believe the methods presented here will benefit anyone interested in designing experiments that are relevant to the behavior of individual organisms.

Conclusion

The benefits of an online setting combined with the rapid rate of data collection will help complement traditional human laboratory and clinical procedures. In this way, the cost and time efficiency afforded by mTurk could provide a resource for generating pilot data that identifies outcomes or relationships of interest and that helps determine optimal parametric parameters for procedures prior to larger, and more expensive, in-person work. This study extended earlier work (Boynton and Richman 2014) by demonstrating the feasibility of data collection over long periods of time, establishing the acceptability of these study procedures, and determining the convergent and discriminant validity of intensive longitudinal alcohol self-report on mTurk. Future studies may leverage mTurk for generating large, geographically diverse samples for prospective research designs in the behavioral and addiction sciences.

Table 6.1. Participant Demographics

	Mean (SD)/%	IQR
Demographics		
Age	35.2 (10.6)	27-40
Male	44.6%	
White	82.7%	
College	68.4%	
Alcohol Use		
Drinks/Week	8.6 (9.3)	3-12
Days/Week	3.0 (2.0)	2-4
AUDIT	10.5 (7.9)	4-14
AUD	40.7%	
Abuse	9.0%	
Dependence	31.7%	
Soda Use		
Soda Drinker[a]	80.2%	
Drinks/Week[b]	7.6 (10.8)	1-10
Days/Week[b]	3.6 (2.5)	1-6
Cigarette Use		
Smoker[c]	31.3%	
CPD[d]	12.9 (7.8)	6-20
FTND[d]	4.3 (2.5)	2-6

Note. SD = standard deviation; IQR = interquartile range; College = college education or greater; AUDIT = Alcohol Use Disorder Identification Test; AUD = Alcohol Use Disorder (DSM-IV criteria); CPD = cigarettes/day; FTND = Fagerström Test for Nicotine Dependence.

[a]Reported drinking soda at any time on baseline survey
[b]Only participants reporting they drink soda on baseline survey
[c]Reported daily tobacco cigarette smoking on baseline survey
[d]Only participants reporting daily tobacco cigarette smoking on baseline survey

Table 6.2. Acceptability Measures (n = 213).

Question	% Endorse/ Mean (SD)
Overall, How Satisfied Were You with the Study Experience?	
Quite Satisfied	73.2%
Mildly Satisfied	20.7%
Mildly Dissatisfied	2.3%
Quite Dissatisfied	3.8%
Would you Participate Again?	
Definitely So	87.3%
Probably So	10.8%
Probably Not	1.4%
Definitely Not	0.5%
Motivations for Participating[a]	
To Gain Self-Knowledge	42.7%
To Kill Time	8.0%
Enjoy Doing Interesting Tasks	54.0%
To Make Money	82.6%
To Have Fun	15.5%
Experiences with mTurk Research	
I find it easier to answer honestly sensitive questions on mTurk compared to an interview	72.8%
I like the idea of participating in research on mTurk as much or more than participating in research in person	73.7%
I would never participate in a research study in person, but would on mTurk	21.6%
Would you participate in future studies on mTurk designed to reduce problem behaviors (e.g., alcohol use, cigarette use, overeating)?	93.0%
Study Acceptability Measures (0-100 VAS)	
Ease of Completion	92.9 (12.7)
Helpful Instructions	89.2 (16.0)
Enjoyable	78.1 (22.7)
Convenient Timing	86.1 (19.5)
Fair Compensation	84.3 (21.8)
Overall Experience	87.3 (16.8)

Note. SD = standard deviation
[a]Participants could select more than one motivation so endorsements will not total to 100%

Table 6.3. Correlates of Prospective Alcohol Use Behaviors

	Drank Alcohol		# of Drinks[a]		Heavy Drinking[a]	
	BI	ADJ	BI	ADJ	BI	ADJ
Between						
Age[c]	1.01	1.03**	0.99*	1.00	0.97*	1.00
Male	1.43	1.11	1.30***	1.17**	0.89	0.60*
White	0.71	0.67	1.02	1.03	0.87	0.84
College	1.04	1.32	0.70***	0.79***	0.23***	0.36***
AUDIT[c]	1.11***	1.13***	1.05***	1.05***	1.19***	1.19***
Within						
Weekend	2.93***	2.93***	1.34***	1.34***	2.63***	2.66***
Intercept	0.72**	0.49*	2.89***	2.85***	0.23***	0.56

Note. BI = bivariate models, single variable included; ADJ = adjusted models, all variables included. Reported are effect sizes for models (odds ratios for drank alcohol and heavy drinking; rate ratios for # of drinks).

[a]Models including days with alcohol use reported

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

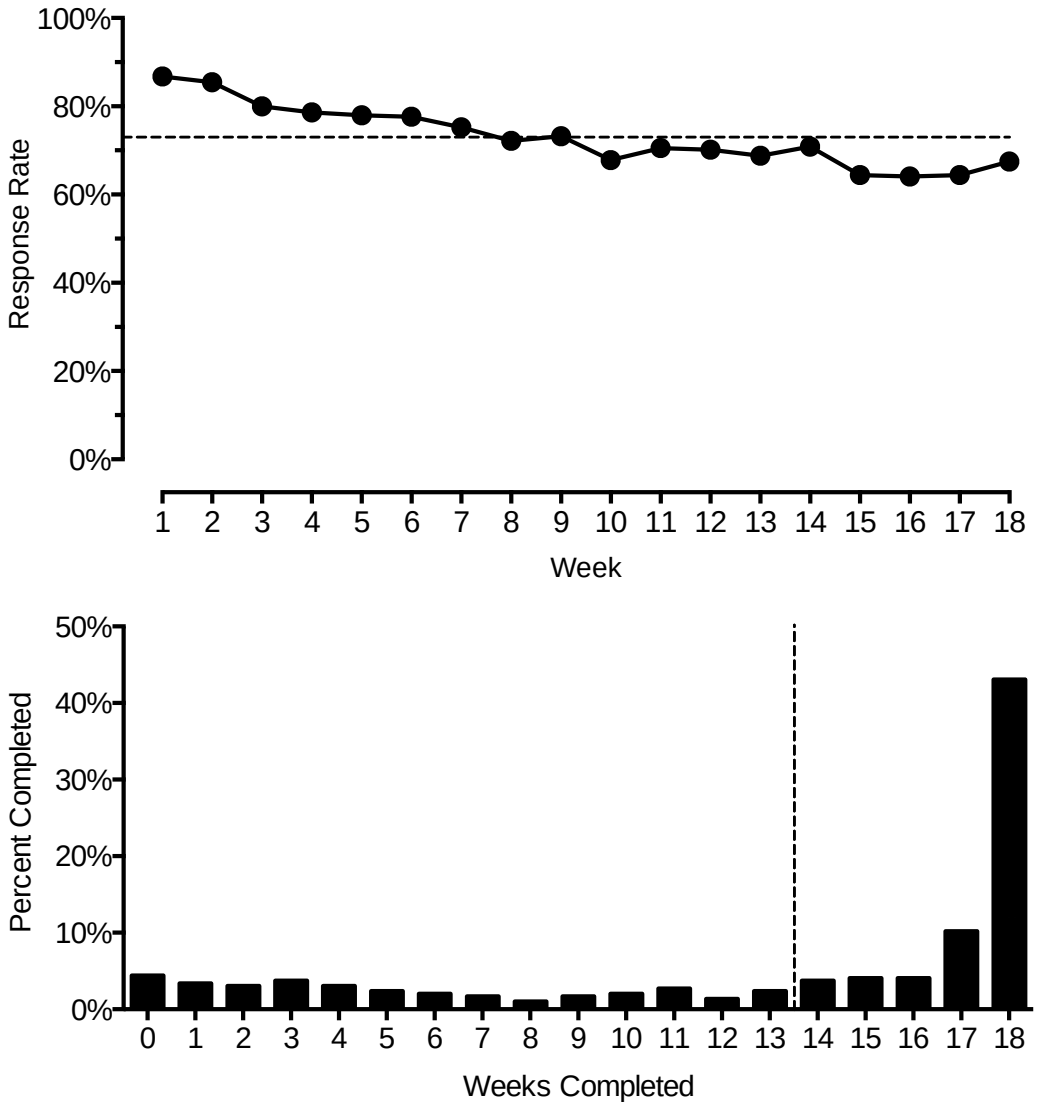

Figure 6.1. Weekly response rates across the 18-week longitudinal phase (top panel) and distribution of individual weeks completed (bottom panel). Top panel: Plotted are response rates for each week (including individuals that did not provide a weekly assessment or did not report alcohol use during the longitudinal period, n = 295). Dotted line is average response rate across the 18 weeks (73%). Bottom Panel: Plotted is the distribution of weeks completed. Vertical dotted line demarcates 14+ weeks (i.e., the necessary weeks for the raffle incentive.

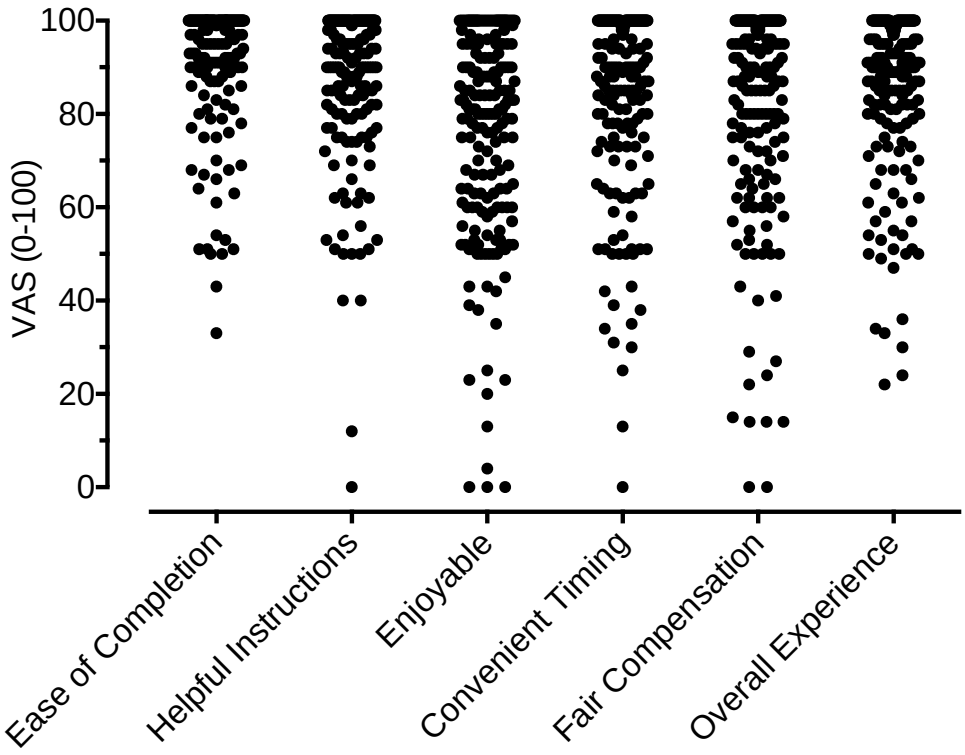

Figure 6.2. Study acceptability measures. Participants reported study acceptability measures at the end of the 18-week longitudinal phase. All items were completed on a 100-point visual analog scale (VAS). Plotted are individual participant data.

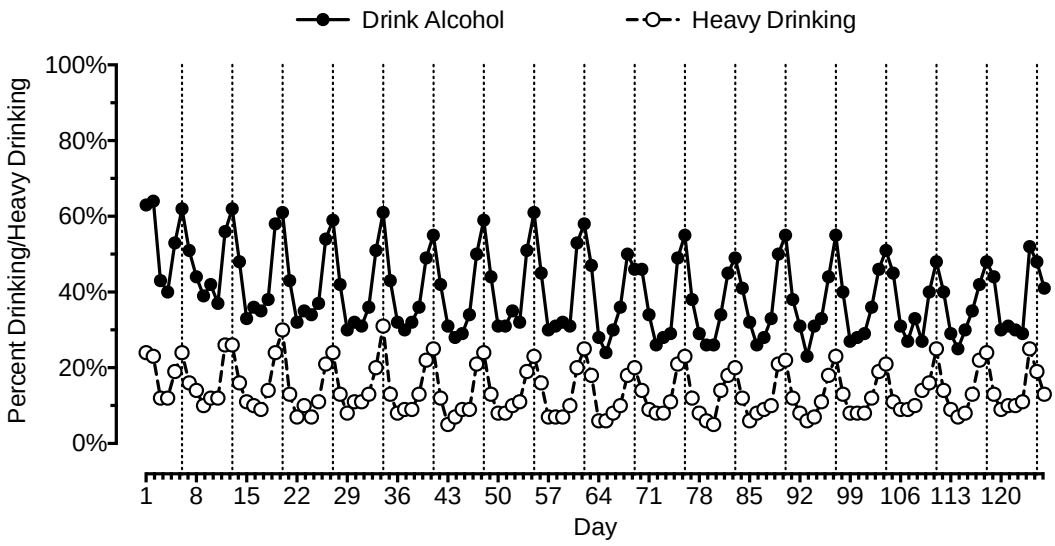

Figure 6.3. Daily fluctuations in alcohol use behaviors across the study. Plotted are the percent participants reporting alcohol use (solid line and closed circle) and participants reporting heavy alcohol use (dotted line and open circle). Each tick on the x-axis refers to a study day (Day 1-126). Vertical dotted lines are Saturdays. Numerically labeled days are Mondays.

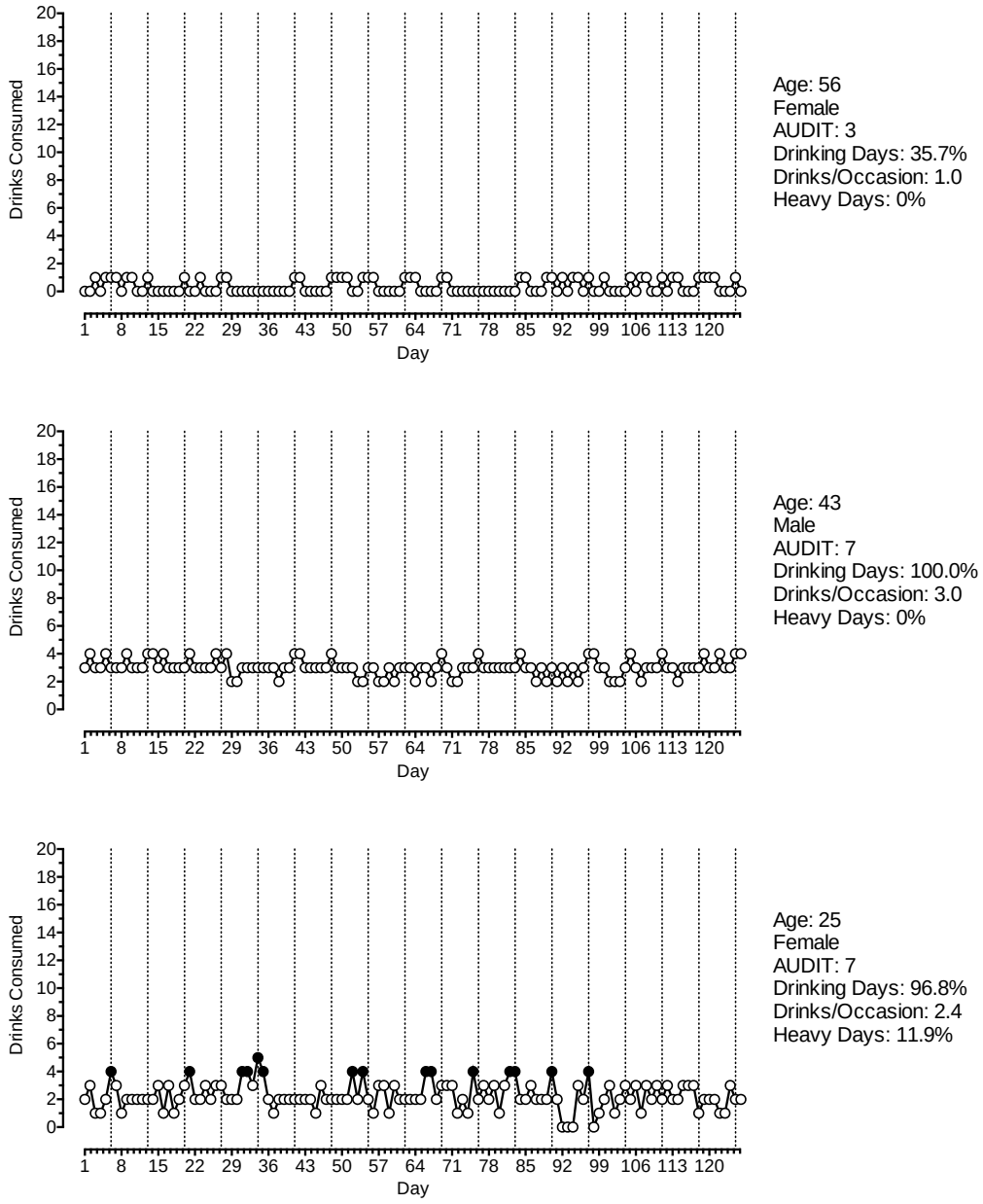

Figure 6.4. Low-risk alcohol use individual participant plots. Plotted are individual participant data for individuals reporting low-risk alcohol use according to AUDIT cutoff criteria (AUDIT < 8). Data points represent daily alcohol consumption (number of drinks) with filled circles representing a heavy use day. Each tick on the x-axis refers to a study day (Day 1-126). Vertical dotted lines are Saturdays. Numerically labeled days are Mondays.

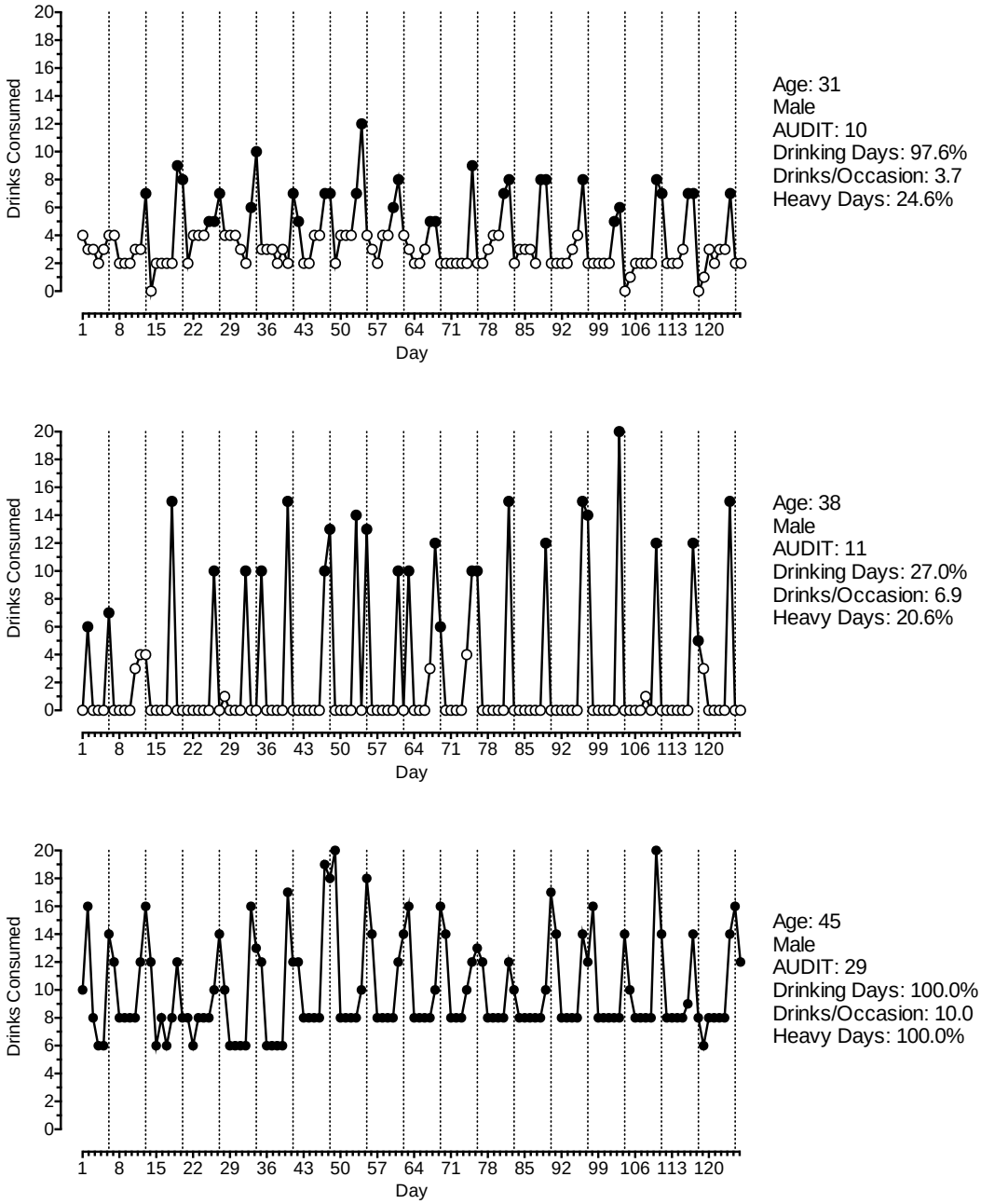

Figure 6.5. High-risk alcohol use individual participant plots. Plotted are individual participant data for individuals reporting high-risk alcohol use according to AUDIT cutoff criteria (AUDIT ≥ 8). Data points represent daily alcohol consumption (number of drinks) with filled circles representing a heavy use day. Each tick on the x-axis refers to a study day (Day 1-126). Vertical dotted lines are Saturdays. Numerically labeled days are Mondays.

Chapter 7

USING BEHAVIORAL ECONOMIC VARIABLES TO PREDICT FUTURE ALCOHOL USE IN A CROWDSOURCED SAMPLE

(Experiment 4b; Strickland et al., under review)

Introduction

Alcohol use disorder is a persistent public health concern. Approximately 14.0 million Americans met criteria for alcohol use disorder in 2017, with the annual economic impact of excessive drinking estimated at $250 billion (Center for Behavioral Health Statistics 2018; Sacks et al. 2015). Alcohol can interact with other licit (e.g., cigarettes) and illicit (e.g., cocaine) substances to increase health risk (e.g., cardiovascular toxicity) (Durazzo et al. 2004; Farre et al. 1997). Approximately half of all violent crimes in the United States involve alcohol consumption and problem drinking plays a particularly salient role in cases of domestic abuse and intimate partner violence (Abbey et al. 2001; Foran and O'Leary 2008; Quigley and Leonard 2000). These evident economic, health, and social implications of alcohol consumption highlight the importance of understanding person-level predictors of alcohol use to inform prevention and treatment efforts.

The mixing of theoretical perspectives from behavioral economics and operant theory has resulted in numerous advances for addiction science, broadly (Bickel et al. 2014; Bickel et al. 2016a) and alcohol research, specifically (MacKillop 2016). Such models propose three core behavioral economic mechanisms, behavioral economic demand, delay discounting, and proportionate alcohol-related reinforcement, which may relate to alcohol use disorder and the development and persistence of problematic alcohol and other substance use (MacKillop 2016).

First, behavioral economic demand refers to the orderly relationship between alcohol consumption and price (Murphy et al. 2009). Demand is commonly measured using commodity purchase tasks wherein participants are asked to report consumption of a

good (e.g., alcohol) across a range of prices (e.g., $0.01, $1.00, $10.00/drink) (Jacobs and Bickel 1999; Reed et al. 2013). This methodology is particularly appealing because of its cost and time efficiency and adaptability for populations with whom drug self-administration is not practically or ethically feasible, such as treatment-seeking patients or those with medical contraindications. Second, delay discounting is the systematic reduction in the value of a reinforcer as a function of the delay to its delivery (Odum 2011; Rachlin and Green 1972). Excessive delay discounting is thought to play a central role in alcohol use disorder and may represent a trans-disease process relating to other substance use and maladaptive health behaviors (Bickel et al. 2012; Koffarnus et al. 2013). Third, proportionate alcohol-related reinforcement is a measure based on the matching law that indexes the relative reinforcement in an individual's daily life that is attributed to alcohol use (Correia et al. 1998; Murphy et al. 2015). This measure provides a more molar determinant of alcohol use compared to alcohol demand by emphasizing the presence and value of alcohol consumption in relation to daily activities.

Although these behavioral economic mechanisms have been extensively studied in the laboratory and clinic using cross-sectional designs, far fewer studies have evaluated their unique relevance for predicting longitudinal patterns of alcohol use. Existing evidence highlights the importance of delay discounting for predicting trajectories of alcohol involvement throughout adolescence and young adulthood (Brody et al. 2014; Fernie et al. 2013). Similarly, a growing literature has identified alcohol demand and proportionate alcohol-related reinforcement as prognostic variables predicting treatment success (e.g., reductions in heavy drinking episodes) following brief interventions targeting alcohol consumption among college students (Dennhardt et al. 2015; MacKillop and Murphy 2007; Murphy et al. 2005; Murphy et al. 2015). These studies provide preliminary support for the importance of behavioral economic variables in predicting future alcohol use behaviors. However, additional research is needed to replicate and

expand this initial work to determine the unique predictive contribution that each mechanism may provide when collected in community samples as well as those outside of an intervention context. Determining these associations in such naturalistic settings is particularly important given the prominent and unique role proposed for behavioral economics in theoretical models of alcohol and substance use (e.g., "reinforcer pathology" models Bickel et al. 2017) as well as empirical evidence supporting unique associations with alcohol and other substance use behaviors at a cross-sectional level (e.g., Acuff et al. 2018; Aston et al. 2016; MacKillop et al. 2010a; Strickland et al. 2017b).

One reason for the relatively limited research on the prediction of behavior in the behavioral economic literature is the cost, time, and geographic constraints related to conducting longitudinal research. An emerging method positioned to address these concerns is crowdsourced sampling (see reviews by Chandler and Shapiro 2016; Strickland and Stoops 2019). Crowdsourcing, such as on Amazon Mechanical Turk (mTurk), allows for the effective and efficient sampling of research participants from diverse geographic regions and with varying alcohol and substance use histories. This is achieved through the posting of a flexible, open call to complete tasks (such as research studies) to the pool of individuals located across the country who are participating on mTurk. Recent research has supported the use of mTurk in addiction science by demonstrating a correspondence between outcomes on mTurk and those obtained in the human laboratory and clinic (e.g., Johnson et al. 2015; Kim and Hodgins 2017; Strickland et al. 2016a). Two recent studies have also demonstrated the feasibility, acceptability, and validity of collecting longitudinal alcohol use data with mTurk samples (Boynton and Richman 2014; Strickland and Stoops 2018b). The first study included daily reports of alcohol use collected over a 14-day period and demonstrated the validity of data collection by replicating typical relationships observed in the alcohol literature

(e.g., heavier drinking on the weekends) (Boynton and Richman 2014). The second study extended these findings by collecting weekly reports of alcohol and soda use over a longer 18-week period (Strickland and Stoops 2018b). Validity and measurement selectivity was also observed in that study with expected relationships involving alcohol consumption that did not extend to soda use.

This overview has highlighted the relevance of behavioral economic measurement in theoretical models of alcohol and substance use as well as critical gaps related to the predictive validity of these behavioral economic variables. A clear rationale was also provided for using crowdsourcing for collecting longitudinal data to this end. The purpose of the current analysis was to evaluate the unique relationship between behavioral economic mechanisms and self-reported future alcohol use. These data were collected as a part of the aforementioned 18-week mTurk study in which participants reported daily alcohol and soda use at weekly intervals (Strickland and Stoops 2018b). The hypotheses were that behavioral economic measures would: 1) associate with alcohol and soda use variables collected at baseline in a stimulus-selective manner (i.e., alcohol use variables associating with alcohol, but not soda use variables), 2) uniquely associate with prospectively collected alcohol use frequency, quantity, and severity outcomes, and 3) show test-retest reliabilities consistent with previous in-person research.

Methods

Participants and Screening

Participants were recruited from mTurk and all surveys were hosted on Qualtrics (Provo, UT, USA). Participants were required to have completed at least 100 mTurk tasks, have a 95% approval rating or higher, and reside in the United States to view the study (see similar qualification restrictions in Cunningham et al. 2017; Reed et al. 2016; Strickland and Stoops 2017). A screening questionnaire was used to determine study eligibility. Inclusion criteria were: 1) age 21 or older, 2) Alcohol Use Disorder

Identification Test (AUDIT) score of 1 or higher (Saunders et al., 1993), 3) self-reported alcohol use in the week prior to screening, and 4) willingness to complete an 18-week study. These inclusion criteria were designed to capture a wide range of individuals with only the limiting constraint of weekly alcohol use (i.e., the population of interest was a community sample of weekly alcohol consumers from the United States). This population was selected to facilitate the evaluation of individual differences as they relate to the behavioral economic measures studied and to provide the necessary variance in alcohol consumption patterns to detect such effects. No inclusion or exclusion criteria were included regarding treatment-seeking status or the level of hazardous or at-risk drinking.

General Procedures

Qualifying participants first completed a baseline survey that included demographic, alcohol and soda use, and behavioral economic measures. Next participants completed a longitudinal phase consisting of 18 weekly surveys in which participants recorded past week alcohol and soda use behaviors by day. The average response rate during this period was 73% (range: 64.1%-86.8% each week). One week after the longitudinal phase, participants were asked to complete a follow up including the baseline behavioral economic measures. For additional details on the study design and feasibility, acceptability, and validity of this data collection see Strickland and Stoops (2018b).

Study Measures

Behavioral Economic Demand

Commodity purchase tasks were used to evaluate behavioral economic demand for alcohol and soda (Morris et al. 2017; MacKillop and Murphy 2007; Strickland and Stoops 2017). Each task presented a similar vignette (see Appendix for sample vignettes). Participants were asked to imagine a typical day over the last month when they used each commodity. In each task, participants were told that they could only get the

commodity from this source, could not stockpile, had no commodity saved from previous days, and would have to consume all purchases in a single day (i.e., 24 hour period). Participants were required to correctly answer questions related to these instructions to verify understanding. Participants were then asked how many drinks (one US standard drink or one 12 oz. serving of soda) they would purchase at 13 monetary increments ranging from $0.00 [free] to $11/unit, presented sequentially (full range: $0.00 [free], $0.01, $0.05, $0.13, $0.25, $0.50, $1, $2, $3, $4, $5, $6, $11).

Price intensity and elasticity were generated from purchase task data using the exponentiated demand equation (Koffarnus et al. 2015):

$$Q = Q_0 * 10^{k(e^{(-\alpha * Q0 * C)} - 1)}$$

where Q = consumption; Q_0 = derived demand intensity; k = a constant related to consumption range (*a priori* set to 2); C = commodity price; and α = derived demand elasticity. Demand intensity refers to the theoretical consumption of a commodity at a unit price of zero (i.e., free). Demand elasticity reflects the sensitivity of consumption to changes in price. Group level purchase task data showed prototypic decreases in consumption with increases in price (Figure 7.1). The exponentiated demand equation also provided an excellent fit to group data (fit for mean demand data R^2: Alcohol = .99; Soda = .99) and individual data (mean of individual demand curve fits R^2: Alcohol = 0.87; Soda = 0.91). Intensity and elasticity were selected as the primary outcomes because prior factor analytic studies have demonstrated that these measures reflect the two factors underlying the purchase task factor structure for alcohol and other substances (Aston et al. 2017; Bidwell et al. 2012; Epstein et al. 2018; Mackillop et al. 2009). Recent evidence also suggests that these derived measures show greater stimulus-selectivity than other purchase task measures (e.g., breakpoint) (Strickland and Stoops 2017).

Demand intensity and elasticity were log-transformed prior to analysis to achieve normality.

Delay Discounting

Delay discounting rates for money were determined using a 5-trial adjusting delay task (Koffarnus and Bickel 2014). Prior research has validated this task against traditional adjusting amount delay discounting tasks (Cox and Dallery 2016; Koffarnus and Bickel 2014). This task version was selected given its benefits for the online setting, including rapid assessment with minimal computing requirements. Participants were instructed to select between $1000 at a delay and $500 available immediately. The first choice was at three-weeks delay, which then adjusted up (longer delay following delayed choice) or down (shorter delay following immediate choice) following each choice. An effective delay 50% (ED_{50}) was determined following five choices across 32 potential delays between 1 hour and 25 years. The primary outcome was delay discounting rates (k) calculated as the inverse of ED_{50} (Koffarnus and Bickel 2014). Delay discounting rates were log-transformed prior to analysis to achieve normality.

Proportionate Alcohol-Related Reinforcement

The Reinforcement Survey Schedule-Alcohol Use Version was used to evaluate proportionate alcohol-related reinforcement (Morris et al. 2017; Murphy et al. 2005). The current study used a 33-item version described by Morris and colleagues (2017). The 33-item version showed good internal consistency (Cronbach's α = .89 to .97 across varying age groups) and construct validity in that previous study, which was also conducted in an online setting. This measure included activities that one might experience over a 30-day period (e.g., go out to eat) that participants were asked to rate on frequency and enjoyability when 1) not drinking alcohol and 2) drinking alcohol. Frequency and enjoyability ratings were multiplied for each item to create a cross-product score. The primary measure was the R-ratio reflecting the ratio of total alcohol-

related reinforcement to total reinforcement (i.e., alcohol-free plus alcohol-related reinforcement). R-ratios were approximately normal and did not require transformation.

Alcohol and Soda Use History

A battery of standardized alcohol use measures was used to index alcohol and other substance use (e.g., AUDIT (Saunders et al. 1993), DSM-IV criteria for alcohol use abuse or dependence (Sheehan et al. 1998)). Individuals were evaluated with DSM-IV criteria because at the time of study design and execution, a brief and validated screening questionnaire was not available for DSM-5 (but see Hagman 2017, for a recently developed measure that could serve this purpose). Retrospective reporting of frequency and quantity of typical alcohol and soda use were also collected (e.g., alcohol drinks/sodas per week). All alcohol referred to one US standard drink and all soda referred to a 12 oz. serving.

Data Analysis

Three hundred and seven participants completed the baseline survey. Thirty participants failed one or more data quality checks throughout the study, did not provide any assessments during the longitudinal phase, and/or did not report drinking alcohol during the longitudinal phase and were removed from initial data analysis (n = 277). Purchase task data were then evaluated for systematic data using standard criteria (Stein et al. 2015). Fifty participants provided non-systematic data either violating these criteria (n = 19) or reporting zero consumption at all prices (n = 31) and four participants did not complete the alcohol purchase task data due to a technical error. This resulted in a final sample of 223 participants with systematic study and alcohol purchase task data. Analyses focused on this sample given the primary hypotheses related to unique prediction by each behavioral economic variable. Sensitivity analyses were conducted to determine the impact the removal of participants could have had on the study outcomes. These analyses compared the demographic and alcohol use characteristics between

those included and not included (233 included versus 74 excluded). These comparisons did not reveal significant differences in demographics, alcohol use, discounting rates, or R-ratio scores suggesting that the sample characteristics were not compromised by these participants' removal. Additionally, of the 223 participants, 166 reported drinking soda regularly and provided systematic soda purchase task data. This sample of 166 was considered for all analyses involving soda demand and/or soda use variables.

First, baseline associations between behavioral economic measures and retrospectively collected alcohol and soda use variables were tested using bivariate correlations (Hypothesis 1). Next, the unique relationship between behavioral economic measures and prospectively collected alcohol use was evaluated using generalized linear mixed effect models (Hypothesis 2). Three outcomes were tested: A) alcohol use (dichotomous), B) drink number (count with a negative binomial distribution), and C) heavy drinking day (dichotomous). Drink number and heavy drinking day models were estimated for drinking days only so as to not conflate drinking frequency (Outcome A) with quantity (Outcome B) and severity (Outcome C). Alcohol use referred to the presence or absence of drinking on a given study day. Drink number referred to the number of drinks consumed on a drinking day. Heavy drinking was defined using National Institute on Alcohol Abuse and Alcoholism (NIAAA) guidelines of ≥5/≥4 drinks/day for men/women (National Institute on Alcohol Abuse Alcoholism 2007). Models were tested in three steps. First, unadjusted models including a single behavioral economic predictor were estimated. Then, models were adjusted for AUDIT scores to determine incremental validity. These AUDIT incremental validity tests were designed to determine whether the relationships between behavioral economic variables and alcohol frequency, quantity, and severity remained after controlling for a gold standard measure of alcohol use. These tests were necessary to determine whether each behavioral economic mechanism provide information above and beyond traditional representations

of alcohol consumption. Finally, unique prediction was tested in a full model containing all behavioral economic predictors and AUDIT scores. Continuous predictors were grand-mean centered prior to analysis. Analyses described in a previous report suggested that attrition across the longitudinal phase was not systematically related to alcohol use behavior (Strickland and Stoops 2018b). Accordingly, data were treated as missing at random (Singer and Willett 2003). Effect size estimates were interpreted for individual models (odds ratios [OR] for dichotomous outcomes and rate ratios [RR] for count outcomes).

Finally, the reliability and temporal stability of each behavioral economic measure was determined (Hypothesis 3). Test-retest reliability (r_{xx}) was measured using bivariate correlations. Temporal stability was evaluated using dependent-samples *t*-tests and Cohen's d_z effect size estimates. These tests were only conducted for individuals providing follow-up data (n = 150 of 223). Fifteen participants in the follow up sample also did not provide analyzable data for the alcohol purchase task (11 due to all zero consumption and 4 due to non-systematic data).

Inferential tests were two tailed with an alpha rate of .05. Maximum likelihood estimation using a Laplace approximation was used for generalized linear mixed effect models. All models were evaluated using *R* statistical language with the *glmmTMB* and *lme4* packages (Bates et al. 2014; Brooks et al. 2017).

Results

Sample Characteristics

Table 7.1 contains demographic variables. The majority of participants were white and employed with a college education. Approximately half of the participants were female and the average age was 35.2 years old.

Table 7.1 also includes alcohol and soda use variables. Participants reported an average of 8.8 standard drinks per week and 3 drinking days per week. Approximately

41% of participants met criteria for alcohol abuse or dependence and 52.9% met AUDIT criteria of hazardous drinking (AUDIT of 8+).

Behavioral Economic Outcomes and Baseline Associations

Bivariate correlations between baseline behavioral economic outcomes and demographic and retrospectively collected alcohol and soda use variables are presented in Table 7.2. Also presented in Table 7.2 are inter-correlations between baseline behavioral economic outcomes. Alcohol and soda demand intensity (r = .57, p < .001) and elasticity (r = .60; p < .001) were significantly and positively correlated.

Commodity-similar demand relationships were generally in the expected direction, statistically significant, and of a medium-to-large effect size (e.g., alcohol and soda demand intensity were positively associated with alcohol and soda drinks/week, respectively; see Table 7.2 for correlations). In contrast, commodity-different demand relationships (e.g., alcohol demand and soda consumption) were generally not statistically significant and/or were of a small effect size. For example, the strongest association among all commodity-different demand relationships was a small-to-medium effect size between soda demand intensity and AUDIT scores (r = .23, p = .003).

R-ratio scores showed medium-sized and significant correlations with all alcohol variables except alcohol drinks/occasion (significant r values .25-.32, p values < .05). R-ratio scores were not significantly related to soda use variables.

Delay discounting rates showed significant, but small effect size relationships with the presence of alcohol abuse or dependence (r = .14, p = .04), AUDIT scores (r = .15, p = .03), and soda drinks/occasion (r = .16, p = .05). Other correlations with alcohol or soda use involving discounting rates were not statistically significant.

Prediction of Alcohol Use by Behavioral Economic Variables

Table 7.3 contains effect size estimates and statistical significance for unadjusted and adjusted comparisons between behavioral economic predictors and alcohol use frequency, quantity, and severity measures during the longitudinal phase.

Unadjusted models predicting drinking frequency (i.e., presence or absence of a drinking day) indicated greater odds of a drinking day for individuals reporting higher alcohol intensity and R-ratios and lower elasticity (more inelastic demand). Only the R-ratio association remained significant in models adjusting for AUDIT scores and in the full model including all predictors and AUDIT scores.

Unadjusted models predicting drinking quantity (i.e., number of drinks on drinking days) also found a higher rate of drinking for individuals reporting higher alcohol demand intensity and R-ratios and lower alcohol demand elasticity. Higher rates of drinking were also observed for individuals with higher soda demand intensity and lower soda demand elasticity in unadjusted comparisons. Higher rates of drinking for individuals with higher alcohol demand intensity and lower demand elasticity remained in models adjusting for AUDIT scores. Only the alcohol intensity effect remained significant in a final model including all predictors, with higher rates of drinking for individuals with higher demand intensity. A significant effect of soda demand intensity was also observed in this final model, however in the opposite direction of that observed in unadjusted comparisons (i.e., higher soda intensity associated with lower rates of drinking when controlling for other behavioral economic variables and AUDIT scores).

Unadjusted models predicting drinking severity (i.e., presence or absence of heavy drinking on drinking days) found a greater odds of heavy drinking for individuals reporting higher discounting rates and alcohol demand intensity as well as lower alcohol demand elasticity. Greater odds of heavy drinking were also observed for individuals with higher soda demand intensity and lower soda demand elasticity in these unadjusted

comparisons. The effects of alcohol intensity and elasticity remained significant in models adjusting for AUDIT scores, but only the alcohol intensity effect remained significant in a final model including all predictors. A significant effect of soda demand intensity was also observed in this final model, however in the opposite direction of that observed in unadjusted comparisons (i.e., higher soda intensity associated with lower odds of heavy drinking when controlling for other behavioral economic variables and AUDIT scores).

AUDIT scores were a significant predictor of all three outcomes in all models tested (all *p* values < .001) with greater AUDIT scores predicting greater drinking frequency, quantity, and severity. Inclusion of soda demand variables in fully adjusted models did not change the direction or significance of estimated effects for other behavioral economic variables. Similarly, inclusion of demographic variables (i.e., age, sex, education, and income) did not change the significance or direction of effects in these fully adjusted models. Inclusion of income in unadjusted and AUDIT-adjusted models similarly did not change the significance or direction of effects.

Additional models were also explored evaluating whether the relationships between soda demand and measures of alcohol frequency, quantity, and severity differed as a function of mixed drink consumption. These models tested the interaction between subjective feeling about mixed drinks (i.e., "What describes how much you like mixed drinks"; Not at all to Very Much) and soda demand intensity/elasticity. These interactions were not statistically significant indicating that the relationship between soda demand and alcohol consumption did not vary as a function of subjective liking of mixed drinks.

Test-Retest Reliability and Stability

Table 7.4 contains test-retest reliabilities and estimates of measurement stability. Statistically significant test-retest reliabilities were observed for all behavioral economic measures, *p* values < .001. The highest reliability was observed for *k* values (r_{xx} = .76)

and the lowest for R-ratios (r_{xx} = .29). Reliabilities for demand intensity were higher than elasticity for alcohol and soda commodities.

Alcohol demand elasticity, R-ratios, soda demand intensity, and soda demand elasticity showed measurement stability (i.e., did not significantly change over the 18-week period) (p values > .05). Significant decreases in k values and alcohol demand intensity were observed. Both changes were of a small effect size (d_z = .24 and .21, respectively).

Discussion

The purpose of the present analysis was to evaluate the association of behavioral economic demand, delay discounting, and proportionate alcohol-related reinforcement with alcohol use frequency, quantity, and severity variables reported retrospectively and during a prospectively collected 18-week period. Baseline patterns of retrospectively reported behavior were consistent with our research hypotheses in that behavioral economic measures closely associated with alcohol and soda use variables in a stimulus-selective manner. Specifically, measures specific to alcohol use (i.e., alcohol demand intensity or elasticity and proportionate alcohol-related reinforcement) were correlated with the majority of alcohol use variables at a medium-to-large effect size. In contrast, associations with soda use outcomes were of a smaller effect size and in most cases not statistically significant. This reciprocal selectivity replicates that of other studies demonstrating the stimulus-selectivity of alcohol and cigarette purchase tasks (Chase et al. 2013; Strickland and Stoops 2017). Such findings collectively support the domain-specific validity of the purchase task methodology for specifically studying drug valuation by demonstrating that the commodity available in single-commodity tasks is the primary determinant of behavioral allocation and demand.

Large effect size correlations were also observed between intensity (r = .57) and elasticity (r = .60) across the alcohol and soda purchase tasks. Similar correlations were

observed in a prior study evaluating alcohol and soda demand as well as cigarette and chocolate demand (Strickland and Stoops 2017; but see Chase et al. 2013). These findings could indicate an overlap in general reinforcer valuation that is reflected in a shared variance for the demand intensity and elasticity measures. Alcohol and soda specifically present similar reinforcer profiles as calorie dense and immediate reinforcers, which may contribute to the correlations in demand valuation observed here. Preclinical and human laboratory research has also identified sweet taste preference and liking as risk factors for alcohol use disorder thereby suggesting a genetic link between sweet taste reactivity and alcohol consumption (Kampov-Polevoy et al. 1999; Kampov-Polevoy et al. 2004; but see Kranzler et al. 2001). Taken together, the current observation of inter-correlated demand measures across commodity type is consistent with prior findings insofar as signifying a shared variance in reinforcer valuation.

Unadjusted comparisons predicting future alcohol use showed significant associations including alcohol demand intensity, elasticity, and R-ratio scores with measures of alcohol use frequency, quantity, and severity. These findings are consistent with the baseline findings reported here as well as other cross-sectional research reported elsewhere (Bertholet et al. 2015; Morris et al. 2017; Murphy and MacKillop 2006; Murphy et al. 2005; Murphy et al. 2015; see review by MacKillop 2016). In particular, these findings are consistent with a recent structural equation modeling study demonstrating the unique association of proportionate alcohol-related reinforcement and alcohol demand with alcohol consumption and related problems in a college-sample of heavy drinkers (Acuff et al. 2018). After adjusting for AUDIT scores, significant associations remained between proportionate-related alcohol reinforcement and alcohol use frequency and between alcohol demand and alcohol use quantity and severity. These findings suggest that increased proportionate-related alcohol reinforcement and behavioral economic demand may uniquely predict differing aspects of alcohol

consumption, namely the frequency of use in the former case and the quantity-related severity measures in the latter. These findings also demonstrate that these behavioral economic measures provide incremental and predictive validity for determining future alcohol consumption above and beyond a commonly used, gold standard measure of problematic alcohol use (AUDIT; Meneses-Gaya et al. 2009; Reinert and Allen 2007). This finding is particularly notable given a recent meta-analysis challenging the incremental validity of the alcohol purchase task for predicting alcohol use and severity beyond traditionally collected measures (Kiselica et al. 2016). That incremental validity was observed indicates that these behavioral economic mechanisms provide unique information about specific aspects of and patterns relevant to alcohol consumption.

Delay discounting rates only modestly predicted heavy drinking in unadjusted models and did not uniquely associate with any alcohol use outcomes during the longitudinal period after accounting for AUDIT scores or other behavioral economic measures. This outcome was not unanticipated given that only monetary discounting rates were collected. Previous work on delay discounting has established the importance of the commodity discounted by showing that commodity-relevant discounting rates provide improved prediction of substance use (Strickland et al. 2017b; Tsukayama and Duckworth 2010) and other health behaviors (Johnson and Bruner 2012; Rasmussen et al. 2010). The absence of a predictive association involving delay discounting rates as well as the associations involving alcohol demand and proportionate-related alcohol reinforcement are also consistent with existing cross-sectional work (Acuff et al. 2018) and longitudinal work evaluating these variables as prognostic variables in brief alcohol interventions (Dennhardt et al. 2015; MacKillop and Murphy 2007; Murphy et al. 2005; Murphy et al. 2015). Dennhardt and colleagues (2015) for example, found that baseline alcohol demand intensity predicted binge drinking and alcohol-related problems at 6-months following a brief alcohol intervention in college students, whereas delay

discounting rates did not significantly predict any study outcomes. Important to note is the distinction between the models tested in this study and those in prior longitudinal research. Specifically, the models used here evaluated use over time whereas those used in previous longitudinal work evaluated behavioral economic variables as predictors of changes in alcohol consumption following brief interventions (Dennhardt et al. 2015; MacKillop and Murphy 2007; Murphy et al. 2005; Murphy et al. 2015). That only delay discounting for monetary goods was collected is a limitation of the current analysis and future work would benefit from evaluating commodity-specific discounting rates.

The stimulus selectivity of the predictive relationships was also explored by including behavioral economic measures relevant to soda use. Increased soda demand intensity and decreased elasticity were modestly related to increased quantity and severity of alcohol use in unadjusted models. However, these associations were of a smaller effect size than those of alcohol demand and did not remain significant in models accounting for AUDIT scores. Interestingly, in models accounting for all behavioral economic variables, increased soda demand intensity was associated with lower rates of drinking quantity and lower odds of heavy drinking. This could indicate a behavioral economic substitution mechanism in which soda consumption increases with decreases in alcohol use (and vice versa), presumably due to increasing cost (for examples of cross-commodity research with drug commodities see (for examples of cross-commodity research with drug commodities see Johnson et al. 2017b; Murphy et al. 2016; Peters et al. 2017; Snider et al. 2017). It is also possible that this inverse relationship represents a narrowing of behavioral repertoire related to exclusive valuation for specific reinforcers at higher demand valuation (Koob et al. 1998). Future investigations could test these hypotheses, for example by using cross-commodity tasks to determine the cross-price elasticity of alcohol and soda.

Temporal reliability and stability were evaluated in participants completing a post-study follow up. Modest-to-strong support for test-retest reliability was observed for most measures over this approximately 18-week period. Reliabilities gathered through this online platform were also remarkably consistent with those observed for previous in-person research. Reliabilities for alcohol demand intensity, for example, have been reported as .89 at two weeks (Murphy et al. 2009) and .73 at one month (Acuff and Murphy 2017), which are similar to our reliability of .69 when considering the longer time interval (i.e., 4+ months). Research evaluating alcohol and alternative drug commodities (e.g., cigarettes) has also found higher reliabilities for demand intensity than elasticity, which is also consistent with the results reported here (Few et al. 2012). A similar consistency between the reliability we observed for delay discounting rates and those reported in research using varying populations, time intervals, and methodologies is also apparent (Anokhin et al. 2015, [rxx = .67-.76]; Baker et al. 2003, [rxx = .71-.90]; Beck and Triplett 2009, [rxx = .64]; Matusiewicz et al. 2013, [rxx = .70]; Ohmura et al. 2006, [rxx = .60]; but see higher reliabilities reported in Simpson and Vuchinich 2000, [rxx = .91]; Weafer et al. 2013, [rxx = .89 at one week]). This correspondence is particularly encouraging given that, to our knowledge, this is the first demonstration of the temporal reliability of the 5-choice task and suggests that reliability is not noticeably compromised when using this rapid assessment technique.

Significant reductions in alcohol demand intensity and delay discounting rates were observed over the 18 weeks, but these changes were of a small effect size and did not appear to impact temporal reliabilities. It is possible that these changes reflected reactivity owing to the self-monitoring of alcohol use for 18 weeks (e.g., Collins et al. 1998; Fremouw and Brown 1980; Uchalik 1979; but see Litt et al. 1998; Sobell et al. 1996). However, correlations between the number of weeks of data collection a participant completed and changes in demand intensity (r = .05) and discounting (r =

.02) were not statistically significant advising against this explanation. Other research has also noted modest changes in discounting rates over time, albeit when recorded in adolescents and over a longer temporal span than used here (e.g., from age 16 to 18) (Anokhin et al. 2015). It is also unclear why reliabilities for the R-ratio scores were substantially lower than those reported in prior research (Hallgren et al. 2016). It is possible that the longer temporal window captured in this study compared to prior research (4+ months versus 2-3 days) and/or the different populations sampled (an online community sample versus college student sample) could have contributed to this difference. The use of a 33-item rather than 45-item version could have also reduced the reliability of the assessment. It is also possible that the online format could have reduced reliability. However, this possibility would not explain why reliabilities consistent with in-person laboratory research were observed for other behavioral economic tasks completed in this online setting. Future studies evaluating these possibilities will be important for establishing the temporal reliability of proportionate alcohol-related reinforcement in community samples. These discrepancies outstanding, these findings collectively support the reliability of common behavioral economic measures when collected through an online crowdsourcing platform.

A central question regarding the association of behavioral economic mechanisms with alcohol and other substance use is the causal direction of this relationship. The majority of studies evaluating behavioral economic variables as a cause or consequence of substance use have focused on discounting of delayed rewards. These studies have revealed evidence for both mechanisms in that discounting may play an etiological role in substance use while also changing as a consequence of substance exposure (see review by Perry and Carroll 2008, regarding tobacco cigarette use). Additional support for an etiological role has been observed in longitudinal studies of alcohol use with steeper discounting predicting future alcohol consumption in adolescent populations

(MacKillop 2016). Preclinical animal work provides some evidence for the relevance of behavioral economic demand and reinforcement ratios as putative causal indicators. With respect to demand, animal laboratory research has demonstrated that baseline levels of cocaine demand can predict increased reinstatement responding and drug self-administration despite negative consequences (i.e., foot shock) (Bentzley et al. 2014). This literature also provides a rich history of research describing the ways in which variations in environmental enrichment and the availability of non-drug reinforcers may contribute to patterns of drug-taking behavior (see review by Bardo et al. 2013; Stairs and Bardo 2009). Reinforcement ratios may prove a more likely causal factor in humans given that its measurement indexes an individual's alcohol use placed within a broader context of environmental influences. However, behavioral economic demand may also represent an important causal indicator because it may reflect an underlying reinforcer sensitivity that is related to genetic and other risks contributing causally or in a causal pathway to substance use disorder. Establishing the relationship between behavioral economic measures and prospectively collected alcohol use variables as described in this study is a necessary, but not sufficient, step in determining a causal relationship between behavioral economic measures and substance use. Ultimately, long-term longitudinal research that evaluates individuals over the varied stages of alcohol and other substance use disorders (e.g., initiation, onset of problematic substance use, abstinence and relapse) and that includes other risk factors relevant for substance use disorder will be important for further addressing the relative contribution of these mechanisms as a cause or consequence of drug-taking behavior.

The limitations of the current design and analysis provide clear directions for future work. First, the use of mTurk means that biological verification of alcohol use was not possible and could have resulted in disingenuous behavior regarding alcohol use histories. Prior research has demonstrated the reliability and validity of self-report for

alcohol and other substance use behaviors (Elman et al. 2000; Kokkevi et al. 1997; Napper et al. 2010). Research on mTurk has also found that participants may be more comfortable reporting sensitive material online than in person (Kim and Hodgins 2017; Strickland and Stoops 2018b). This reported comfort is consistent with other work suggesting that online data collection can help reduce reporting biases related to stigmatized behaviors (Harrison and Hughes 1997; Turner et al. 1998). Second, although the use of crowdsourcing helped to improve the heterogeneity of the sampled demographic and health characteristics, deviations from a truly nationally representative sample did exist. In general, mTurk samples have tended to be younger and more education and less likely to be fully employed and a racial minority than those generated in nationally representative studies (Chandler and Shapiro 2016). Deviations from these expectations that were observed in the current study could be attributed to the focus on individuals with weekly or greater alcohol consumption. It is less likely that departures were systematically related to the approval rating restrictions given that previous studies have shown that individuals differing in qualification restriction do not differ with respect to demographic and substance use characteristics (Peer et al. 2014; Strickland and Stoops 2018a). Regarding the validity of this convenience sampling approach, prior research has found that when compared to other forms of convenience sampling, mTurk samples can provide similar or sometimes improved representation of the United States population (Berinsky et al. 2012; Huff and Tingley 2015) and does not result in greater rates of problematic responding (e.g., social desirability bias) (Necka et al. 2016). Although the limitations of convenience sampling should be considered, it is likely that a combination of sampling approaches from laboratory, clinic, and online settings that balance the strengths and weaknesses of these respective approaches will serve to enhance the rigor and scope of alcohol and other substance use research. Future work

evaluating these relationships within the laboratory and clinical setting will nevertheless provide collateral support for the relationships described here.

Third, as noted above, commodity-specific discounting and cross-commodity demand were not collected. Similarly, information on other substance use was not collected during the longitudinal period precluding statements about the predictive relationship of behavioral economic measures indexing valuation for other drugs of abuse. Fourth, the maximum drink price included on the alcohol purchase task was relatively low compared to other studies utilizing the purchase task procedure. This low maximum price resulted in 29.6% of participants reporting some level of consumption at the final unit price. Future studies would benefit from including a higher price range to ensure that the maximum range of purchasing behavior is observed (for more information on purchase task design see review by Kaplan et al. 2018). Relatedly, the proportionate alcohol use measure specified between "when you were not drinking alcohol" and "when you were drinking alcohol". Therefore, it is possible that other substance use that occurred outside of the context of alcohol use (e.g., cigarette use, illicit substance use) was captured in the alcohol-free activity assessments. This version was selected to help isolate the specific influence of alcohol-related valuation. However, future work would benefit from parametric manipulations of these instructions to determine the potential influence of the reliability and validity of the measure and its association with alcohol consumption.

Finally, this analysis focused on a subset of the initially enrolled sample due to exclusions for non-systematic or inattentive data. Sensitivity analyses suggested that the exclusion of these participants was not systematically related to demographic profiles or alcohol use characteristics. Of the participants removed, 39% (33 of 84) were also due to low rates of alcohol consumption (i.e., not reporting alcohol use during the study follow up or reporting zero consumption across all values on the alcohol purchase task)

rather than inattentive or irregular responding. This high percentage is likely a partial artifact of the inclusion criterion used (one or more alcohol drinks/week), which could have resulted in individuals with low rates of alcohol consumption. Prior studies have often focused on individuals reporting heavy drinking or other forms of problematic drinking (for review of sample characteristics and drinking criteria in alcohol purchase task studies see Kaplan et al. 2018). Other studies, however, have used community samples that report more widely varying use patterns similar to the sample analyzed here (e.g., Bertholet et al. 2015; Morris et al. 2017). The rationale for our inclusion criteria and target population was to provide a wide variety of patterns to index these behavioral economic relationships across a range of alcohol use patterns and to generalize these findings to a general community sample. Future work could instead target individuals with alcohol use disorder within or outside a treatment context to determine if similar relationships are observed in a problematic alcohol use context.

The present study adds to the growing literature developed at the intersection of behavioral economics and addiction science. Our findings suggest that behavioral economic variables, such as behavioral economic demand and proportionate-related alcohol reinforcement, provide unique, predictive, and incremental validity for future determining variations in alcohol use frequency, quantity, and severity. Future work will be important for generalizing these findings to samples collected using alternative methods as well as other drugs of abuse and participant populations (e.g., those seeking treatment). Nevertheless, such associations provide support for the continued use of behavioral economic measures in the addiction science as valuable measures for the development of prevention and treatment interventions targeting alcohol use.

Table 7.1. Participant Demographics and Behavioral Economic Variables (N = 223)

	Mean/%	SD	IQR
Demographics			
Age	35.2	10.5	27 to 41
Male	47.1%		
White	83.0%		
Unemployed	14.8%		
College	68.6%		
Income (in thousands)	$49.6	$30.5	$20.0 to $70.0
Alcohol Use			
Drinks/Weeks	8.8	9.3	3 to 12
Days/Week	3.0	2.0	2 to 4
Drinks/Occasion	3.1	2.0	2 to 4
DSM-IV Abuse	9.0%		
DSM-IV Dependence	31.8%		
AUDIT	10.3	7.5	4 to 14
Hazardous Drinking (8+)	52.9%		
Soda Use (n = 166)[a]			
Drinks/Weeks	7.2	9.4	1 to 10
Days/Week	3.6	2.5	1 to 6
Drinks/Occasion	1.9	1.4	1 to 2
BE Variables			
Delay Discounting (k) [log]	-2.34	0.78	-2.71 to -1.87
Alcohol Q_0 [log]	0.76	0.36	0.52 to 0.95
Alcohol α [log]	-1.85	0.61	-2.26 to -1.55
R-ratio [log]	0.36	0.16	0.25 to 0.49
Soda Q^a [log]	0.62	0.40	0.36 to 0.81
Soda α^a [log]	-1.33	0.53	-1.68 to -1.06

Note. BE = behavioral economic; AUDIT = Alcohol Use Disorder Identification Test; *k* = discounting rates; Q_0 = demand intensity; α = demand elasticity; IQR = interquartile range.

[a]Only subjects reporting soda use and providing systematic soda purchase task data (n = 166)

Table 7.2. Baseline Correlations with Behavioral Economic Variables (N = 223)

		Alcohol			Soda (n = 166)	
	k	Q_0	A	R-ratio	Q_0	α
Demographics						
Age	-.04	-.17*	.06	-.04	-.07	.06
Male	-.14*	.17*	-.07	.06	-.03	-.06
White	-.12	-.11	.01	-.07	-.04	-.12
Unemployed	.01	-.01	-.06	-.02	.01	-.08
College	-.13	-.17*	.10	.03	-.04	.12
Income	-.15*	-.03	-.03	-.05	-.16*	.02
Alcohol						
Drinks/Weeks	.10	.42***	-.34***	.25***	.12	-.11
Days/Week	.12	.14*	-.18**	.30***	.06	-.01
Drinks/Occasion	.10	.46***	-.32***	.09	.17*	-.20**
DSM-IV	.14*	.31***	-.20**	.25***	.17*	-.07
AUDIT	.15*	.49***	-.32***	.32***	.23**	-.16*
Soda (n = 166)						
Drinks/Weeks	.04	.04	-.10	-.01	.30***	-.29***
Days/Week	.00	.10	-.09	-.12	.33***	-.37***
Drinks/Occasion	.16*	.10	-.14	.03	.35***	-.35***
BE Variables						
k	-	-	-	-	-	-
Alcohol Q_0	.19**	-	-	-	-	-
Alcohol α	-.12	-.57***	-	-	-	-
R-ratio	.14*	.19**	-.11	-	-	-
Soda Q^a	.18*	.58***	-.25**	.04	-	-
Soda α^a	-.21**	-.41***	.60***	.03	-.58***	-

Note. DSM-IV = meets criteria for alcohol abuse or dependence; AUDIT = Alcohol Use Disorder Identification Test; BE = behavioral economic; *k* = discounting rates; Q_0 = demand intensity; α = demand elasticity. Correlations involving soda demand or soda use variables included 166 participants.
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 7.3. Effect Size Estimates for Generalized Mixed Effect Models

		Alcohol			Soda (n = 166)	
	k	Q_0	α	R-ratio	Q_0	α
Unadjusted						
Drinking Day	1.12	2.24*	0.62*	27.03***	1.07	0.79
Drinks/Drinking Day	1.09	2.46***	0.73***	1.87*	1.32*	0.81**
Heavy Drinking Day	1.60*	41.50***	0.26***	7.57	2.63*	0.48*
Adjusted for AUDIT						
Drinking Day	0.96	0.58	0.98	5.54*	0.58	1.06
Drinks/Drinking Day	1.02	1.63***	0.88*	0.87	1.07	0.90
Heavy Drinking Day	1.22	7.58***	0.56*	0.32	1.13	0.73
Full Model						
Drinking Day	0.96	0.45	0.79	6.02*	0.46	0.69
Drinks/Drinking Day	0.99	1.65***	1.01	0.85	0.76*	0.88
Heavy Drinking Day	1.12	7.15***	0.97	0.26	0.29*	0.71

Note. AUDIT = Alcohol Use Disorder Identification Test; *k* = discounting rates; Q_0 = Demand Intensity; α = Demand Elasticity. Estimates involving soda demand or soda use variables included 166 participants. All values represent effect size estimates (odds ratios [OR] for presence or absence of drinking day and heavy drinking day and rate ratios [RR] for drinks/drinking day).
* $p < .05$; ** $p < .01$; *** $p < .001$

Table 7.4. Test-Retest Reliabilities and Effect Size for Change in Behavioral Economic Measures

	Pre	Post	*N*	r_{xx}	d_z
k^a	-2.377	-2.512	150	.76	0.24**
Alcohol $Q_0{}^a$	0.744	0.687	135	.69	0.21*
Alcohol α^a	-1.848	-1.789	135	.50	0.10
R-ratio	0.353	0.346	150	.29	0.03
Soda Q^a	0.571	0.521	99	.70	0.20
Soda α^a	-1.334	-1.251	99	.42	0.14

Note. k = discounting rates; Q_0 = Demand Intensity; α = Demand Elasticity; r_{xx} = test-retest reliability correlation; d_z = Cohen's d effect size for repeated designs (Lakens 2013). Sample sizes reflect the number of participants providing follow up data and providing systematic data at pre- and post-test follow

[a]Pre and post-study values reflect log-transformed values only for subjects reporting values at both timepoints.

* $p < .05$; ** $p < .01$ comparing pre and post values

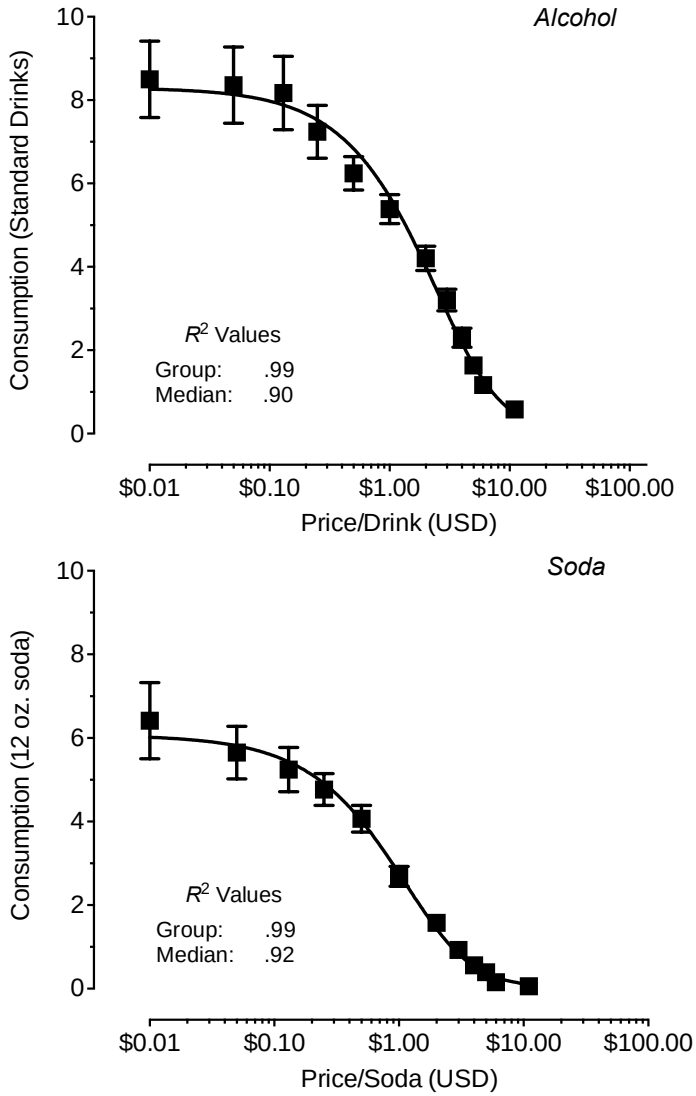

Figure 7.1. Behavioral economic demand for alcohol (top) and soda (bottom). Participants completed commodity purchase tasks in which hypothetical alcohol (quantified as one US standard drink) or soda (quantified as one 12 oz. serving) were available. Price varied in United States dollars (USD). Plotted are mean (SEM) group data fit using the exponentiated model. Group R^2 refers to the model fit for the plotted data (i.e., mean data values for each commodity. Median R^2 refers to the median value for individually fit demand curves.

Chapter 8

FEASIBILITY, ACCEPTABILITY, AND INITIAL EFFICACY OF DELIVERING ALCOHOL USE COGNITIVE INTERVENTIONS VIA CROWDSOURCING

(Experiment 5a; Strickland et al., in press)

Introduction

Alcohol use disorder (AUD) presents a persistent public health concern. Globally, 5.3% of global mortality is attributable to alcohol consumption (World Health Organization 2018). In the United States, 14.0 million Americans met criteria for AUD in 2017 with recent estimates indicating an annual economic impact of excessive drinking of $250 billion (Center for Behavioral Health Statistics 2018; Sacks et al. 2015). Alcohol use also results in numerous health and social impacts for the individual, including a substantive contribution to intimate partner violence and domestic abuse, exacerbation of existing physical and mental health conditions, and interactions with other substances (e.g., cocaine, cigarettes) that can further increase health risk (Cargiulo 2007; Durazzo et al. 2004; Farre et al. 1997; Foran and O'Leary 2008; Rehm 2011). Despite the identification of approved interventions for AUD, treatment gaps still exist and many of those approaches that do exist are still not widely utilized or universally effective (Litten et al. 2016; Kufahl et al. 2014). These evident consequences of AUD combined with the lack of widely effective or utilized treatment modalities highlight the importance of novel approaches for intervention development to address AUD.

Cognitive training has received a great deal of attention in interventions development for AUD. Training may be broadly divided into two categories: 1) cognitive bias modification and response inhibition training (i.e., inhibitory control) and 2) working memory interventions (Verdejo-Garcia 2016). Inhibitory control training attempts to retrain prepotent responses away from drug-related cues by specifically pairing those cues with no-go signals in training tasks (e.g., Houben et al. 2011a; Houben et al. 2012).

Working memory training uses cognitive-behavioral tasks (e.g., letter/digit strings, visual search, N-Back) to improve information maintenance and manipulation (e.g., Bickel et al. 2011c; Houben et al. 2011b). These interventions hold particular appeal because they present few contraindications and may be easily incorporated into and potentially enhance comprehensive treatment approaches. For example, it is possible that the cognitive improvements owing to training could improve engagement with and attention to cognitive-behavioral therapy and compliance with homework and other program-related activities.

Existing studies on cognitive training targeting substance use behaviors have provided mixed results. Training has consistently resulted in improvements in performance on the trained or closely related tasks (i.e., near-transfer) (e.g., Houben et al. 2011a; Houben et al. 2011b; Snider et al. 2018). In contrast, improvements on dissimilar tasks or those within different cognitive domains are generally not observed (i.e., far-transfer) (Snider et al. 2018). Several studies have reported reductions in weekly alcohol consumption and/or laboratory alcohol consumption following inhibitory control training (Houben et al. 2011a; Houben et al. 2012; Jones and Field 2013) or working memory training (Houben et al. 2011b). However, negative or mixed outcomes have also been reported (Bowley et al. 2013; Smith et al. 2017; Wanmaker et al. 2018). Although fewer studies have addressed other substance use disorders, the results of existing research are similar with consistent improvements in near-transfer performance (Alcorn et al. 2017; Bickel et al. 2011c), but more varied with respect to changes in far-transfer performance or substance use (e.g., Adams et al. 2017; Rass et al. 2015b; Schulte et al. 2018; for general meta-analysis on working memory effects see Melby-Lervag and Hulme 2013).

Important gaps in the cognitive training literature may explain the mixed results observed. One significant limitation is the relatively small and selective samples typically

evaluated (e.g., 20-40 participants per training condition; college student samples). It is likely that the effects of cognitive training are of a small-to-moderate effect size and the use of small samples could preclude the detection of significant effects due to low power or the inflation of observed effect size estimates when significant effects are detected (see discussion of these problems associated with low statistical power in Button et al. 2013). It is also likely that individual characteristics will moderate the utility of these interventions for impacting substance use or related health behaviors. The exact parameters that result in the greatest improvements in cognitive performance or substance use behaviors are also unknown due to the lack of parametric studies evaluating how the depth or breadth of training influences treatment effects.

An emerging method positioned to address these gaps is crowdsourcing. Crowdsourcing, such as on Amazon Mechanical Turk (mTurk), allows for the effective and efficient recruitment of research participants by using online sampling of large and varied pools of potential participants from across the United States (and, if desired, world). Recent research in psychological science, broadly, and addiction science, specifically, has demonstrated the reliability and validity of conducting research on mTurk (see reviews in Chandler and Shapiro 2016; Strickland and Stoops 2019). These studies, including those conducted in individuals with AUD and other substance use disorders, have demonstrated a close correspondence between findings obtained using in-person samples and those recruited through mTurk thereby providing support for the methodological approach (Johnson et al. 2015; Kim and Hodgins 2017; Strickland et al. 2016a). mTurk has proved a particularly flexible platform for research with varied methodological approaches successfully applied from basic cross-sectional survey designs to measure development to intensive longitudinal research.

The purpose of this study was to test the feasibility and acceptability of delivering cognitive training interventions via mTurk. Participants were randomized to 1) inhibitory

control training (ICT), 2) working memory training (WMT), or 3) control training (Control). Training was completed daily for a two-week period. Follow up assessments were conducted immediately following and two weeks after training to evaluate the impact on alcohol use behaviors as well as the acceptability of the training tasks. Our primary hypothesis was that delivering cognitive training via mTurk would be feasible and acceptable consistent with previous intensive longitudinal research related to alcohol use on the platform (Boynton and Richman 2014; Strickland and Stoops 2018b). We also evaluated the initial efficacy of cognitive training for reducing alcohol consumption. Our secondary hypothesis was that cognitive training would produce small effect size reductions in the proportion of drinking days and heavy drinking days.

Methods

Participants and Screening

Participants were recruited from mTurk. To view the study, participants had to have completed at least 100 mTurk tasks, have a ≥ 95% approval rating on prior tasks, and reside in the United States (see similar qualifications in Cunningham et al. 2017; Strickland and Stoops 2015). A short screening questionnaire was used to determine study eligibility. Inclusion criteria were: 1) self-reported past week alcohol use, 2) 21 years of age or older, 3) interest in a 2-week study on mTurk, and 4) meet criteria for DSM-5 AUD according to a validated brief questionnaire (Hagman 2017). All surveys were hosted on Qualtrics (Provo, UT, USA). Participants received $0.05 for completing the screening survey. The University of Kentucky Medical Institution Review Board reviewed and approved all procedures and the protocol was pre-registered on ClinicalTrials.gov (NCT03438539).

General Procedures

Qualifying participants first completed a baseline survey that contained questions about health and alcohol use history and a timeline follow back assessment (TLFB) (see

Study Measures below for details). Participants were then randomized to receive normative feedback information at the end of this baseline survey. Briefly, half of participants were randomly assigned to receive an alcohol normative feedback delivery in which they were directed to a statement standardized based on reported average number of standard drinks per week, age, and gender with individual percentile rank of weekly alcohol consumption compared to values from the National Survey on Drug Use and Health (Center for Behavioral Health Statistics 2017). The other half of participants received control feedback consisting of information on daily television usage. No impact of normative feedback alone or in combination with the training conditions was observed on the alcohol efficacy outcomes reported here so this will not be discussed further.

Participants were also randomized to one of three training conditions (i.e., Inhibitory Control, Working Memory, or Control) following completion of the baseline survey (see details on training in Training Tasks below). Randomization was stratified based on AUD status (mild, moderate, severe). Tasks were programmed in PsyToolkit, an open-source web-based platform that provides reliable reaction time data for online delivery (Stoet 2017). Each training task was completed daily for a continuous 14-day period. Tasks were designed to take approximately five minutes (actual median times of completion: Inhibitory Control = 6 minutes; Working Memory = 6 minutes; Control = 4 minutes). Participants were paid $0.50 for each training survey.

Follow up surveys were completed immediately following and two weeks after the end of the training phase. These surveys included alcohol use information collected at baseline (e.g., TLFB) to allow for comparisons across the study phases. Participants were paid $2 and $0.75 for these follow ups, respectively. An additional incentive raffle for one of five $50 bonuses was used to encourage active completion. Participants received escalating entries based on the number of training task completed (10 = 1

entry; 11 = 3 entries; 12 = 5 entries; 13 = 10 entries; 14 = 15 entries) as well as 5 additional entries for each follow up survey completed.

Training Tasks

Inhibitory Control Training

The inhibitory control training task was a modified version of Cued Go/No-Go tasks (Miller et al. 1991; Weafer and Fillmore 2012). The task consisted of two blocks of 50 trials. Each trial began with a fixation point presented for 800 ms followed by a blank black screen for 500 ms. A cue image (alcohol or neutral) was then presented for one of five stimulus onset asynchronies (SOA; i.e., 100, 200, 300, 400, 500 ms). Finally, a go or no-go target was displayed until a response occurred or 1000 ms elapsed. The color green was the go target and signaled that a response should be made, whereas the color blue was the no-go target and signaled that a response should be withheld. Feedback on the response and reaction time for go responses was provided for 1250 ms followed by a new trial. Alcohol images served as no-go cues and visually matched neutral images served as go cues. Cues predicted which target was presented 100% of the time (e.g., alcohol images were always followed by no-go targets). The primary outcome was correct responses to the no-go targets as a measure of inhibitory control. Performance on go trials and response time were also recorded.

Working Memory Training

A battery of working memory tasks was used during the intervention period. These tasks were selected from previous research evaluating working memory training in substance use disorders (Bickel et al. 2011c; Houben et al. 2011b). Tasks included a visuospatial working memory task (i.e., Corsi task), digit span task (forward and reverse), letter span task (forward and reverse), and the N-Back. Broadly speaking, these tasks required information recall and/or categorization of that information based on short-term retention. Task difficulty in the recall tasks was adaptive and increased or

decreased based on task performance (i.e., two consecutive correct responses increased the tested span and two incorrect responses decreased the tested span). The maximum recall span across 25 trials was the primary outcome for recall tasks. Task difficulty increased in the N-Back task by transitioning through the 1-Back, 2-Back, and 3-Back within the task. The primary outcome for the N-Back task was error rates for targets and non-targets. Participants completed one task randomly selected during each session.

Control Training Tasks

Control training tasks included completion of 60 arithmetic problems. These problems included simple single digit arithmetic (20 addition, 20 subtraction, and 20 multiplication) that did not increase or decrease in difficulty within or across sessions. This control condition was selected to provide a task that required active engagement, but did not provide adaptive training. Accuracy was not recorded or reported to participants.

Study Measures

Alcohol Use and Soda Use History

Participants completed a battery of standardized assessments of alcohol use during the baseline survey, including the Alcohol Use Disorder Identification Test (AUDIT) (Saunders et al. 1993) and Short Index of Problems (SIP) (Kiluk et al. 2013). Other quantity-frequency measures were also included to evaluate recent alcohol and soda consumption patterns. All alcohol units referred to standard US drinks and all soda units referred to 12 oz. servings.

Timeline Followback (TLFB)

The TLFB was used at baseline and follow up surveys to assess alcohol and soda consumption for the pre-training, training, and post-training phases. Participants were provided a calendar of two-week periods included in those phases and asked to report

the number of standard drinks consumed and number of sodas consumed on each day. Drinks were reported by type (e.g., beer, wine, liquor), but were totaled to standard drinks for the purpose of analysis. Participants were explicitly instructed that soda did not include soda mixed with alcoholic drinks. Previous research has demonstrated the reliability of the TLFB when delivered using computerized methods (Sobell et al. 1996).

Acceptability Measures

The primary acceptability measure was responses on a version of the Treatment Acceptability Questionnaire (TAQ) (Raiff et al. 2013). Participants rated six statements about the daily training tasks (i.e., ease of completion, helpful instructions, enjoyability, convenient timing, fair compensation, and overall satisfaction) on a 100-point visual analog scale (0 = Low; 100 = High). Secondary acceptability measures evaluated overall satisfaction, future participation, participation motives, and experience with mTurk. Participants were explicitly instructed that their choices would not affect future payments and to respond honestly. Acceptability measures were included in both follow-up surveys to maximize completion rates, but participants were only asked to complete acceptability measures once if they completed both follow-ups.

Data Analysis

Figure 8.1 presents a CONSORT diagram describing study enrollment and the sample sizes for feasibility, acceptability, and efficacy analyses. Four hundred and seventy-six participants qualified and completed the baseline survey. Thirty-two of these participants failed one or more attention or data quality checks throughout the study and were removed from analysis. This systematic sample of 444 was used for analysis of adherence data in order avoid artificial inflation of feasibility assessments. However, acceptability and efficacy analyses focused on individuals who completed at least one training session (N = 402).

Demographic and substance use history was first evaluated using descriptive statistics for the total sample and by intervention group. Baseline group differences were compared using a one-way analysis of variance (ANOVA) with group as the between-subject factor.

Feasibility was evaluated in two ways. First, adherence rates over the 14-day intervention period were determined and compared between groups using a one-way ANOVA. Bivariate correlations were also computed between adherence and demographic and alcohol use variables. Second, intervention fidelity was assessed by evaluating task performance over the 14-day intervention period. The effect of intervention day was determined using general linear mixed models that accounted for the within-subject design and continuous predictor variable (i.e., day). Additional mixed models tested expected relationships within tasks types by parameterizing features of the task design (e.g., decreased performance with increasing N-Back span).

Acceptability was first evaluated using descriptive statistics of the primary and secondary acceptability measures. Median values are presented for the TAQ given the observation of a non-normal distribution and non-parametric tests used to compare intervention groups (Kruskal-Wallis test with Mann-Whitney U test post-hoc). Categorical secondary acceptability measures were compared using chi-square tests with group as the independent variable.

Two primary endpoints were selected for tests of initial efficacy: 1) proportion of drinking days and 2) proportion of heavy drinking days over each 2-week period. Heavy drinking was defined using National Institute on Alcohol Abuse and Alcoholism (NIAAA) guidelines of ≥5/≥4 drinks/day for men/women (U.S. Department of Health and Human Services, 2016). These endpoints were selected to provide simple continuous measures for model fitting that were also clinically meaningful. General linear mixed models tested the full factorial effect of Group (ICT, WMT, Control) x Phase (Baseline, Training, Follow

Up). The control group and baseline phase served as reference categories in each model. Secondary moderation analyses were also conducted evaluating demographic (age and sex) and alcohol use (AUDIT scores and AUD severity) variables as putative moderators of treatment efficacy (i.e., Group x Phase x Moderator interactions). An additional selectivity endpoint of proportion of soda consumption days was also tested. This model focused only on individuals who reported soda consumption at baseline (N = 348). However, similar magnitude and significance effects were observed when using the total sample rather than a soda consumption sub-sample (data not shown). Analyses described below suggested that adherence was not systematically or meaningfully related to alcohol use behavior. Accordingly, data were treated as missing at random and maximum likelihood used (Singer and Willett 2003).

SPSS Statistics (IBM; Armonk, NY) and R statistical language with the *nlme* package (Pinheiro et al. 2018) were used for analyses. All inferential tests were two tailed and used an alpha rate of .05.

Results

Sample Characteristics

Table 8.1 contains demographics and substance use behaviors collected at baseline for the total sample and separated by intervention group. Overall, participants were an average of 34.3 years old with an approximately equal distribution of men and women. The majority of participants were white, employed, and had a college education. Participants endorsed an average of 5.5 *DSM-5* AUD diagnostic criteria. No baseline differences were observed as a function of intervention group.

Feasibility

Adherence

Daily response rates by intervention group are plotted in Figure 8.2. The highest overall response rate was observed on Day 1 (74.5%) and the lowest response rate

observed on Day 13 (56.1%). Approximately 90.5% of participants completed at least one training session with an average of 9.1 sessions (65.0%) completed in the total sample. One-half (47.3%) of participants completed 80% or more daily sessions.

Significant, but small effect size, relationships were observed between days adherent and lower SIP scores, r = -.11, and fewer sodas per occasion, r = -.18 as well as lower endorsement of daily cigarette smoking, r = -.14, and past week cannabis use, r = -.14, p values < .05. No significant relationships were observed with other alcohol or soda use behaviors, r values < .07. Similarly, individuals with 80% or greater adherence reported fewer sodas per occasion, r = -.11, were less likely to report past week cannabis use, r = -.10, and tended to be older, r = .11, and male, r = .11, p values < .05. No significant relationships were observed with other alcohol or soda use behaviors, r values < .09.

The average number of sessions completed was significantly different by treatment condition, $F_{2,443}$ = 3.94, p = .02. This effect represented lower adherence in the WMT (8.1 sessions [SD = 5.3]) compared to ICT (9.7 sessions [SD = 4.8]) and Control (9.4 sessions [SD = 4.9]) groups, p values < .027.

Intervention Fidelity

No-go response accuracy (i.e., the inverse of inhibitory failures) increased over the test period, effect of Day β = 0.001, p < .001. This effect indicated a small, but statistically significant improvement in response accuracy over the intervention period (e.g., Day 1 estimate = 95.9% to Day 14 estimate = 97.3%; Figure 8.3 top panel). No significant change in Go response accuracy (i.e., the inverse of commission errors) was observed, effect of Day β = < 0.001, p = .12. Average reaction times on Go trials decreased by approximately 3.4 ms per day, β = -3.40, p < .001.

A significant effect of Day was observed on recall performance collapsing across the recall tasks, β = 0.05, p < .001. This effect represented an approximately 0.7 unit increase in maximum recall span over the 14-day period. Model estimated performance

on each task is plotted in Figure 8.3 (middle panel). Varied parameterizations of model effects indicated decreased performance on backward compared to forward recall, letter compared to number recall, and spatial compared to alpha/numeric recall, p values < .05. Approximately 2.7% of responses included a recall span greater than 10 (maximum possible = 15).

A significant effect of Day was observed collapsing across response type and N-Back span, $\beta = -1.13$, $p < .001$. This effect represented an approximate decrease of 15.8% in overall failure rate over the 14-day period. Significant interactions between response type (target versus non-target) and N-Back span (one, two, three) were also observed for failure rate, p values < .05. As indicated in estimates plotted in Figure 8.3 (bottom panel), these effects represented a general increase in Target response failures and a general decrease in Non-Target response failures with increases in N-Back span.

Acceptability

Median ratings on the TAQ by treatment condition are presented in Figure 8.4. Significant group effects were observed for Ease of Completion, $\chi^2 = 45.87$, $p < .001$, and Overall Experience, $\chi^2 = 7.01$, $p = .03$. Post-hoc comparisons indicated that the WMT group reported lower ratings of Ease of Completion than the ICT or Control groups and lower ratings of Overall Experience than the Control group, p values < .009.

Table 8.2 contains responses for secondary acceptability measures. A majority of participants indicated that they satisfied with the study procedures, would participate again, and would consider incorporating the training task in their daily life. The most common motive for participation was monetary compensation (83.9%) followed by completing interesting tasks (59.3%). A majority of participants also indicated that they like participating in mTurk research as much or more than in-person studies (73.7%) and that it is easier to answer sensitive questions honestly on mTurk compared to an

interview (75.1%). No significant group differences were observed for secondary acceptability measures or motives, *p* values > .09.

Initial Efficacy

Coefficient estimates from linear mixed effect models predicting proportion drinking days, heavy drinking days, and soda consumption days are presented in Table 8.3. A significant effect of Training Phase, β = 0.042, *p* = .033, and significant Follow Up Phase x ICT interaction, β = -0.064, *p* = .026, were observed for proportion drinking days. Planned follow-up tests indicated a significant reduction in proportion drinking days in the ICT group, β = -0.068, *p* < .001, a trend towards reduction in the WMT group, β = -0.035, *p* = .097, and no reduction in the Control group, β = -0.004, *p* = .85, during the two-week follow-up period compared to pre-training period (Figure 8.5 top panel).

A significant effect of Follow Up Phase, β = -0.038, *p* = .024, was observed for proportion heavy drinking days. Planned follow-up tests indicated a significant reduction in proportion heavy drinking days in the ICT group, β = -0.050, *p* = .011, and Control group, β = -0.039, *p* = .008, but no reduction in the WMT group, β = -0.017, *p* = .28, during the two-week follow-up period compared to pre-training phase (Figure 8.5 bottom panel).

No Group or Phase effects were observed for proportion soda drinking days. Moderation analyses also did not reveal significant interactions including age, sex, AUDIT scores, or AUD severity for either outcome variable, *p* values > .05.

Discussion

The purpose of this study was to evaluate the feasibility and acceptability of delivering cognitive interventions through the crowdsourcing website mTurk. Additional tests of efficacy evaluating reductions in alcohol consumption and selectivity tests evaluating soda consumption were also conducted. Response rates were satisfactory over the two-week intervention period and performance on training tasks was consistent

with existing laboratory and clinical research (Houben et al. 2011a; Houben et al. 2011b; Snider et al. 2018). Participants also indicated that the intervention delivery was acceptable and that they would participate in similar research again. Modest reductions in alcohol consumption were observed, primarily in the ICT group, and these effects were selective and did not extend to soda consumption. Taken together, these findings demonstrate the feasibility and acceptability of utilizing crowdsourcing methods for interventions development and support this sampling method for future work in AUD and other substance use disorders.

The feasibility of online delivery was demonstrated in two distinct ways. First, approximately two-thirds of the possible sessions were completed by participants with one-half of participants completing 80% or more of the daily sessions (i.e., a standard cutoff of clinical adherence) (Brown and Bussell 2011). These response rates could be considered satisfactory given the low intensity of contact with participants (i.e., once daily email reminders) and relatively lean compensation schedule used. No studies, to our knowledge, have evaluated the effects of repeated inhibitory control training in a population with substance use disorders making comparisons to existing literature difficult. Our response rates were similar, however, to completion rates in another study of working memory in individuals with AUD in which 39% of participants were non-completers following study randomization for reasons such as lost contact or too much time between sessions (Snider et al. 2018; but see Houben et al. 2011b, for higher response rates). Lower response rates were observed in the WMT group compared to the ICT or Control groups. This difference could be explained by the increased difficulty and burden associated with the working memory task (see further discussion of acceptability below). This finding suggests that additional incentives may be needed to enhance compliance for WMT delivery.

Feasibility was also demonstrated by examining the fidelity of intervention delivery. Performance on the inhibitory control, recall, and N-Back tasks improved over the training period consistent with the near-transfer effects widely reported in previous studies (Houben et al. 2011a; Houben et al. 2011b; Snider et al. 2018). For example, the 0.7 increase in recall span we observed is similar to increases reported in previous working memory research (e.g., ~1.2 in working memory span over 14 sessions in Houben et al. 2011b; ~1.0 increase in backward recall span over 25 sessions in Rass et al. 2015b). More broadly, canonical effects expected from the cognitive psychology literature (e.g., poorer performance on backward than forward recall span Dempster 1981) were replicated thereby providing further support for the feasibility and validity of online delivery. In only 2.7% of sessions did participants have a recall span of greater than 10 and in only one session was perfect performance observed indicating that participants were not inappropriately writing down or otherwise recording the requested recall span. This result is particularly important given that a common critique of online research is the loss of control over the testing environment that is argued to promote disingenuous or dishonest behavior. That participants did engage in such behavior when given a clear opportunity to do so further supports the validity of data collected through online crowdsourcing methods.

Participants reported a positive experience in the study with a clear majority indicating that they were satisfied with the study procedures and would participate again. Ratings on the TAQ were also high indicating that participants found the intervention easy to complete, enjoyable, and convenient and that they were adequately compensated for their time. Lower ratings for ease of completion and overall experience were observed in the WMT group potentially accounting for some of the decrements in adherence observed. No differences were observed in ratings of satisfaction or likelihood of future participation between groups, however, which suggests that the

differences observed on the TAQ did not uniformly impact study acceptability. It is important to note that acceptability measures were collected at the end of training, which meant that acceptability data were not available for a subset of participants that completed any training tasks (~10%). Nevertheless, the consistent positive response across the measures that is in accordance with acceptability measures in a prior longitudinal study on mTurk (Strickland and Stoops 2018) supports the acceptability of this delivery method.

Two endpoints were selected to assess the initial efficacy of cognitive training for reducing alcohol consumption. Significant decreases in the proportion of drinking days and heavy drinking days were observed in the ICT group consistent with previous studies showing decreases in drinking in the laboratory (Bowley et al. 2013; Di Lemma and Field 2017) and naturalistic setting (Houben et al. 2011a; Houben et al. 2012) (but see null findings in Smith et al. 2017). Only a modest decrease in alcohol use that was not statistically significant was observed in the WMT group. These finding is partially consistent with previous studies, which shown mixed results for the effects of working memory training on alcohol or other substance use (Houben et al. 2011b; Rass et al. 2015b; Schulte et al. 2018; Wanmaker et al. 2018). The modest reductions observed in each group support the study of combining ICT and WMT approach similar to the effective use of combination pharmacotherapies and/or behavioral therapies for substance use disorders (Stead and Lancaster 2012; Stoops and Rush 2014). A combination of ICT and WMT could address independent neurobehavioral mechanisms thought to contribute to maladaptive patterns of substance use. Specifically, IC training could help to regulate an otherwise overactive bottom-up impulsive system (i.e., an overactive “hot” system) whereas WMT could help to engage poor top-down executive control (i.e., an underactive “cold” system) (Goldstein and Volkow 2011; McClure and Bickel 2014).

Decreases in heavy drinking were observed in the control group, which could be due to reactivity effects owing to recording and self-monitoring of alcohol consumption over this extended period (Collins et al. 1998; Fremouw and Brown 1980; but see Litt et al. 1998). It is also possible that the reductions observed in the control condition were due to the cognitively challenging nature of the arithmetic task, albeit without an adaptive difficulty. Also important to note is that these observed effects were selective to alcohol consumption and did not extend to soda consumption. This finding is particularly relevant for the IC group given that this treatment approach specifically targets response inhibition to alcohol cues. Taken together, these findings combined with prior literature provide tentative support for the efficacy of cognitive training programs for producing small effect size reductions in alcohol consumption.

This study represents the first to our knowledge to evaluate the delivery of an intensive, daily intervention through the mTurk platform. Previous studies have demonstrated the feasibility and acceptability of delivering brief interventions related to substance use through mTurk, including personalized alcohol feedback (Cunningham et al. 2017; Kuerbis et al. 2016; Kuerbis et al. 2017), episodic future thinking training (Stein et al. 2017; Stein et al. 2018b), and knowledge-based education (Huhn et al. 2018; Wen et al. 2016). These studies have evaluated the impact of brief intervention delivery immediately or in a single follow up assessment. The current study extends this literature by demonstrating the feasibility and acceptability of conducting repeated delivery and assessments of putative interventions via crowdsourcing. Several benefits of this approach are clear. The effective and efficient recruitment that crowdsourcing allows provides for an ideal platform to test varied parametric manipulations that are otherwise overlooked in the laboratory or clinic. For example, this study only evaluated short daily sessions of training over a two-week period and it is possible that training that was more intense and/or over a longer duration could produce more robust changes in alcohol

consumption. The large samples possible also allow for the testing of putatively small effect size interventions, such as cognitive training, that could have a significant impact because of their ease of implementation and low cost.

These strengths should be considered in the context of the limitations of mTurk. mTurk remains a form of convenience sampling and will therefore depart from a nationally representative sample with respect to demographic and substance use characteristics. Existing studies have found that mTurk samples tend to report higher rates of substance use as well as tend to be younger, more educated, and less employed compared to national samples (see reviews in (see reviews in Chandler and Shapiro 2016; Strickland and Stoops 2019). However, this research has also demonstrated that when compared to other forms of convenience sampling (e.g., college student samples or samples from college towns), mTurk samples can provide similar or sometimes improved representation of the US population (Berinsky et al. 2012; Huff and Tingley 2015). Concerns about the attention and honesty of mTurk participants are also a possible limitation given the lack of control over the testing environment and inability to biologically verify substance use. As noted above, we provided evidence that participants were honestly attending to the training task despite the open opportunity to engage in dishonest behavior (e.g., recording digit or numeric strings in the recall tasks). A majority of participants also reported that they felt more comfortable reporting sensitive material, such as substance use, over an online platform than in person. Similar results have been reported in previous mTurk research in populations reporting licit and illicit substance use (Kim and Hodgins 2017; Strickland and Stoops 2018b). More broadly, this comfort in reporting is consistent with other studies demonstrating reductions in underreporting biases that may occur with stigmatized behaviors when online data collection is used (Harrison and Hughes 1997; Turner et al. 1998). These limitations outstanding, it is ultimately likely that a combination of sampling approaches from

laboratory, clinic, and online setting that balance the strengths and weaknesses of respective approaches will serve to enhance the rigor and scope of substance use research.

This study focused on alcohol use given the ease of collection and clinical acceptance of alcohol use self-report as a primary outcome (Sobell et al. 2003). However, it is likely that the methods described here would extend to other populations given the effective recruitment of participants reporting varying illicit substance use histories through mTurk (e.g., Huhn et al. 2018; Peters et al. 2017; Strickland et al. 2016a). Taken together, this study provides comprehensive support for the delivery of cognitive interventions via crowdsourcing. This feasibility and acceptability helps to establish a setting for future large sample studies testing novel interventions and/or individual characteristic moderating intervention efficacy related to AUD and other substance use disorders.

Table 8.1. Participant Demographics and Substance Use by Intervention Condition

	Total (N = 444) Mean (SD)/%	ICT (n = 145) Mean (SD)/%	WMT (n = 150) Mean (SD)/%	Control (n = 149) Mean (SD)/%
Age	34.3 (9.7)	33.4 (10.1)	34.4 (9.0)	34.9 (9.9)
Male	48.9%	47.6%	52.7%	46.3%
White	79.1%	84.1%	76.7%	76.5%
College	64.6%	62.8%	68.0%	63.1%
Unemployed	6.1%	5.5%	6.0%	6.7%
Alcohol Use				
Drinks/Week	11.3 (14.9)	12.7 (16.1)	9.8 (11.2)	11.4 (16.9)
Drinks/Occasion	4.2 (3.6)	4.4 (3.7)	3.9 (2.5)	4.2 (4.5)
AUDIT	12.7 (7.3)	12.8 (7.6)	12.7 (7.2)	12.6 (7.0)
SIP	10.5 (10.0)	10.6 (10.1)	10.4 (9.9)	10.7 (10.1)
AUD Symptom Count	5.5 (2.9)	5.5 (2.9)	5.4 (2.9)	5.5 (3.0)
Mild	34.0%	33.1%	34.7%	34.2%
Moderate	21.2%	21.4%	21.3%	20.8%
Severe	44.8%	45.5%	44.0%	45.0%
Soda Use				
Soda Use	82.4%	80.7%	84.7%	81.9%
Sodas/Week	7.6 (9.4)	7.7 (9.3)	6.5 (7.1)	8.7 (11.3)
Sodas/Occasion	2.2 (2.3)	2.1 (1.7)	2.3 (2.7)	2.3 (2.3)
Other Drug Use				
Daily Cigarette Use	37.6%	41.4%	32.7%	38.9%
Cigarettes/Day	11.9 (7.5)	11.5 (8.7)	11.8 (7.1)	12.5 (6.5)
Past Week Cannabis	32.0%	32.4%	36.7%	26.9%

Note. ICT = Inhibitory Control Training; WMT = Working Memory Training; AUDIT = Alcohol Use Disorder Identification Test; SIP = Short Inventory of Problems [Alcohol]; AUD = DSM-5 Alcohol Use Disorder. Alcohol use referred to US standard drinks. Soda use referred to 12 oz. serving. No significant baseline group differences were observed.

Table 8.2. Study Acceptability Measures

Question	Total (N = 354)	ICT (n = 121)	WMT (n = 115)	Control (n = 118)
Overall, How Satisfied Were You with the Study Experience?				
Quite Satisfied	62.3%	64.5%	54.4%	67.8%
Mildly Satisfied	32.3%	30.6%	37.7%	28.8%
Mildly Dissatisfied	3.7%	3.3%	7.0%	0.8%
Quite Dissatisfied	1.7%	1.7%	0.8%	2.5%
Would you Participate Again?				
Definitely So	77.1%	78.5%	73.0%	79.7%
Probably So	20.3%	18.2%	24.3%	18.6%
Probably Not	2.0%	2.5%	2.6%	0.8%
Definitely Not	0.6%	0.8%	0.0%	0.8%
Incorporate in Your Daily Life?				
Definitely So	46.9%	47.1%	40.9%	52.5%
Probably So	34.2%	34.7%	39.1%	28.8%
Probably Not	15.5%	14.0%	15.7%	16.9%
Definitely Not	3.4%	4.1%	4.3%	1.7%
Motivations for Participating[a]				
To Gain Self-Knowledge	42.1%	41.3%	40.0%	44.9%
To Kill Time	16.1%	16.5%	17.4%	14.4%
Enjoy Doing Interesting Tasks	59.3%	61.2%	63.5%	53.4%
To Make Money	83.9%	81.8%	85.2%	84.7%
To Have Fun	24.3%	22.3%	31.3%	19.5%
Experiences with mTurk Research				
I find it easier to answer honestly sensitive questions on mTurk compared to an interview	75.1%	72.7%	76.5%	76.3%
I like the idea of participating in research on mTurk as much or more than participating in research in person	73.7%	68.6%	76.5%	76.3%
I would never participate in a research study in person, but would on mTurk	20.6%	15.7%	23.5%	22.9%

Note. No significant differences were observed between groups using a chi-square test.
[a]Participants could select more than one motivation so endorsements do not total to 100%

Table 8.3. Generalized Linear Mixed Effect Models Predicting Alcohol and Soda Consumption

Variables	Drinking Days	Heavy Drinking Days	Soda Days
Fixed Effects			
Training Phase	0.04 (0.00, 0.08)*	0.00 (-0.03, 0.03)	0.01 (-0.05, 0.06)
Follow Up Phase	0.00 (-0.04, 0.04)	-0.04 (-0.07, -0.01)*	-0.03 (-0.09, 0.03)
ICT	0.01 (-0.07, 0.09)	0.02 (-0.05, 0.09)	-0.05 (-0.15, 0.05)
WMT	-0.03 (-0.11, 0.05)	-0.01 (-0.08, 0.06)	-0.05 (-0.16, 0.05)
Training x ICT	-0.03 (-0.09, 0.02)	0.01 (-0.03, 0.05)	0.00 (-0.07, 0.08)
Follow Up x ICT	-0.06 (-0.12, -0.01)*	-0.01 (-0.06, 0.03)	-0.03 (-0.11, 0.05)
Training x WMT	-0.01 (-0.07, 0.04)	0.01 (-0.03, 0.06)	-0.02 (-0.10, 0.05)
Follow Up x WMT	-0.03 (-0.09, 0.03)	0.02 (-0.03. 0.07)	0.04 (-0.04, 0.12)
(Intercept)	0.56 (0.51, 0.62)***	0.22 (0.17, 0.27)***	0.57 (0.50, 0.64)***
Participants	402	402	348
Observations	1053	1053	849

Note. Values represent parameter estimates and 95% confidence intervals. All outcomes were proportion days over the two-week period. Reference categories were baseline for phase and control condition for group.
* $p < .05$; *** $p < .001$

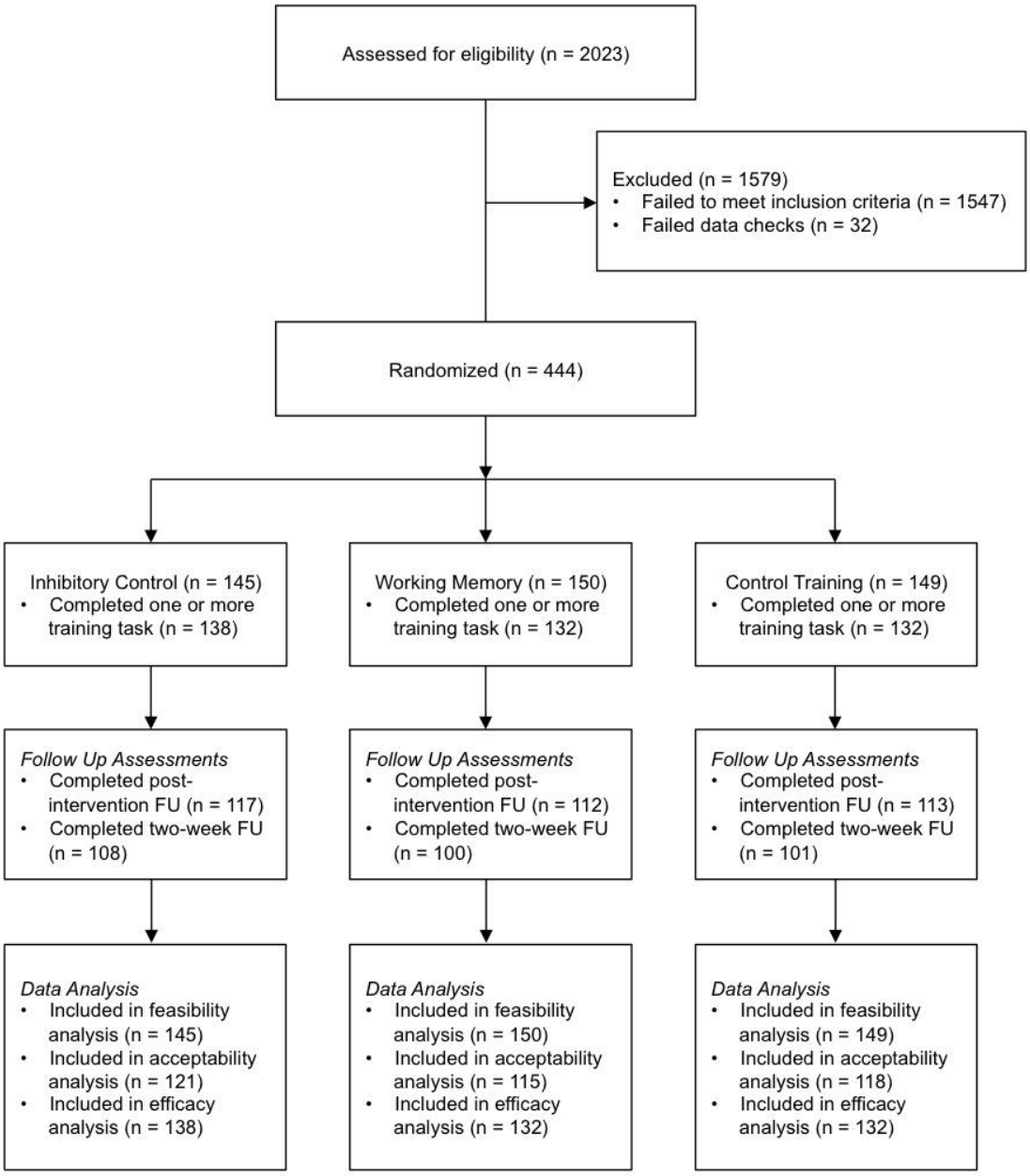

Figure 8.1. CONSORT flow diagram.

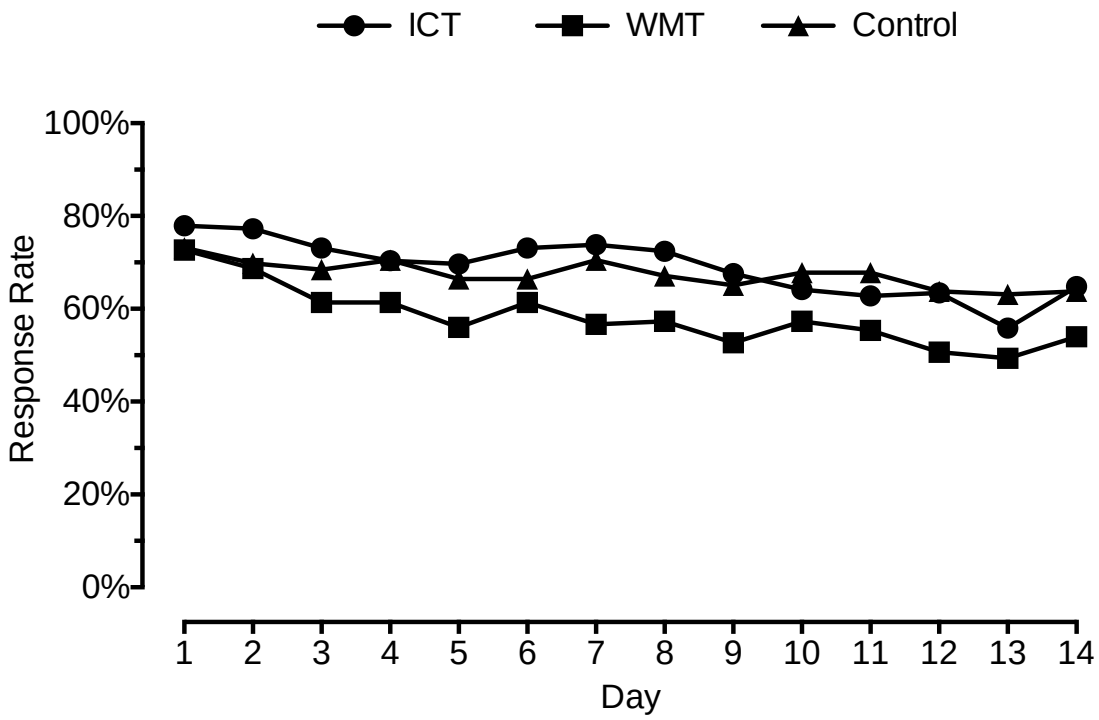

Figure 8.2. Daily response rates across the 14-day training phase. Plotted are response rates for each training day. Responses are separated by intervention group (Inhibitory Control [IC] = circles; Working Memory [WM] = squares; Control = triangles).

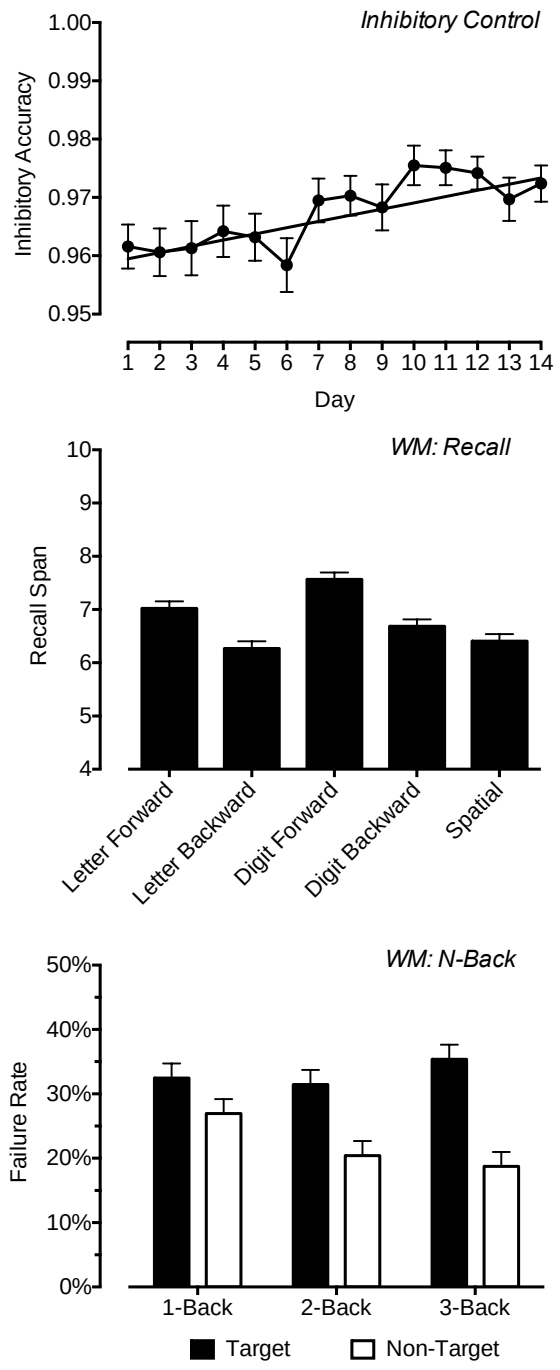

Figure 8.3. Cognitive task performance. Plotted is cognitive performance on the inhibitory control task (top panel), working memory recall tasks (middle panel), and working memory N-Back (bottom panel). Point estimates represent mean values of best fit from linear mixed models with standard errors.

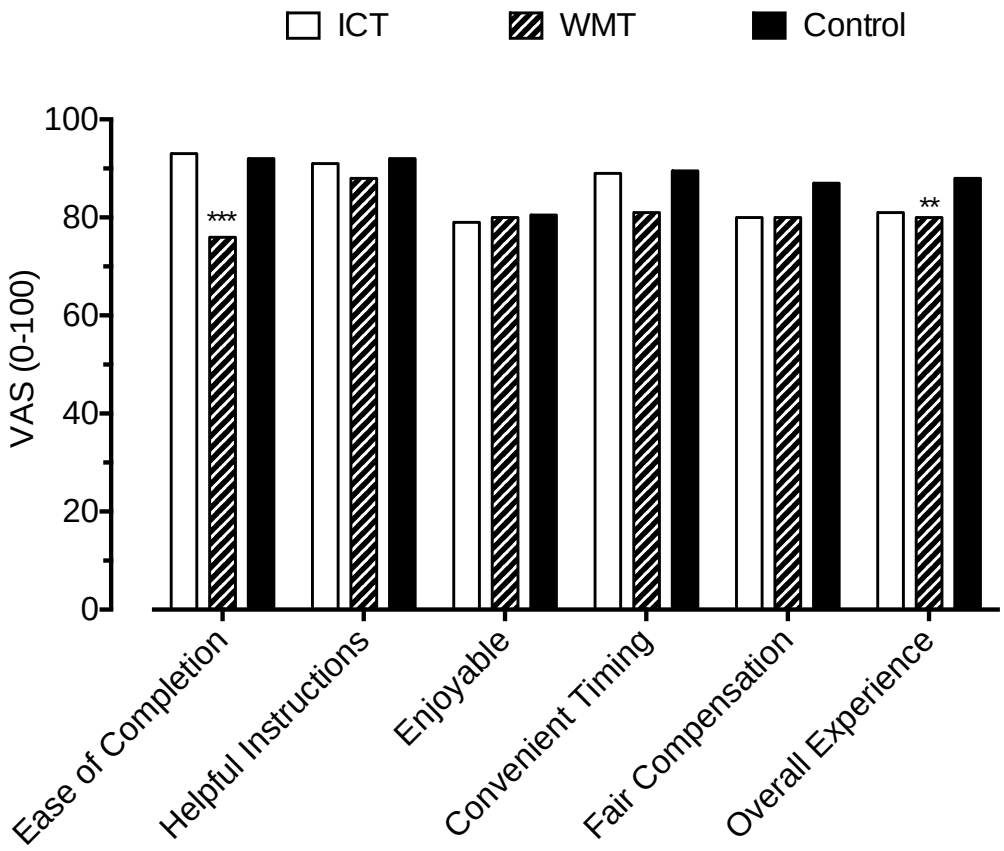

Figure 8.4. Study acceptability measures. Median values for acceptability measures on the Treatment Acceptability Questionnaire (TAQ) completed at the end of the training phase. All items were completed on a 100-point visual analog scale (VAS).
** $p < .01$; *** $p < .001$ compared to the Control group

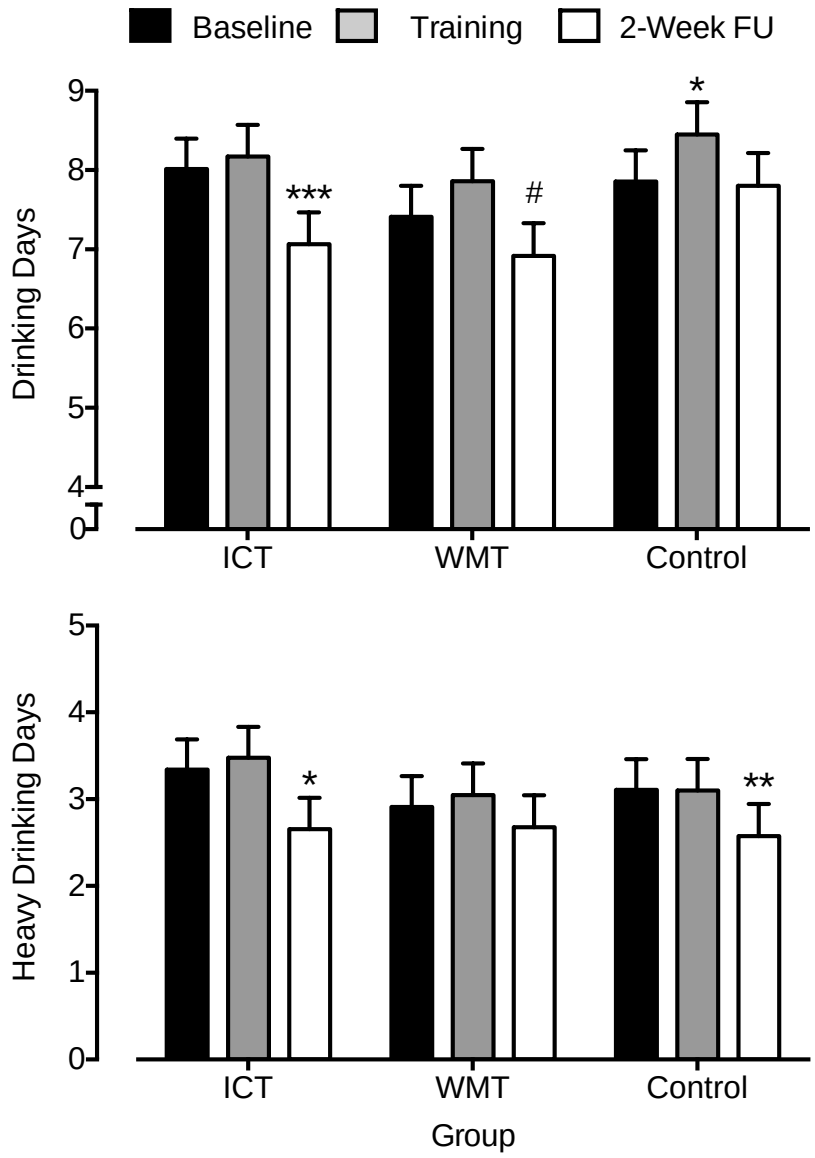

Figure 8.5. Drinking and heavy drinking outcomes at pre-, during-, and post-intervention. Plotted are number of drinking days and heavy drinking days over a 14-day period before (black bar), during (gray bar), and after (white bar) the training phase. Point estimates represent mean values of best fit from linear mixed models with associated standard errors estimates separated by Inhibitory Control Training (ICT), Working Memory Training (WMT), and Control groups.

# $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$ as compared to pre-training baseline value within a group

Chapter 9

USING BEHAVIORAL ECONOMIC VARIABLES TO PREDICT CHANGES IN ALCOHOL CONSUMPTION FOLLOWING COGNITIVE INTERVENTION

(Experiment 5b; Strickland et al., in preparation)

Introduction

Alcohol use disorder (AUD) remains a persistent and pervasive public health concern. The United States experiences annual economic costs upwards of $250 billion due to excessive drinking and 5.3% of global mortality is attributable to alcohol consumption (Sacks et al. 2015; World Health Organization 2018). Treatment approaches including pharmacological, cognitive-behavioral, and brief interventions have shown some success in reducing alcohol consumption (see reviews by Carroll and Kiluk 2017; Kranzler and Soyka 2018; Moyer et al. 2002). However, less is known about the person-level characteristics that may predict changes in alcohol consumption over time. Such research is important given the benefits that individualized treatment approaches may have for producing effective and efficient reductions in alcohol consumption (see discussion of the benefits of personalized medicine in Hamburg and Collins 2010). Determining individual factors that predict reduced alcohol use may also afford the opportunity to identify novel risk or protective factors that can be targeted for modified or novel intervention approaches.

One set of factors potentially related to AUD and other substance use disorders is choice and decision-making bias explained by behavioral economic theory. The application of behavioral economics to drug-taking behavior has proven useful for addiction science, broadly, and alcohol research, specifically (see reviews in Bickel et al. 2016a; MacKillop 2016). Behavioral economics is broadly concerned with the mechanisms by which an individual's decision-making is informed and described by concepts at the intersection of psychological and microeconomic theory. Although

traditional research has focused on commodities such as food or other commercial goods, more recent work has translated these theories to understand drug use as an experimental outcome.

Two popular behavioral economic mechanisms are demand and delay discounting. Demand is operationally defined as the consumption of a good at a given cost and a behavioral economic demand curve describes this functional relationship across a range of costs or prices. Demand curve analyses parameterize two independent behavioral mechanisms underlying drug consumption, demand intensity (i.e., consumption at unconstrained cost) and demand elasticity (i.e., sensitivity of consumption to changing cost) (Hursh and Silberberg 2008). The promise of examining drug valuation and reinforcement within a demand framework is that this approach accounts for and describes the multi-dimensional nature of reinforcement rather than viewing reinforcement as a homogenous construct (Johnson and Bickel 2006; Hursh and Silberberg 2008). Studies in the human laboratory and clinic have demonstrated that behavioral economic demand for alcohol and other substances is characterized by prototypic decreases in consumption with increases in price, is consistently related to measures of drug consumption and severity, and is sensitive to state-level changes in drug valuation, such as during withdrawal and cue presentation (e.g., Acker and MacKillop 2013; Aston et al. 2015; Amlung and MacKillop 2014; Bruner and Johnson 2014; MacKillop et al. 2008; MacKillop et al. 2012; Murphy and MacKillop 2006).

Delay discounting, and more specifically delayed reward discounting, refers to the systematic devaluation of a reinforcer with delay to its delivery (Odum 2011; Rachlin 2006). Considerable evidence suggests that excessive delay discounting, or greater reductions in reinforcer value with delay, contributes to the etiology and persistence of alcohol and substance use disorders (see meta-analyses and reviews by Amlung et al. 2017b; Bickel et al. 2012; MacKillop et al. 2011). It has also been suggested that delay

discounting represents a trans-disease process linking substance use with other negative health behaviors, such as gambling and overeating (Bickel et al. 2012). A reinforcer pathology perspective further posits that delay discounting and behavioral economic demand independently and interactively contribute to substance use through an extreme preference for immediate consumption of drug commodities combined with high valuation for those reinforcers (Bickel et al. 2017).

Considerable research has evaluated behavioral economic demand and delay discounting in AUD within a cross-sectional context (see review in MacKillop 2016). However, far less research has studied these concepts prospectively within the context of pharmacological or behavioral interventions. In this regard for alcohol use, a growing body of literature has evaluated behavioral economic demand and discounting in college students receiving brief alcohol interventions (e.g., personalized drinking feedback, brief motivational interviewing) (Dennhardt et al. 2015; MacKillop and Murphy 2007; Murphy et al. 2015). These studies have demonstrated less intense and more elastic alcohol demand following brief intervention (Dennhardt et al. 2015; Murphy et al. 2015). Furthermore, greater reductions in alcohol demand, but not monetary delay discounting, have been associated with lower drinking quantity (e.g., drinks/week) and severity (e.g., binge drinking, alcohol-related problems) at follow-up assessments (Dennhardt et al. 2015; MacKillop and Murphy 2007; Murphy et al. 2015). These findings suggest that changes in behavioral economic demand could be predictive of prospective changes in alcohol use for AUD, however are limited in generalizability by the focus on college students and use of relatively small samples (n = 51-133). These studies have also focused on more general discounting rates for money rather than commodity-specific discounting rates for alcohol. This distinction is important given the observation that commodity-specific values can provide improved prediction of health outcomes (see examples in Johnson and Bruner 2012; Tsukayama and Duckworth 2010)

The purpose of the current analysis was to evaluate behavioral economic demand and delay discounting as predictors of alcohol consumption in a sample of adults completing brief cognitive training interventions. Participants completed two weeks of daily cognitive training tasks as a part of a study on the feasibility and acceptability of delivering cognitive training via crowdsourcing (Strickland et al. in press; see Chapter 8). This secondary analysis focused on three primary questions: 1) whether brief training produced changes in alcohol demand or delay discounting, 2) whether baseline behavioral economic measure outcomes could predict future alcohol consumption, and 3) whether changes in behavioral economic measure outcomes would correspond to changes in alcohol use. Based on prior research (Dennhardt et al. 2015; MacKillop and Murphy 2007; Murphy et al. 2015), our hypotheses were that demand, but not discounting, would decrease following intervention, that behavioral economic outcomes would predict future alcohol use behaviors with improved prediction by commodity-specific measures, and that greater reductions in demand would be associated with greater reductions in alcohol consumption.

Methods

Participants and Screening

Participants were recruited from the crowdsourcing website Amazon Mechanical Turk (mTurk). Crowdsourcing leverages Internet resources to sample a large number of individuals from varied geographic regions and with varied health histories. Recent research has demonstrated the utility and validity of using crowdsourcing in psychological and addiction science (see reviews by Chandler and Shapiro 2016; Strickland and Stoops 2019). Inclusion criteria for this study were 1) self-reported past week alcohol use, 2) 21 years of age or older, 3) interest in a 2-week study on mTurk, and 4) meet criteria for *DSM-5* AUD according to a validated brief questionnaire (Hagman 2017). Participants completed a short screening questionnaire to determine

study eligibility. Access to this screening survey was limited to individuals with at least 100 completed mTurk tasks, a ≥ 95% approval rating on prior tasks, and United States residence (see similar qualifications in Cunningham et al. 2017; Strickland and Stoops 2015). The University of Kentucky Medical Institution Review Board reviewed and approved all procedures and the study was registered on clinicaltrials.gov (NCT03438539)

General Procedures

Detailed information on study procedures and primary outcomes regarding feasibility, acceptability, and initial efficacy are described previously (Strickland et al. 2019; see Chapter 8). Briefly, qualifying participants first completed a baseline survey containing alcohol use history and behavioral economic variables. Participants were then randomized to training conditions, including a working memory training, an inhibitory control training, or a control training condition (i.e., control training involved completing arithmetic problems). Participants were also randomized to receive alcohol normative feedback or control feedback during the baseline survey. Alcohol normative feedback delivery involved delivery of a statement standardized based on the reported average number of standard drinks per week and participant age and gender with individual percentile rank of weekly alcohol consumption compared to values from the National Survey on Drug Use and Health (Center for Behavioral Health Statistics 2017).

After randomization, participants completed daily training tasks for two weeks. Follow up surveys were completed at the end of the training phase and two weeks after the end of the training phase. These surveys included information about alcohol use for comparisons across the study phases (i.e., Baseline, Post-Training, and Two-Week Follow Up). The post-training follow up also included the behavioral economic measures completed at baseline.

Measures

Behavioral Economic Demand

Behavioral economic demand for alcohol and soda was evaluated using commodity purchase tasks (e.g., Morris et al. 2017; Murphy and MacKillop 2006; Strickland and Stoops 2017). Standard vignettes were used for each task in which participants were instructed that they would have to consume all purchases in a single day, that they could only get the commodity from this source, could not stockpile, and had no commodity available from previous days. Understanding of these stipulations was verified by two questions related to the instructions. Purchases were evaluated across 13 monetary increments ranging from $0.00 [free] to $11/unit, presented sequentially (full range: $0.00 [free], $0.01, $0.05, $0.13, $0.25, $0.50, $1, $2, $3, $4, $5, $6, $11). Commodities were one standard US drink (alcohol task) or one 12 oz. serving of soda (soda task). Task order was randomized.

Data from commodity purchase tasks were analyzed using the exponentiated demand equation (Koffarnus et al. 2015):

$$Q = Q_0 * 10^{k*(e^{(-\alpha*Q_0*C)}-1)}$$

where Q = consumption; Q_0 = derived demand intensity; k = a constant related to consumption range (*a priori* set to 2); C = commodity price; and α = derived demand elasticity. Demand intensity reflects the theoretical consumption of a commodity at a unit price of zero or near-zero. Demand elasticity reflects the sensitivity of consumption to changes in price. The exponentiated demand equation provided an excellent fit to individual data (mean of individual demand curve fits R^2: Alcohol = .87; Soda = .91). Intensity and elasticity were selected for analysis because prior factor analytic studies have demonstrated that these measures reflect the two factors underlying the purchase task factor structure for alcohol and other substances (Aston et al. 2017; Bidwell et al.

2012; Epstein et al. 2018; Mackillop et al. 2009). Other evidence also suggests that these derived measures show greater stimulus-selectivity than other purchase task measures (e.g., breakpoint) (Strickland and Stoops 2017). Intensity and elasticity were log-transformed prior to analysis to achieve normality.

Delay Discounting

Delay discounting rates for money and alcohol were evaluated using a 5-trial adjusting delay task (Koffarnus and Bickel 2014). Participants were instructed to select between $1000 at a delay and $500 available immediately for money or alcohol commodities in respective tasks (order randomized). Initial choices were set at a three-week delay and adjusted down (shorter delay following immediate choice) or up (longer delay following delayed choice) after each choice. The effective delay 50% (ED_{50}) was determined after five choices. The primary outcome was delay discounting rates (k) calculated as the inverse of ED_{50} (Koffarnus and Bickel 2014). Delay discounting rates were log-transformed prior to analysis to achieve normality. Prior research has validated this 5-trial adjusting delay task by showing correspondence with traditional adjusting amount delay discounting tasks (Cox and Dallery 2016; Koffarnus and Bickel 2014). This task version was selected because of its benefits for the online setting, including rapid assessment with minimal computing requirements.

Study Outcomes

The Timeline Followback (TLFB) assessment was used during each study phase to evaluate alcohol consumption. Participants completed a calendar over a two-week period that encompassed each study phase and recorded the number of standard drinks consumed. Drinks were reported by type (e.g., beer, wine, liquor), but were totaled to US standard drinks for the purpose of analysis. Previous research has demonstrated the reliability of the TLFB when delivered using computerized methods (Sobell et al. 1996). Two study outcomes were determined from the TLFB: a) proportion drinking days and b)

proportion heavy drinking days. Heavy drinking was defined using National Institute on Alcohol Abuse and Alcoholism (NIAAA) guidelines of ≥5/≥4 drinks/day for men/women (National Institute on Alcohol Abuse Alcoholism 2007).

Alcohol craving was also evaluated at each study phase. Craving was assessed using a best-practice approach of using multiple items indexing a continuum of urges and desires (Kozlowski et al. 1989; Sayette et al. 2000). Five items rated on an 11-point Likert scale (0 [not at all] to 10 [constantly]) were used to evaluate craving over each two-week period (e.g., "How often did you have a strong urge to drink alcohol?"). Scores were totaled for a possible range of 0 to 50 with higher values indicating greater alcohol craving.

Data Analysis

Four hundred and seventy-six participants qualified and completed the baseline survey. Seventy-four participants were removed from data analysis for failing to pass attention and data quality checks (n = 32) or failing to complete at least one training session (n = 42). Given the focus of this manuscript on behavioral economic measures, we limited the sample to individuals providing systematic data on the alcohol purchase task. Purchase task data were evaluated for non-systematic data using standardized criteria (Stein et al. 2015). Sixty-four individuals provided non-systematic or non-analyzable data, 39% of which were due to zero consumption at all price points (n = 25). This resulted in a final sample for analysis of 338 participants. Participants reporting systematic purchase task data reported lower SIP scores (9.4 versus 13.3) and were more likely to be male (51% versus 36%). Other demographic and alcohol use history did not significantly differ between groups.

Bivariate associations involving behavioral economic variable and participant demographic and alcohol use characteristics collected at baseline were first evaluated using Pearson correlations. Changes in behavioral economic demand and delay

discounting rates from baseline to post-training were then evaluated using linear mixed effect models parameterizing the effect of time. Additional models were evaluated that evaluated the differential effects of training group and normative feedback group.

Linear mixed effect models were then used to test the relationship between behavioral economic variables and the three study outcomes: 1) proportion drinking days, 2) proportion heavy drinking days, and 3) alcohol craving. These endpoints were selected to provide clinically relevant measures of objective and subjective alcohol consumption that were continuous so as to facilitate model fitting and enhance specificity. Each model tested two effects of interest involving behavioral economic measures. First, the association between baseline behavioral economic values and outcomes across the study phases were evaluated through baseline scores included in each model. Second, the association between changes in behavioral economic measures and changes in alcohol consumption was evaluated by parameterizing interactions between change scores and study phase. Models were first tested without covariates (i.e., unadjusted models). Covariates were then included to account for demographic characteristics (age, sex, education, employment, and income) and AUD severity (severe versus non-severe AUD). These models were conducted collapsing across intervention groups given the lack of significant interventions observed for changes in behavioral economic variables by condition (see Results).

SPSS Statistics (IBM; Armonk, NY) and R statistical language with the *nlme* package (Pinheiro et al. 2018) were used for analyses. All inferential tests were two tailed and used an alpha rate of .05.

Results

Demographics and Alcohol Use Behaviors

Table 9.1 contains participant demographic and alcohol use variables. Participants were on average 34 years old and a majority had a college education and reported

current part- or full-time employment. Participants endorsed an average of 5.3 DSM-V AUD symptoms and 12 standard alcohol drinks per week with 4 drinks per drinking occasion. Baseline behavioral economic variables are also included in Table 9.1. Delay discounting rates for alcohol were significantly higher than discounting rates for money, $t_{337} = 9.11$, $p < .001$, $d_z = 0.50$.

Baseline Associations

Table 9.2 contains bivariate correlations involving demand and discounting variables with demographic and alcohol use history. Alcohol demand was significantly associated with all alcohol use variables with the exception of days of alcohol use at baseline, significant *r* values = .14 to .38. Men also reported significantly higher alcohol demand intensity and less elastic alcohol demand compared to women. Correlations involving monetary or alcohol delay discounting rates were also significantly related to all alcohol use variables, *r* values = .13 to .38.

Soda demand intensity or elasticity were significantly associated with the number of AUD criteria endorsed, presence of severe AUD, and alcohol craving, although these correlations were of a small effect size, significant *r* values = .15 to .18. No other alcohol use variables were significantly associated with soda demand intensity or elasticity.

Changes in Behavioral Economic Variables

A significant decrease in alcohol demand intensity and significant increase in alcohol demand elasticity were observed from baseline to post-training, *p* values < .001 (see Figure 9.1 for model estimated values). No changes in delay discounting rates for money, $p = .26$, or alcohol, $p = .65$, were observed (Figure 9.1). These effects did not differ by cognitive training group or normative feedback group as indicated by non-significant interactions in tested models.

No changes in soda demand intensity or elasticity were observed (Figure 9.1). Similarly, these effects again did not differ by cognitive training group or normative feedback group.

Predictive Utility of Behavioral Economic Variables

Baseline Predictors

Table 9.3 contains model estimates for prediction of proportion days alcohol use, proportion heavy drinking days, and alcohol craving over the three study phases by baseline behavioral economic variables (see *Baseline* rows). Alcohol demand intensity was positively and significantly associated with all three outcomes and the relationships involving heavy drinking days and alcohol craving remained significant in adjusted models. Alcohol demand elasticity was negatively and significantly associated with heavy drinking days in an unadjusted model, however was not significant after controlling for demographic and alcohol use covariates.

Monetary and alcohol delay discounting rates were significantly and positively associated with heavy drinking days and alcohol craving in unadjusted and adjusted models. Alcohol discounting rates were also significantly associated with days drinking in unadjusted and adjusted models. These outcomes represented greater alcohol use and craving with steeper delay discounting rates.

Change as a Predictor

Table 9.3 also contains model estimates predicting changes in alcohol use as a function of changes in alcohol demand and discounting rates. Changes in alcohol demand intensity significantly predicted changes in alcohol craving at post-training and two-week follow up. These effects represented greater reductions in alcohol craving with greater reductions in alcohol demand intensity. In contrast, changes in alcohol demand elasticity were a significant predictor of days of alcohol consumption at post-training and two-week follow up and heavy drinking days at post-training. These effects represented

greater reductions in alcohol use with more elastic alcohol demand (i.e., greater price elasticity).

Changes in monetary delay discounting rates were a significant predictor of heavy drinking days and alcohol craving at post-training, although this second effect was not significant in adjusted models. These effects represented greater reductions in heavy drinking and alcohol craving with greater reductions in monetary delay discounting (i.e., less steep discounting). Changes in alcohol delay discounting rates were not significant predictors in any model tested.

Soda Demand

Table 9.3 also contains baseline and change predictor estimates for soda demand intensity and elasticity. Baseline soda demand intensity was significantly and positively related to alcohol craving; however, this effect was not significant in an adjusted model. Baseline soda demand intensity and elasticity were not significantly related to any other variables.

Changes in soda demand intensity were significantly associated with the proportion of drinking days at post-intervention. This effect represented a greater proportion of drinking days with larger decreases in soda demand. No other significant effects were observed for changes in soda demand intensity or elasticity.

Discussion

The purpose of the present study was to evaluate behavioral economic demand and delay discounting as predictors of alcohol consumption in a sample of adults receiving brief cognitive training. Baseline associations involving behavioral economic variables and alcohol use history were significant and in the expected direction (MacKillop 2016). Alcohol demand intensity and elasticity, for example, were significantly associated with alcohol drinking quantity (e.g., drinks/occasion) and severity (e.g., severe AUD, SIP scores). Soda demand, in contrast, only showed small effect size correlations with AUD

severity. This finding supports the selectivity of demand measures for capturing commodity-specific valuation similar to prior research demonstrating the stimulus-selectivity of the purchase task procedure (Chase et al. 2013; Strickland and Stoops 2017). Associations evaluating the relationship between delay discounting rates for money and alcohol with alcohol use history were also significant, which is consistent with a considerable body of work linking excessive delay discounting with negative health behaviors such as substance use (Bickel et al. 2012; MacKillop 2016). That alcohol was more steeply discounted than money also reproduces the well-described finding that consumable goods are discounted at a greater rate than money (e.g., Baker et al. 2003; Bickel et al. 2011b; Charlton and Fantino 2008; Johnson et al. 2007). These baseline associations collectively provide clear replication of existing behavioral economic findings thereby lending support for the validity of the data collection and sampling procedure.

The first goal of this study was to test whether cognitive training would produce changes in alcohol demand and delay discounting. Alcohol demand became less intense and more elastic following training independent of training condition. Prior research has observed similar reductions in alcohol demand following brief motivational interviewing or personalized feedback in college student samples (Dennhardt et al. 2015; Murphy et al. 2015). Importantly, reductions in demand intensity in one of these studies were also observed in an assessment only control condition (Murphy et al. 2015). This finding could explain the reductions observed in the control training group here, which included assessment of alcohol use behaviors as well as daily engagement in a control training task. That changes did not emerge as a function of training conditions is also not surprising given that the effect of training was modest in the overall study and, in the case of heavy drinking days, also observed in the control training group (for more discussion of these overall effects see Strickland et al. in press; Chapter 8). No changes

were observed in soda demand indicating that the reductions in alcohol demand were commodity specific. Similarly, no changes in discounting were observed, which suggests a mechanistic specificity to behavioral economic demand (and see similar results by Dennhardt et al. 2015). Future studies will benefit from evaluating more intensive or potent interventions as well as including additional control groups that experience minimal assessment.

The second goal of this study was to test behavioral economic variables collected at baseline as predictors of alcohol consumption and craving throughout the training and follow-up period. Each behavioral economic variable was a significant predictor of at least one evaluated outcome. These significant relationships ranged from alcohol demand elasticity as a significant predictor of heavy drinking days in unadjusted models to alcohol discounting rates as a significant predictor of all variables in unadjusted and adjusted models. These findings are again similar to those previously observed in college students receiving brief interventions targeting alcohol consumption (Dennhardt et al. 2015; MacKillop and Murphy 2007; Murphy et al. 2015) and extend these findings to a broader community sample. These prospective relationships are also consistent with a recent study in which alcohol demand intensity and elasticity were significant and incremental predictors of heavy drinking collected over a prospective 18-week period (Strickland et al. under review; Chapter 7). Participants in that study were also adults recruited using crowdsourcing and reported alcohol use that freely varied in the absence of intervention exposure. Although monetary delay discounting was not a significant predictor of heavy drinking in that study and was significant here, this discrepancy could be attributed to the differences in the populations sampled (i.e., individuals with past week alcohol use versus those meeting criteria for AUD). Taken together, these findings demonstrate that prospective relationships between behavioral economic variables and

alcohol use generalize across varied conditions, including participant characteristics and intervention context.

The third goal of this study was to evaluate if changes in behavioral economic measures corresponded to changes in alcohol use during training and follow-up periods. Models evaluating this interaction indicated significant relationships represented by greater reductions in alcohol consumption and craving with greater reductions in demand and discounting. Specifically, greater reductions in alcohol intensity were associated with greater reductions in alcohol craving whereas increases in alcohol elasticity were associated with greater reductions in days of drinking and heavy drinking. Similar correspondence between reductions in alcohol demand and drinking frequency and severity has been reported elsewhere (Dennhardt et al. 2015; Murphy et al. 2015). This is the first study, to our knowledge, to test such relationships with alcohol craving as an outcome. Cross-sectional research has found that higher levels of craving are related to greater demand, and in particular demand intensity, for alcohol and other substances (Aston et al. 2017; MacKillop et al. 2010a; Metrik et al. 2016; Strickland et al. 2016c). Human laboratory studies have also found that demand is sensitive to state-level changes in craving, such as following cue exposure and drug prime (Amlung et al. 2012; Amlung et al. 2015a; MacKillop et al. 2012). That changes in demand were associated with fluctuations in craving may prove particularly relevant in augmenting interventions for AUD given craving's recent inclusion as a diagnostic criterion for AUD in the *DSM-V* (American Psychiatric Association 2013).

Decreases in discounting rates (i.e., less steep discounting) were associated with modest reductions in heavy drinking and craving, although these effects were only observed for monetary discounting and only during the training period. These effects were not unexpected given the lack of significant changes in discounting observed in the overall sample as well as previous studies reporting similar outcomes in colleges

students (Dennhardt et al. 2015; Murphy et al. 2012a). That we evaluated alcohol discounting in a community sample indicates that the lack of significant results in these prior studies were not likely attributable to differences between general and commodity-specific discounting or the use of college student samples.

Additional limitations of the analysis should also be noted. The sample was recruited using a crowdsourcing method that necessitated some loss over experimental control and potential concerns with generalizability. However, previous research has demonstrated that participants recruited using crowdsourced methods do not systematically differ between from community and college sources in problematic responding (e.g., responding in socially desirable ways) (Necka et al. 2016). Other studies have also found that participants find conveying sensitive material, such as substance use histories, easier over an online than in-person format (Kim and Hodgins 2017; Strickland and Stoops 2018b). With respect to generalizability, some research suggests that mTurk can provide similar or sometimes improved representation of the United States population when compared to other forms of convenience sampling (e.g., college student samples or those drawn from college towns) (Berinsky et al. 2012; Huff and Tingley 2015). This analysis also focused on a subset of individuals from the parent study due to non-systematic and/or inattentive responding. Sensitivity analyses indicated that individuals excluded due to non-systematic data did not significantly differ on the majority of demographic and alcohol use history variables (the exceptions being gender and SIP scores). Such outcomes suggest that the sample characteristics were not likely compromised by these participants' removal from the primary analyses.

The present study provides further evidence demonstrating the relevance of behavioral economic theory in addiction science. Specifically, this study replicated and extended previous findings relating behavioral economic measures relevant to alcohol use with changes in alcohol consumption following brief interventions. We found that

delay discounting and behavioral economic demand provided predictive and incremental knowledge about prospective alcohol consumption. We also found that changes in alcohol use and craving over the course of cognitive training and follow-up periods corresponded to changes in alcohol demand. Clinically, these findings suggest that demand and discounting are related to clinically relevant alcohol use outcomes and that demand specifically may serve as a dynamic marker of changes in alcohol consumption over time. Such findings and their clinical implications support the continued utility of applying behavioral economics within interventions development efforts.

Table 9.1. Participant Demographics and Alcohol Use Variables (N = 338)

	Mean	SD	IQR
Demographics			
Age	34.2	9.6	27-39
Male	50.6%	50.1%	
College	64.8%	47.8%	
Unemployed	6.5%	24.7%	
Income	49,000	30,000	30-70k
Alcohol Use			
AUD Number	5.3	2.8	3-7
Severe	42.3%	49.5%	
Drinks/Week	12.0	16.1	3-14
Drinks/Occasion	4.1	3.7	2-5
SIP	9.5	9.3	2-14
Baseline Outcomes			
Days Alcohol Use	0.55	0.31	0.29-0.86
Heavy Drinking Days	0.22	0.29	0.0-0.29
Alcohol Craving	22.3	12.8	11-33
Behavioral Economic			
Q_0 Alcohol	0.86	0.35	0.64-1.10
α Alcohol	-1.96	0.69	-2.38 - -1.72
k Money	-2.30	0.76	-2.71 - -1.87
k Alcohol	-1.89	1.03	-2.71 - -1.23
Q_0 Soda	0.59	0.39	0.34-0.80
α Soda	-1.31	0.57	-1.71 - -0.99

Note. AUD = DSM-V alcohol use disorder; SIP = short inventory of problems-alcohol; Q_0 = demand intensity; α = demand elasticity; *k* = discounting rates.

Table 9.2. Baseline Bivariate Correlations

	Q_0 Alcohol	α Alcohol	*k* Money	*k* Alcohol	Q_0 Soda	α Soda
Demographics						
Age	-.10	**.12***	-.03	.06	**-.18****	**.20*****
Male	**.29*****	**-.12***	.02	.08	.03	-.02
College	-.10	-.03	**-.20*****	**-.11***	-.11	.11
Unemployed	-.06	.02	.03	.03	-.04	.03
Income	.00	-.04	**-.17****	**-.13***	**-.15***	.04
Alcohol Use						
AUD Number	**.33*****	**-.18****	**.32*****	**.37*****	**.18****	**-.17****
Severe	**.26*****	**-.14***	**.21*****	**.26*****	**.15***	**-.15***
Drinks/Week	**.38*****	**-.26*****	**.21*****	**.17****	.03	-.07
Drinks/Occasion	**.47*****	**-.24*****	**.21*****	**.14***	.12	-.09
SIP	**.27*****	**-.12***	**.33*****	**.38*****	.12	-.03
Baseline Outcomes						
Days Alcohol Use	**.11***	-.10	**.13***	**.17****	.02	-.05
Heavy Drinking Days	**.33*****	**-.18****	**.25*****	**.25*****	.08	-.10
Alcohol Craving	**.27*****	**-.11*****	**.24*****	**.33*****	**.15***	-.10

Note. AUD = DSM-V alcohol use disorder; SIP = short inventory of problems-alcohol; Q_0 = demand intensity; α = demand elasticity; *k* = discounting rates. Statistically significant values are **bolded**.

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

Table 9.3. Behavioral Economic Predictors of Alcohol Use Variables

		Days Alcohol		HDD		Craving	
		UN	ADJ	UN	ADJ	UN	ADJ
Q_0 Alcohol	Baseline	**0.14***	0.06	**0.39*****	**0.31*****	**10.8*****	**5.18***
	Change*FU1	0.09	0.09	-0.02	-0.02	**-5.47****	**-5.47****
	Change*FU2	-0.02	-0.02	-0.05	-0.05	**-6.61****	**-6.60****
α Alcohol	Baseline	-0.03	0.00	**-0.07***	-0.05	-0.59	0.71
	Change*FU1	**0.05***	**0.05***	**0.05***	**0.05***	-0.86	-0.86
	Change*FU2	**0.06***	**0.06***	0.02	0.02	-0.08	-0.04
k Money	Baseline	0.06*	0.04	**0.09*****	**0.06****	**3.51*****	**2.13***
	Change*FU1	-0.03	-0.03	**-0.05****	**-0.05****	**-2.03***	-1.95
	Change*FU2	0.00	0.00	-0.02	-0.02	-1.43	-1.34
k Alcohol	Baseline	**0.06*****	**0.03***	**0.08*****	**0.05****	**4.01*****	**2.72*****
	Change*FU1	0.00	0.00	0.00	0.00	-0.55	-0.55
	Change*FU2	0.01	0.01	0.01	0.01	-0.47	-0.43
Q_0 Soda	Baseline	0.04	0.02	0.09	0.02	**6.71****	3.26
	Change*FU1	**0.10***	**0.10***	-0.01	-0.01	-1.30	-1.30
	Change*FU2	-0.06	-0.06	-0.03	-0.03	-0.75	-0.73
α Soda	Baseline	0.02	0.03	-0.03	0.00	-0.71	0.84
	Change*FU1	0.02	0.02	0.05	0.05	-0.76	-0.76
	Change*FU2	0.06	0.06	0.03	0.04	0.85	0.91

Note. HDD = heavy drinking days; Q_0 = demand intensity; α = demand elasticity; *k* = discounting rates. FU1 = post-training follow up; FU2 = two-week follow up. All values represent coefficient estimates for unadjusted models (UN) and models adjusting for demographic and alcohol covariates (ADJ). Statistically significant values are **bolded**.

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

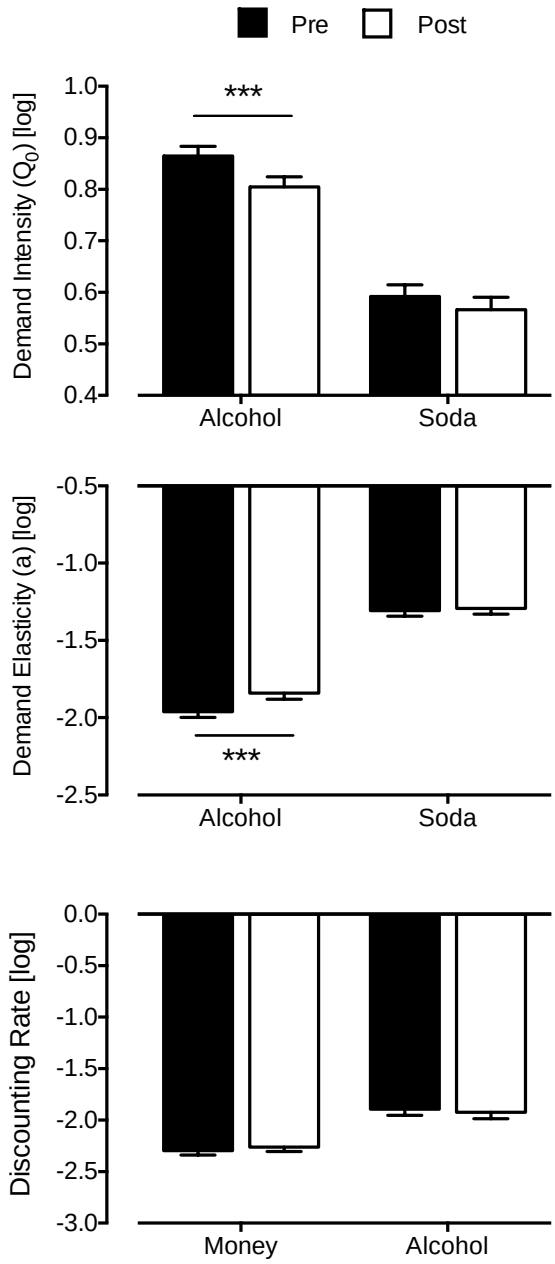

Figure 9.1. Changes in behavioral economic variables. Plotted are predicted changes in behavioral economic variables from baseline to post-training. All values represent estimate means and standard error based on linear mixed effects models. Values are collapsed across training and normative feedback groups.
*** $p < .001$ comparing pre-to-post values.

Chapter 10

GENERAL CONCLUSIONS

Introduction

This dissertation project presents a programmatic series of studies designed to demonstrate the utility of a behavioral economic demand framework for understanding substance use disorder in basic and applied settings. This final chapter will provide a general overview of the primary findings of each study as they relate to the dissertation aims, synthesize crosscutting conclusions from this research, and offer future directions for work on behavioral economic demand.

Summary of Findings

The first aim of this dissertation was to describe the contribution of behavioral economic demand to addiction science theory. Experiment 1 found that behavioral economic demand provided unique information about cannabis use above and beyond that provided by delay discounting rates. These findings were consistent with predictions from reinforcer pathology models thereby extending empirical support for this theory to cannabis use disorder.

The second aim of this dissertation was to provide novel assessments of the psychometric properties of the purchase task procedure. Experiment 2 demonstrated that the purchase task procedure provided a stimulus selective measure of alcohol and cigarette demand and that demand values adequately reflected valuation for the specific commodity under study. Experiment 3 evaluated a battery of purchase task measures relevant to prescription opioid use and found that these measures were construct valid across two independent samples and temporally reliable over a one-month period. These findings extended prior work by demonstrating stimulus selectivity in traditionally used purchase tasks as well as validating the use of the purchase task procedure for non-medical prescription opioid use.

The third aim of this dissertation was to establish the predictive and incremental validity of behavioral economic demand for describing prospectively collected substance use. Experiment 4 demonstrated the unique, predictive, and incremental validity of alcohol demand for evaluating alcohol use severity over an 18-week period. Experiment 5 similarly demonstrated that alcohol demand collected at baseline was predictive of patterns of alcohol use within the context of a brief cognitive intervention for alcohol use. These results collectively demonstrate that behavioral economic demand can provide valid, prospective prediction of alcohol use within and outside an intervention context.

The fourth aim of this dissertation was to demonstrate the utility of behavioral economic demand for interventions development research. Experiment 5 found that greater reductions in alcohol intensity following brief cognitive intervention were associated with greater reductions in alcohol craving whereas greater increases in alcohol elasticity were associated with greater reductions in days drinking and heavy drinking. Although not specific to a training condition, these results indicate that changes in demand are clinically relevant for alcohol use and may serve as a dynamic marker of fluctuations in alcohol consumption over time.

Crosscutting Conclusions

Several crosscutting conclusions may be generated based on these five experiments. The first and most consistently demonstrated across each experiment is that behavioral economic demand provides a reliable and valid measure of drug valuation. These findings are best summarized in a comparison and aggregation of effect sizes across these studies. Figures 10.1 and 10.2 depict forest plots for meta-analytic comparisons of four effect sizes reflecting the bivariate correlations (unadjusted) for intensity and elasticity measures with use frequency and use severity. Demand intensity was significantly associated with use frequency (r = .36 95% CI .23 to .49) and

use severity (r = .42 95% CI .35 to .50) with a medium-to-large effect size (Figure 10.1). Similarly, demand elasticity was associated with use frequency (r = -.28 95% CI -.34 to -.21) and use severity (r = -.24 95% CI -.31 to -.17) with a medium effect size (Figure 10.2). These effect size estimates were also generally consistent across each experiment with only a few exceptions (e.g., opioid demand and use frequency). These findings are consistent with a prior meta-analytic study showing a significant relationship between alcohol demand and alcohol use variables (Kiselica et al. 2016). Although these relationships only reflect simple, bivariate associations, that demand was systematically related to measures of substance use frequency and severity across these experiments supports the idea that behavioral economic demand reflects a measure of substance valuation adaptable for varied drug classes.

The second conclusion is that behavioral economic demand is a stimulus-selective measure reflecting valuation for the commodity under study. Although Experiment 2 was specifically designed to measure stimulus-selectivity, a similar commodity-specific relationship was observed in all five experiments. For example, Experiment 3 found that opioid and cannabis demand were most closely associated with the corresponding use frequency and use disorders. Similarly, alcohol and soda demand were associated with alcohol and soda use measures, respectively in three independent experiments (Experiments 2, 4, and 5). Stimulus-selectivity is an important and desirable quality of the purchase task procedure. These tasks, as typically utilized, are considered to index a commodity specific valuation and changes in demand thought primarily to reflect changes specific to that commodity valuation. The alternative to this exactness is a general representation of valuation for reinforcers without regard to the commodity or commodity type investigated. A domain-general response, while interesting for the potential of evaluating hypo- or hyper-valuation of reinforcers, would ultimately weaken the fidelity of purchase tasks as a behaviorally specific measure. The findings from this

dissertation provide consistent and clear support for such pharmacological exactness described for behavioral pharmacology and interventions development research.

The third conclusion is that demand measures provide incremental information about substance use above and beyond traditional measures of reinforcer valuation and beyond other behavioral economic variables. Specifically, demand provided relevant prediction of alcohol use variables above and beyond the AUDIT and measures of alcohol use disorder. Similarly, demand provided relevant information about alcohol, cannabis, and opioid use above and beyond measures of delay discounting. These findings are important because they indicate that demand measures are not simply a retooled means of evaluating existing measures of substance use, but instead represent a distinct behavioral mechanism underlying drug-taking behavior.

The fourth and final conclusion is that crowdsourcing is a valuable tool for the behavioral and addiction sciences. Each of the five experiments in this dissertation was conducted using crowdsourced sampling. This methodology afforded the opportunity to recruit populations that are difficult to sample in the human laboratory and clinic as well as to do so with much larger sample sizes than traditionally utilized. These experiments also expanded the existing literature on crowdsourcing by demonstrating that the sampling method can be effectively utilized for intensive longitudinal designs and interventions studies. Although limitations related to sampling bias and issues related to the convenience methodology should be acknowledged (see more extensive discussion of this issue in Chapter 2), the combination of findings from multiple sources, including college, community, clinic, and crowdsourced, that balances the relative strengths and weakness of these methods should ultimately benefit the research literature.

Theoretical Implications

The data presented in this dissertation may help inform theory development in behavioral economic demand in several ways. The first and primary contribution is to the

study of behavioral economic demand as a construct independent of delay discounting processes. To empirically synthesize some of this work from this dissertation, a principal components analysis (PCA) was conducted to evaluate where demand, discounting, and other alcohol use variables lie within a dimensional space. This PCA combined data collected in Experiments 4 and 5 (N = 608) and included relevant demand variables (alcohol demand intensity and elasticity), discounting variables (monetary discounting rates), and alcohol use measures (AUDIT, craving, and drinks/week). An un-rotated, two-component solution indicated that demand variables loaded in an orthogonal dimension from discounting rates (Figure 10.3). Craving loaded in a similar dimensional space as discounting rates whereas AUDIT scores and drinks/week fell between discounting and demand intensity. These findings support the notion that demand and discounting measures provide unique information, and that this information is also unique when compared to simple quantity-frequency and severity measures of alcohol use.

Such findings are consistent with contemporary theoretical perspectives that argue for an independent and interactive contribution of reward valuation and delay discounting processes in substance use and substance use disorder. These perspectives, such as reinforcer pathology and competing neurobehavioral decisions system theories, predict that maladaptive health behaviors are characterized by an interaction between excessive valuation for particular commodities and an extreme preference for immediate reinforcers (Bickel et al. 2017). These decision-making processes are thought to reflect a shift in the balance of neurobiological systems away from controlled and future-focused processes towards more reward and present-focused ones that co-occurs with negative health behaviors such as substance use.

Although a specific connection has not been emphasized in the literature, this reinforcer pathology approach seems largely concordant with the additive utility model of discounting proposed by Killeen (2009, 2015). This model posits that it is the utility of a good, and not its value (e.g., monetary value), that is discounted in a delayed or probabilistic setting:

$$v_t = (v^{\alpha} - kt^{\beta})^{1/\alpha}$$

The above model parameterizes an additive influence of utility and time with parameters governing power functions of utility (α) and time (β) scaling in addition to traditional discounting rates (*k* or λ in some versions of the model). It is possible that valuation as indexed by the purchase task procedure and behavioral economic demand may help to conceptualize and quantify individual and group differences in this hedonic scaling of utility (the α parameter). A relationship with utility scaling would represent a distinction mechanism from the impact of temporal discounting (*k*) and the influence of temporal scaling (β). Such a correspondence would then be consistent with the orthogonal loadings of discounting and demand parameters in the PCA evaluated here. Suggesting differences in utility scaling in substance-using populations is also consistent with previous observations of decreased loss aversion (and therefore loss-gain utility weighting) by individuals with a history of cocaine use (Strickland et al. 2017a). Empirical work comparing the parameters, and more precisely the utility tuning parameter, of discounting functions fit using an additive utility approach with those valuation parameters derived within a behavioral economic demand framework will be relevant for testing this theoretical connection.

Concepts of delay and reinforcer valuation from traditional learning theory, specifically work conducted by Rachlin and colleagues may also provide some insight into this distinction between discounting and demand measures. Rachlin (1992) elegantly proposed that delay discounting is a major (and maybe the single) contributor

to diminishing marginal utility of a commodity (see Raineri and Rachlin 1993 for additional empirical data and discussion). This perspective posits that the last unit of a larger commodity will be by necessity of sequential consumption consumed later than the last unit of a smaller commodity. Therefore, the marginal utility of the last unit of the larger commodity will be lower because it is discounted by the longer delay to its consumption. Borrowing from one of Rachlin's examples, the final apple in a purchase of ten apples will be eaten later and, therefore, devalued greater by its delay, than the last apple in a purchase of two apples. Diminishing marginal utility is the primary economic mechanism by which consumption is constrained within a purchasing situation (so much so that it is almost tautological to state as such in an econometric framework). This idea offered by Rachlin suggests that consumption rates are the foundational constraint on the value of a good. Such a viewpoint then places a heavy influence on an individual's discounting processes in determining consumption patterns and by inference suggests that discounting and demand measures should overlap significantly, if not entirely.

However, there is an additional pathway by which valuation and measures of demand could contribute independent of a temporal discounting process. Specifically, individuals presented with diminishing marginal utility constrained by consumption rate may learn to consume a commodity with increased efficiency and at higher rates to increase value (i.e., value increases directly with consumption rate in this model). Particular commodities, such as alcohol or other substances of abuse, are thought to be particularly susceptible to a feed-forward loop in which more efficient consumption leads to an increased value, increased consumption, and finally back to further increases in consumption efficiency (and so forth until an absolute limit of consumption is met). This mechanism would suggest that individuals who are more capable of learning to consume efficiently, whether for physiological or environmental reasons, will be more susceptible to substance use and a progression to substance use disorder. It is possible that

behavioral economic demand provides a measure of this valuation process that contributes to marginal utility independent of delay discounting processes. Although this connection to demand is certainly speculative, it provides an interesting explanation for independence between discounting and demand measures and an appealing theoretical model to test in further work.

The second potential contribution to theory development concerns the status of demand as a multifactorial construct. Demand is positioned as a multifaceted concept in most, if not all, approaches applying behavioral economic demand to reinforcer valuation. This multifaceted nature is argued to be one of the primary benefits of using behavioral economic demand over traditional drug self-administration measures, which treat reinforcement as a unitary concept. However, several findings from the experiments reported here are inconsistent with or argue against this idea of multidimensionality. For example, demand intensity tended to explain the majority/all of the variance related to demand variables in multivariable models predicting substance use variables. High correlations were also observed between demand intensity and elasticity variables that suggest a substantive overlap. The PCA described above, for example, found that demand intensity and elasticity loaded similarly within a two-component space.

There are a variety of reasons that could describe these findings of non-orthogonality, including theoretical and methodological ones. A theoretical explanation is that demand intensity and elasticity measure the same underlying concept of reinforcer valuation and are not orthogonal in nature as previously suggested. Although appealing in its parsimony, this explanation is largely unsatisfying when considering the rich biological and behavioral rationales underlying proposals that demand intensity and elasticity represent independent aspects of reinforcer valuation. From a behavioral standpoint, to use ideas from matching law, there is little reason to argue that the

asymptote of behavior (i.e., demand intensity) should show substantive overlap with the rate of external reinforcement that reduces an organism's behavior towards the target reinforcer (i.e., demand elasticity).

Methodological and measurement issues, rather than theoretical misspecification, could instead explain these discrepancies. The area representing transition from inelastic to elastic consumption on a demand curve that heavily influences elasticity estimation could be susceptible to increased within-person volatility and measurement error. This is because these portions of the curve, particularly those in the higher price ranges, may be experienced less frequently in everyday consumption and may be more difficult to accurately and consistently estimate by participants in hypothetical settings. Prices approaching zero are likely experienced more often in the natural environment in situations such as open bars or as gifted/free access commodities (this is not to say that there are no costs in these "free" situations, simply that the monetary cost is perceived as low or absent). Experience-dependent measurement error that differentially impacts elasticity would also explain the lower temporal reliabilities observed for demand elasticity compared to demand intensity measures. This reasoning suggests that demand curves generated using experienced, operant approaches should improve estimation of elasticity compared to those using the hypothetical purchase task approach. Improved measurement using experiential operant procedures could explain why research in the animal laboratory more frequently identifies elasticity as a critical variable underlying substance use behaviors (see example and discussion in Bentzley et al. 2014). No study, to date, has directly compared operant demand with purchase task approaches in the human laboratory. Such a study would be beneficial for testing this hypothesis of experience-dependent measurement error.

Directions for Future Work

The following details three distinct future directions for research on behavioral economic demand. These recommendations are rooted in a directive to advance both a basic science understanding of behavioral economic demand as well as its clinical applicability for the prevention and treatment of substance use disorders.

Advances in the Analysis of Behavioral Economic Demand Data

The past decade has witnessed a dramatic increase in the utilization of the purchase task methodology. However, studies critically evaluating the means by which behavioral economic demand data are analyzed have not kept a similar pace. One of the most prominent and discussed of these analytic challenges is how to appropriately account for zero consumption data in the modeling process. A number of solutions have been proposed, including a modified demand equation (Koffarnus et al. 2015; Strickland et al. 2016) and variations of non-linear mixed effect modeling (Liao et al. 2013; Yu et al. 2014; Zhao et al. 2016). However, consensus has yet to be reached about which of these methods, if any, appropriately account for zero values and have not yet been adapted for the analysis of cross-commodity demand.

Zero consumption values are not the only challenge in the analysis of behavioral economic demand data. For example, the selection of a scaling parameter (*k*) strongly influences the value of alpha in the exponential demand equation thereby preventing comparisons when scaling parameters differ (Hursh and Roma 2016). Evaluating analytic and methodological issues represents an area of active inquiry that should not be discounted in place of clinical research on demand.

Elucidating an Etiological Influence Utilizing Longitudinal Designs

Experiments 4 and 5 of this dissertation provided clear evidence for the prospective utility of behavioral economic demand for predicting future patterns of alcohol consumption. Longitudinal research would benefit from studies that evaluate behavior

over a longer time period and index these relationships across various stages of substance use disorder (e.g., during the development of or relapse to substance use disorder). Such studies would help to identify the extent to which behavioral economic demand and associated reinforcer valuation represent a developmental risk factor. As discussed in Chapter 7, research in the discounting literature suggests that steep delay discounting may play both an etiological role in substance use disorder as well as being a consequence of substance exposure (see reviews in MacKillop 2016; Perry and Carroll 2008). Whether and to what extent behavioral economic demand plays a causal or consequential role in alcohol and other substance use disorders remains a largely empirical question that future research is well positioned to address.

Large-Scale Clinical Application and Interventions Development

The inclusion of behavioral economic demand measures in phase II and phase III clinical research will provide an important test of clinical utility. Clinical research involving behavioral economic demand has largely focused on human laboratory studies (e.g., (e.g., Stoops et al. 2016; Bujarski et al. 2012) or changes in alcohol consumption in college students (e.g., MacKillop and Murphy 2007; Dennhardt et al. 2015). The use of demand measures in other clinical venues will be important for establishing the predictive validity for clinical efficacy as well as for helping to elucidate the behavioral mechanisms underlying effective and novel interventions.

One of the benefits of the purchase task procedure in this regard is its ease and efficiency of delivery. That the procedure may be completed in less than 5 minutes using simple pen and paper or computer delivery should facilitate this application in varied clinical venues. This format should also allow for the remote completion of measures at points in which laboratory or clinic visits are costly or time prohibitive. As such, the measurement of demand could provide a more fine-grained analysis of reinforcer valuation throughout intervention delivery than traditional clinic-based measures afford.

Interventions specifically developed to target behavioral economic demand may also prove useful given the consistent associations with measures of substance use and severity. In this regard, Murphy and colleagues (2012b) have explored interventions based on behavioral economics, namely proportionate reinforcement and delay discounting, as a supplement to other brief interventions. These methods have shown some efficacy for reducing alcohol and other substance use in college students (Murphy et al. 2012a; Murphy et al. 2012b; Yurasek et al. 2015). Developing similar interventions specifically targeting behavioral economic demand and ones applicable for a broader community setting is an important direction for future clinical work.

Final Impressions

Substance use disorders remain a prevalent and persistent economic and public health concern. Behavioral economic demand represents a framework by which the behavioral mechanisms underlying and environmental factors contributing to these disorders may be understood. This dissertation has provided a comprehensive overview of the basic and applied science applications of behavioral economic demand in addiction science. Specific advances in the understanding of stimulus-selectivity, novel applications to illicit substance use, and the prospective and incremental validity of the purchase task procedure have been demonstrated. These results emphasize an exciting platform for future applications of behavioral economics and behavioral economic demand in addressing substance use disorders.

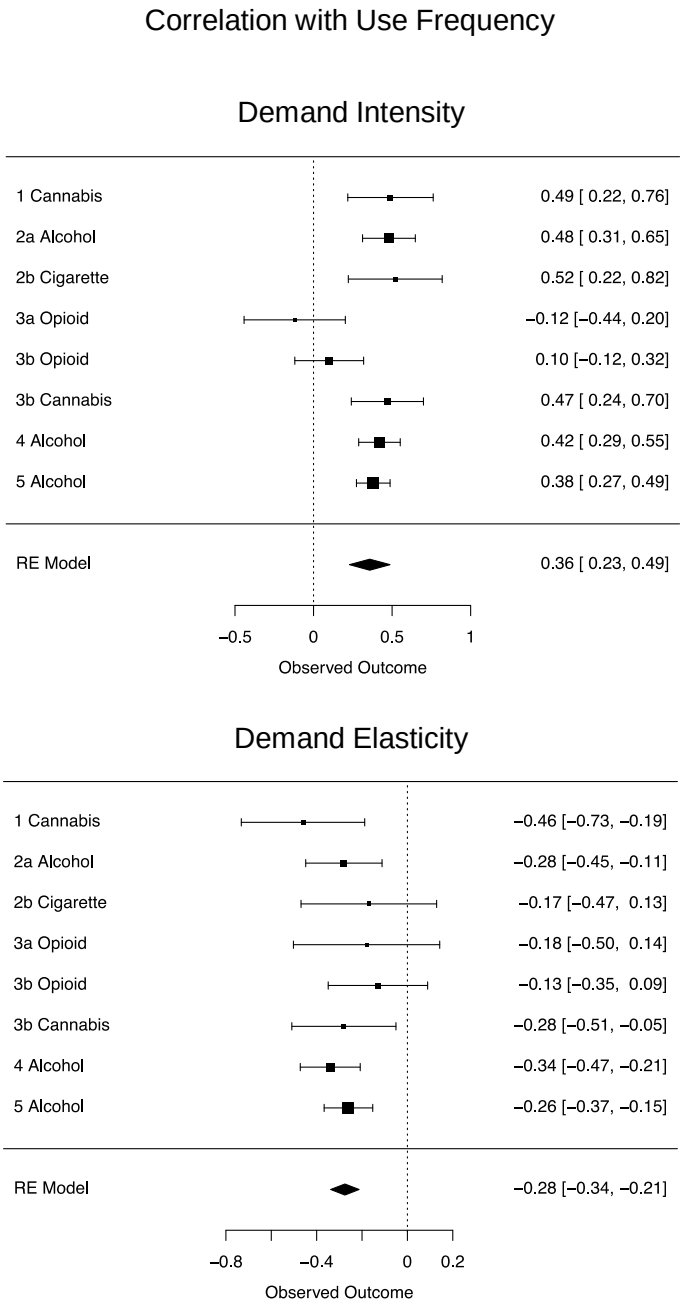

Figure 10.1. Meta-analysis of substance use frequency correlations. The above depicts a forest plot of the correlation between demand intensity (top panel) and demand elasticity (bottom panel) with substance use frequency measures in the five presented experiments.

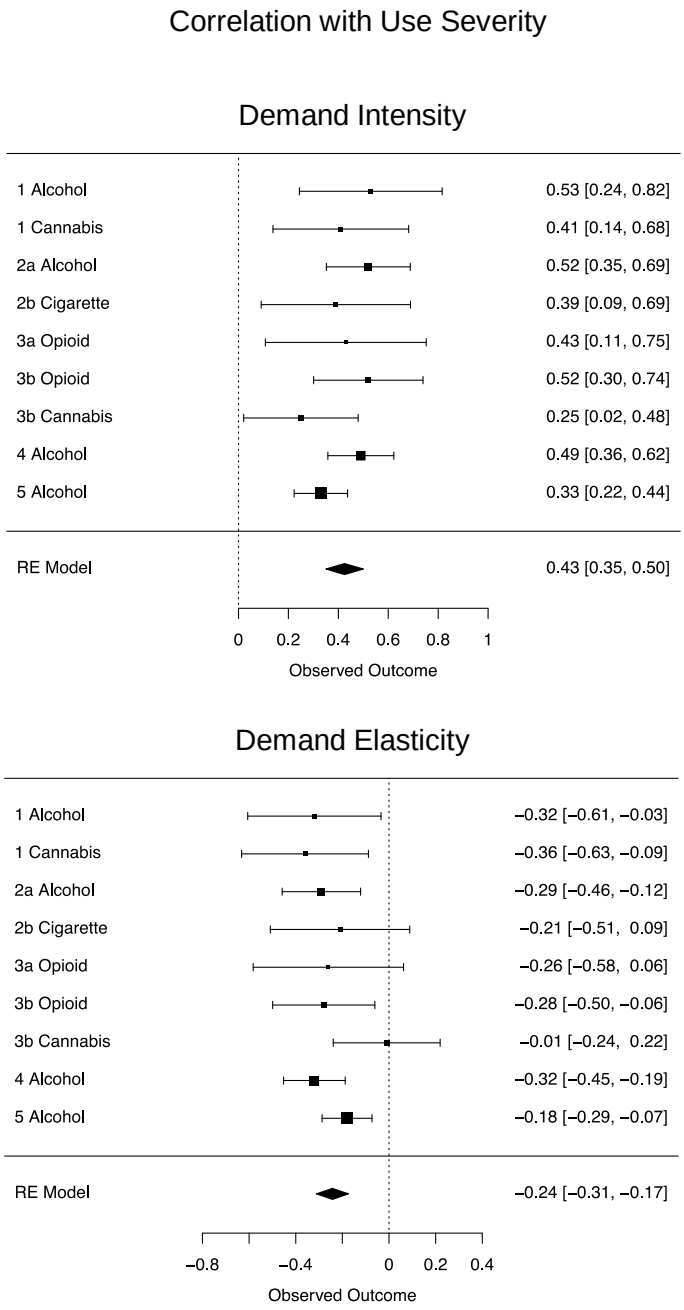

Figure 10.2. Meta-analysis of substance use severity correlations. The above depicts a forest plot of the correlation between demand intensity (top panel) and demand elasticity (bottom panel) with substance use severity measures in the five presented experiments.

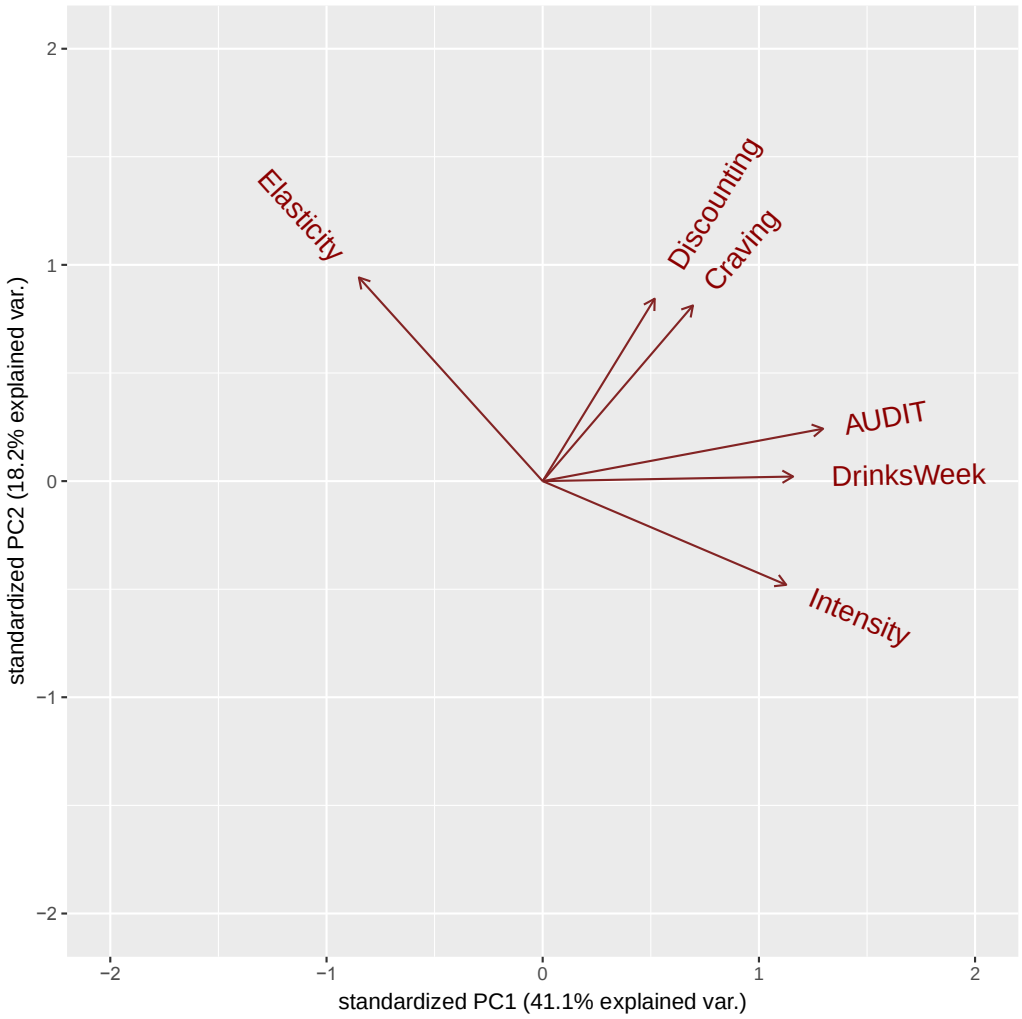

Figure 10.3. Principal components analysis. Plotted are the variable loadings on the first two components of a PCA evaluating alcohol use and behavioral economic variables.

Appendix

Example Commodity Purchase Task Instructions

This is a series of questions designed to assess choices for **alcohol** across changes in price. This information is entirely for research purposes. All questions about purchasing alcohol are completely hypothetical (pretend).

Imagine a TYPICAL DAY over the last month when you would drink alcohol. Assume that:
1) Alcohol refers to your preferred brand of alcohol.
2) The alcohol in question is the only alcohol available to you for the next 24 hours.
3) You have NO ACCESS to any other alcohol products other than those offered at these prices for the next 24 hours.
4) You have the same income/savings that you have now and you may buy as much or as little as you'd like.
5) You can drink without restriction, but you must drink all the alcohol you purchase in the next 24 hours.
6) You cannot stockpile or save alcohol for a later date.
7) You cannot sell the alcohol you purchase or give it away.
8) You did not drink any alcohol before making these decisions.

Think about how much **alcohol** you would purchase at each price. For the purpose of the task, one drink equals one 12 oz bottle/can of beer, one 5 oz glass of wine, or one 1.5 oz shot alone or in a mixed drink.

References

Abbey, A., T. Zawacki, P. O. Buck, A. M. Clinton, and P. McAuslan. 2001. Alcohol and sexual assault. *Alcohol Res Health* 25 (1):43-51.

Acker, J., and J. MacKillop. 2013. Behavioral economic analysis of cue-elicited craving for tobacco: a virtual reality study. *Nicotine Tob Res* 15 (8):1409-1416.

Acuff, S. F., and J. G. Murphy. 2017. Further examination of the temporal stability of alcohol demand. *Behav Processes* 141 (Pt 1):33-41.

Acuff, S. F., K. E. Soltis, A. A. Dennhardt, K. S. Berlin, and J. G. Murphy. 2018. Evaluating behavioral economic models of heavy drinking among college students. *Alcohol Clin Exp Res* 42 (7):1304-1314.

Adams, S., C. Mokrysz, A. S. Attwood, and M. R. Munafo. 2017. Resisting the urge to smoke: inhibitory control training in cigarette smokers. *R Soc Open Sci* 4 (8):170045.

Adkison, S. E., R. J. O'Connor, M. Chaiton, and R. Schwartz. 2015. Development of measures assessing attitudes toward contraband tobacco among a web-based sample of smokers. *Tob Induc Dis* 13 (1):7.

Alcorn, J. L., 3rd, E. Pike, W. S. Stoops, J. A. Lile, and C. R. Rush. 2017. A pilot investigation of acute inhibitory control training in cocaine users. *Drug Alcohol Depend* 174:145-149.

Allais, M. 2008. *Allais paradox*: Springer.

Allison, J. 1983. Behavioral substitutes and complements. In *Advances in Psychology*: Elsevier, 1-30.

Almeida, D. M., E. Wethington, and R. C. Kessler. 2002. The Daily Inventory of Stressful Events: An interview-based approach for measuring daily stressors. *Assessment* 9 (1):41-55.

Amazon. *Amazon Mechanical Turk requester tour* 2018 [cited. Available from https://requester.mturk.com/tour.

American Psychiatric Association. 2013. *Diagnostic and statistical manual of mental disorders (DSM-5®)*: American Psychiatric Pub.

Amlung, M., L. R. Few, J. Howland, D. J. Rohsenow, J. Metrik, and J. MacKillop. 2013. Impulsivity and alcohol demand in relation to combined alcohol and caffeine use. *Exp Clin Psychopharmacol* 21 (6):467-474.

Amlung, M., and J. MacKillop. 2014. Understanding the effects of stress and alcohol cues on motivation for alcohol via behavioral economics. *Alcohol Clin Exp Res* 38 (6):1780-1789.

Amlung, M., and J. MacKillop. 2015. Further evidence of close correspondence for alcohol demand decision making for hypothetical and incentivized rewards. *Behav Processes* 113:187-191.

Amlung, M., J. MacKillop, P. M. Monti, and R. Miranda Jr. 2017a. Elevated behavioral economic demand for alcohol in a community sample of heavy drinking smokers. *J Stud Alcohol Drugs* 78 (4):623-628.

Amlung, M., K. N. McCarty, D. H. Morris, C. L. Tsai, and D. M. McCarthy. 2015a. Increased behavioral economic demand and craving for alcohol following a laboratory alcohol challenge. *Addiction* 110 (9):1421-1428.

Amlung, M., D. D. Reed, V. Morris, E. R. Aston, J. Metrik, and J. MacKillop. 2019. Price elasticity of illegal vs legal cannabis: A behavioral economic substitutability analysis. *Addiction* 114:112-118.

Amlung, M., L. Vedelago, J. Acker, I. Balodis, and J. MacKillop. 2017b. Steep delay discounting and addictive behavior: a meta-analysis of continuous associations. *Addiction* 112 (1):51-62.

Amlung, M., A. Yurasek, K. N. McCarty, J. MacKillop, and J. G. Murphy. 2015b. Area under the curve as a novel metric of behavioral economic demand for alcohol. *Exp Clin Psychopharmacol* 23 (3):168-175.
Amlung, M. T., J. Acker, M. K. Stojek, J. G. Murphy, and J. MacKillop. 2012. Is talk "cheap"? An initial investigation of the equivalence of alcohol purchase task performance for hypothetical and actual rewards. *Alcohol Clin Exp Res* 36 (4):716-724.
Andersson, G., and N. Titov. 2014. Advantages and limitations of Internet-based interventions for common mental disorders. *World Psychiatry* 13 (1):4-11.
Anokhin, A. P., S. Golosheykin, and R. C. Mulligan. 2015. Long-term test-retest reliability of delayed reward discounting in adolescents. *Behav Processes* 111:55-59.
Arner, S., and B. Meyerson. 1988. Lack of analgesic effect of opioids on neuropathic and idiopathic forms of pain. *Pain* 33 (1):11-23.
Ashburn, M. A., and P. S. Staats. 1999. Management of chronic pain. *The Lancet* 353 (9167):1865-1869.
Aston, E. R., S. G. Farris, J. MacKillop, and J. Metrik. 2017. Latent factor structure of a behavioral economic marijuana demand curve. *Psychopharmacology (Berl)* 234 (16):2421-2429.
Aston, E. R., J. Metrik, M. Amlung, C. W. Kahler, and J. MacKillop. 2016. Interrelationships between marijuana demand and discounting of delayed rewards: Convergence in behavioral economic methods. *Drug Alcohol Depend* 169:141-147.
Aston, E. R., J. Metrik, and J. MacKillop. 2015. Further validation of a marijuana purchase task. *Drug Alcohol Depend* 152:32-38.
Atance, C. M., and D. K. O'Neill. 2001. Episodic future thinking. *Trends Cogn Sci* 5 (12):533-539.
Athamneh, L. N., J. S. Stein, and W. K. Bickel. 2017. Will delay discounting predict intention to quit smoking? *Exp Clin Psychopharmacol* 25 (4):273-280.
Backinger, C. L., C. N. Michaels, A. M. Jefferson, P. Fagan, A. L. Hurd, and R. Grana. 2008. Factors associated with recruitment and retention of youth into smoking cessation intervention studies--a review of the literature. *Health Educ Res* 23 (2):359-368.
Baker, F., M. W. Johnson, and W. K. Bickel. 2003. Delay discounting in current and never-before cigarette smokers: similarities and differences across commodity, sign, and magnitude. *J Abnorm Psychol* 112 (3):382-392.
Bardo, M. T., J. L. Neisewander, and T. H. Kelly. 2013. Individual differences and social influences on the neurobehavioral pharmacology of abused drugs. *Pharmacol Rev* 65 (1):255-290.
Bartneck, C., A. Duenser, E. Moltchanova, and K. Zawieska. 2015. Comparing the similarity of responses received from studies in Amazon's Mechanical Turk to studies conducted online and with direct recruitment. *PLoS One* 10 (4):e0121595.
Bates, D., M. Maechler, B. Bolker, and S. Walker. 2014. lme4: Linear mixed-effects models using Eigen and S4. *R package version* 1 (7):1-23.
Bauhoff, S., A. Montero, and D. Scharf. 2017. Perceptions of e-cigarettes: a comparison of adult smokers and non-smokers in a Mechanical Turk sample. *Am J Drug Alcohol Abuse* 43 (3):311-323.
Beard, C. L., A. L. Haas, R. E. Wickham, and V. Stavropoulos. 2017. Age of initiation and Internet gaming disorder: The role of self-esteem. *Cyberpsychol Behav Soc Netw* 20 (6):397-401.
Beck, R. C., and M. F. Triplett. 2009. Test-retest reliability of a group-administered paper-pencil measure of delay discounting. *Exp Clin Psychopharmacol* 17 (5):345-355.

Behrend, T. S., D. J. Sharek, A. W. Meade, and E. N. Wiebe. 2011. The viability of crowdsourcing for survey research. *Behav Res Methods* 43 (3):800-813.
Bentzley, B. S., and G. Aston-Jones. 2015. Orexin-1 receptor signaling increases motivation for cocaine-associated cues. *Eur J Neurosci* 41 (9):1149-1156.
Bentzley, B. S., and G. Aston-Jones. 2017. Inhibiting subthalamic nucleus decreases cocaine demand and relapse: therapeutic potential. *Addiction Biology* 22 (4):946-957.
Bentzley, B. S., K. M. Fender, and G. Aston-Jones. 2013. The behavioral economics of drug self-administration: a review and new analytical approach for within-session procedures. *Psychopharmacology (Berl)* 226 (1):113-125.
Bentzley, B. S., T. C. Jhou, and G. Aston-Jones. 2014. Economic demand predicts addiction-like behavior and therapeutic efficacy of oxytocin in the rat. *Proc Natl Acad Sci U S A* 111 (32):11822-11827.
Berinsky, A. J., G. A. Huber, and G. S. Lenz. 2012. Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis* 20:351-368.
Berman, H. L., and M. P. Martinetti. 2017. The effects of next-day class characteristics on alcohol demand in college students. *Psychology of Addictive Behaviors* 31 (4):488.
Bertholet, N., J. G. Murphy, J. B. Daeppen, G. Gmel, and J. Gaume. 2015. The alcohol purchase task in young men from the general population. *Drug Alcohol Depend* 146:39-44.
Bickel, W. K., R. J. DeGrandpre, S. T. Higgins, and J. R. Hughes. 1990. Behavioral economics of drug self-administration. I. Functional equivalence of response requirement and drug dose. *Life Sci* 47 (17):1501-1510.
Bickel, W. K., R. J. DeGrandpre, J. R. Hughes, and S. T. Higgins. 1991. Behavioral economics of drug self-administration. II. A unit-price analysis of cigarette smoking. *J Exp Anal Behav* 55 (2):145-154.
Bickel, W. K., D. P. Jarmolowicz, E. T. Mueller, and K. M. Gatchalian. 2011a. The behavioral economics and neuroeconomics of reinforcer pathologies: implications for etiology and treatment of addiction. *Curr Psychiatry Rep* 13 (5):406-415.
Bickel, W. K., D. P. Jarmolowicz, E. T. Mueller, M. N. Koffarnus, and K. M. Gatchalian. 2012. Excessive discounting of delayed reinforcers as a trans-disease process contributing to addiction and other disease-related vulnerabilities: emerging evidence. *Pharmacol Ther* 134 (3):287-297.
Bickel, W. K., M. W. Johnson, M. N. Koffarnus, J. MacKillop, and J. G. Murphy. 2014. The behavioral economics of substance use disorders: reinforcement pathologies and their repair. *Annu Rev Clin Psychol* 10:641-677.
Bickel, W. K., R. D. Landes, D. R. Christensen, L. Jackson, B. A. Jones, Z. Kurth-Nelson, and A. D. Redish. 2011b. Single-and cross-commodity discounting among cocaine addicts: the commodity and its temporal location determine discounting rate. *Psychopharmacology (Berl)* 217 (2):177-187.
Bickel, W. K., L. A. Marsch, and M. E. Carroll. 2000. Deconstructing relative reinforcing efficacy and situating the measures of pharmacological reinforcement with behavioral economics: a theoretical proposal. *Psychopharmacology (Berl)* 153 (1):44-56.
Bickel, W. K., L. Moody, and S. T. Higgins. 2016a. Some current dimensions of the behavioral economics of health-related behavior change. *Prev Med* 92:16-23.
Bickel, W. K., A. L. Odum, and G. J. Madden. 1999. Impulsivity and cigarette smoking: delay discounting in current, never, and ex-smokers. *Psychopharmacology (Berl)* 146 (4):447-454.

Bickel, W. K., S. E. Snider, A. J. Quisenberry, and J. S. Stein. 2017. Reinforcer pathology: The behavioral economics of abuse liability testing. *Clin Pharmacol Ther* 101 (2):185-187.
Bickel, W. K., A. G. Wilson, C. Chen, M. N. Koffarnus, and C. T. Franck. 2016b. Stuck in time: Negative income shock constricts the temporal window of valuation spanning the future and the past. *PLoS One* 11 (9):e0163051.
Bickel, W. K., R. Yi, R. D. Landes, P. F. Hill, and C. Baxter. 2011c. Remember the future: working memory training decreases delay discounting among stimulant addicts. *Biol Psychiatry* 69 (3):260-265.
Bidwell, L. C., J. MacKillop, J. G. Murphy, J. W. Tidey, and S. M. Colby. 2012. Latent factor structure of a behavioral economic cigarette demand curve in adolescent smokers. *Addict Behav* 37 (11):1257-1263.
Bidwell, L. C., J. Metrik, J. McGeary, R. H. Palmer, S. Francazio, and V. S. Knopik. 2013. Impulsivity, variation in the cannabinoid receptor (CNR1) and fatty acid amide hydrolase (FAAH) genes, and marijuana-related problems. *J Stud Alcohol Drugs* 74 (6):867-878.
Biernacki, P., and D. Waldorf. 1981. Snowball sampling: Problems and techniques of chain referral sampling. *Sociological methods & research* 10 (2):141-163.
Bock, B. C., R. Lantini, H. Thind, K. Walaska, R. K. Rosen, J. L. Fava, N. P. Barnett, and L. A. Scott-Sheldon. 2016. The Mobile Phone Affinity Scale: Enhancement and Refinement. *JMIR Mhealth Uhealth* 4 (4):e134.
Bohannon, J. 2016. Mechanical Turk upends social sciences: American Association for the Advancement of Science.
Bold, K. W., L. M. Fucito, W. R. Corbin, K. S. DeMartini, R. F. Leeman, H. R. Kranzler, and S. S. O'Malley. 2016. Daily relations among affect, urge, targeted naltrexone, and alcohol use in young adults. *Exp Clin Psychopharmacol* 24 (5):367.
Bolger, N., and J. Laurenceau. 2013. Intensive longitudinal methods. *New York, NY: Guilford*.
Borgida, E., B. Loken, A. L. Williams, J. Vitriol, I. Stepanov, and D. Hatsukami. 2015. Assessing constituent levels in smokeless tobacco products: A new approach to engaging and educating the public. *Nicotine Tob Res* 17 (11):1354-1361.
Borodovsky, J. T., L. A. Marsch, and A. J. Budney. 2018. Studying cannabis use behaviors with Facebook and web surveys: Methods and insights. *JMIR public health and surveillance* 4 (2).
Bowley, C., C. Faricy, B. Hegarty, J. J. S, L. S. J, J. K. P, and A. R. J. 2013. The effects of inhibitory control training on alcohol consumption, implicit alcohol-related cognitions and brain electrical activity. *Int J Psychophysiol* 89 (3):342-348.
Boynton, M. H., and L. S. Richman. 2014. An online daily diary study of alcohol use using Amazon's Mechanical Turk. *Drug Alcohol Rev* 33 (4):456-461.
Braitman, A. L., A. N. Linden-Carmichael, and J. M. Henson. 2017. Protective behavioral strategies as a context-specific mediator: A multilevel examination of within- and between-person associations of daily drinking. *Exp Clin Psychopharmacol* 25 (3):141-155.
Briones, E. M., and G. Benham. 2017. An examination of the equivalency of self-report measures obtained from crowdsourced versus undergraduate student samples. *Behav Res Methods* 49 (1):320-334.
Brody, G. H., T. Yu, J. MacKillop, G. E. Miller, E. Chen, E. M. Obasi, and S. R. Beach. 2014. Catecholamine levels and delay discounting forecast drug use among African American youths. *Addiction* 109 (7):1112-1118.

Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Maechler, and B. M. Bolker. 2017. Modeling zero-inflated count data with glmmTMB. *bioRxiv*:132753.
Brown, J. D., A. J. Goodin, and J. C. Talbert. 2018. Rural and Appalachian disparities in neonatal abstinence syndrome incidence and access to opioid abuse treatment. *The Journal of Rural Health* 34 (1):6-13.
Brown, M. T., and J. K. Bussell. 2011. Medication adherence: WHO cares? *Mayo Clin Proc* 86 (4):304-314.
Brown, S. A., P. W. Vik, T. L. Patterson, I. Grant, and M. A. Schuckit. 1995. Stress, vulnerability and adult alcohol relapse. *J Stud Alcohol* 56 (5):538-545.
Bruner, N. R., and M. W. Johnson. 2014. Demand curves for hypothetical cocaine in cocaine-dependent individuals. *Psychopharmacology (Berl)* 231 (5):889-897.
Buhrmester, M., T. Kwang, and S. D. Gosling. 2011. Amazon's Mechanical Turk: A New Source of Inexpensive, Yet High-Quality, Data? *Perspect Psychol Sci* 6 (1):3-5.
Bujarski, S., J. MacKillop, and L. A. Ray. 2012. Understanding naltrexone mechanism of action and pharmacogenetics in Asian Americans via behavioral economics: a preliminary study. *Exp Clin Psychopharmacol* 20 (3):181-190.
Bulley, A., and M. J. Gullo. 2017. The influence of episodic foresight on delay discounting and demand for alcohol. *Addict Behav* 66:1-6.
Button, K. S., J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, and M. R. Munafo. 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14 (5):365-376.
Carbonaro, T. M., M. P. Bradstreet, F. S. Barrett, K. A. MacLean, R. Jesse, M. W. Johnson, and R. R. Griffiths. 2016. Survey study of challenging experiences after ingesting psilocybin mushrooms: Acute and enduring positive and negative consequences. *J Psychopharmacol* 30 (12):1268-1278.
Cargiulo, T. 2007. Understanding the health impact of alcohol dependence. *Am J Health Syst Pharm* 64 (5 Suppl 3):S5-11.
Carlson, R. G., R. W. Nahhas, S. S. Martins, and R. Daniulaityte. 2016. Predictors of transition to heroin use among initially non-opioid dependent illicit pharmaceutical opioid users: A natural history study. *Drug Alcohol Depend* 160:127-134.
Carroll, K. M., and B. D. Kiluk. 2017. Cognitive behavioral interventions for alcohol and drug use disorders: Through the stage model and back again. *Psychology of Addictive Behaviors* 31 (8):847.
Carroll, K. M., and B. J. Rounsaville. 2010. Computer-assisted therapy in psychiatry: be brave-it's a new world. *Curr Psychiatry Rep* 12 (5):426-432.
Cassidy, R. N., J. W. Tidey, S. M. Colby, V. Long, and S. T. Higgins. 2017. Initial development of an e-cigarette purchase task: A mixed methods study. *Tobacco regulatory science* 3 (2):139-150.
Caulkins, J. P., B. Kilmer, P. H. Reuter, and G. Midgette. 2015a. Cocaine's fall and marijuana's rise: questions and insights based on new estimates of consumption and expenditures in US drug markets. *Addiction* 110 (5):728-736.
Caulkins, J. P., J. Sussell, B. Kilmer, and A. Kasunic. 2015b. How much of the cocaine market are we missing? Insights from respondent-driven sampling in a mid-sized American city. *Drug Alcohol Depend* 147:190-195.
Center for Behavioral Health Statistics. 2017. *2016 National Survey on Drug Use and Health: Detailed tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration.
Center for Behavioral Health Statistics. 2018. *2017 National Survey on Drug Use and Health: Detailed tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Chandler, J., P. Mueller, and G. Paolacci. 2014. Nonnaïveté among Amazon Mechanical Turk workers: Consequences and solutions for behavioral researchers. *Behav Res Methods* 46 (1):112-130.
Chandler, J., G. Paolacci, E. Peer, P. Mueller, and K. A. Ratliff. 2015. Using nonnaive participants can reduce effect sizes. *Psychol Sci* 26 (7):1131-1139.
Chandler, J., and D. Shapiro. 2016. Conducting clinical research using crowdsourced convenience samples. *Annu Rev Clin Psychol* 12:53-81.
Charlton, S. R., and E. Fantino. 2008. Commodity specific rates of temporal discounting: does metabolic function underlie differences in rates of discounting? *Behav Processes* 77 (3):334-342.
Chase, H. W., J. Mackillop, and L. Hogarth. 2013. Isolating behavioural economic indices of demand in relation to nicotine dependence. *Psychopharmacology (Berl)* 226 (2):371-380.
Chavarria, J., N. P. Allan, A. Moltisanti, and J. Taylor. 2015. The effects of present hedonistic time perspective and past negative time perspective on substance use consequences. *Drug Alcohol Depend* 152:39-46.
Chivers, L. L., D. J. Hand, J. S. Priest, and S. T. Higgins. 2016. E-cigarette use among women of reproductive age: Impulsivity, cigarette smoking status, and other risk factors. *Prev Med* 92:126-134.
Chivers, L. L., and S. T. Higgins. 2012. Some observations from behavioral economics for consideration in promoting money management among those with substance use disorders. *Am J Drug Alcohol Abuse* 38 (1):8-19.
Choo, E. K., S. W. F. Ewing, and T. I. Lovejoy. 2016. Opioids out, cannabis in: negotiating the unknowns in patient care for chronic pain. *JAMA* 316 (17):1763-1764.
Collins, R. L., E. T. Morsheimer, S. Shiffman, J. A. Paty, M. Gnys, and G. D. Papandonatos. 1998. Ecological momentary assessment in a behavioral drinking moderation training program. *Exp Clin Psychopharmacol* 6 (3):306-315.
Collins, R. L., P. C. Vincent, J. Yu, L. Liu, and L. H. Epstein. 2014. A behavioral economic approach to assessing demand for marijuana. *Exp Clin Psychopharmacol* 22 (3):211-221.
Compton, W. M., C. M. Jones, and G. T. Baldwin. 2016. Relationship between Nonmedical Prescription-Opioid Use and Heroin Use. *N Engl J Med* 374 (2):154-163.
Conrad, C., H. M. Bradley, D. Broz, S. Buddha, E. L. Chapman, R. R. Galang, D. Hillman, J. Hon, K. W. Hoover, and M. R. Patel. 2015. Community outbreak of HIV infection linked to injection drug use of oxymorphone--Indiana, 2015. *MMWR. Morbidity and mortality weekly report* 64 (16):443-444.
Contractor, A. A., S. B. Frankfurt, N. H. Weiss, and J. D. Elhai. 2017. Latent-level relations between DSM-5 PTSD symptom clusters and problematic smartphone use. *Comput Human Behav* 72:170-177.
Correia, C. J., J. Simons, K. B. Carey, and B. E. Borsari. 1998. Predicting drug use: application of behavioral theories of choice. *Addict Behav* 23 (5):705-709.
Cox, D. J., and J. Dallery. 2016. Effects of delay and probability combinations on discounting in humans. *Behav Processes* 131:15-23.
Crump, M. J., J. V. McDonnell, and T. M. Gureckis. 2013. Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS One* 8 (3):e57410.
Cunningham, J. A., A. Godinho, and V. Kushnir. 2017. Using Mechanical Turk to recruit participants for internet intervention research: experience from recruitment for four trials targeting hazardous alcohol consumption. *BMC Med Res Methodol* 17 (1):156.

Cunningham, J. A., P. L. Selby, K. Kypri, and K. N. Humphreys. 2006. Access to the Internet among drinkers, smokers and illicit drug users: is it a barrier to the provision of interventions on the World Wide Web? *Med Inform Internet Med* 31 (1):53-58.

Cunningham, J. A., T. C. Wild, J. Cordingley, T. van Mierlo, and K. Humphreys. 2009. A randomized controlled trial of an internet-based intervention for alcohol abusers. *Addiction* 104 (12):2023-2032.

Czoty, P. W., W. W. Stoops, and C. R. Rush. 2016. Evaluation of the "Pipeline" for Development of Medications for Cocaine Use Disorder: A Review of Translational Preclinical, Human Laboratory, and Clinical Trial Research. *Pharmacol Rev* 68 (3):533-562.

Dallery, J., A. Kurti, and P. Erb. 2015. A New Frontier: Integrating Behavioral and Digital Technology to Promote Health Behavior. *Behav Anal* 38 (1):19-49.

Daly, T. M., and R. Nataraajan. 2015. Swapping bricks for clicks: crowdsourcing longitudinal data on Amazon Turk. *Journal of Business Research* 68 (12):2603-2609.

De Leeuw, J. R. 2015. jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behav Res Methods* 47 (1):1-12.

Degenhardt, L., W. Hall, and M. Lynskey. 2001. Alcohol, cannabis and tobacco use among Australians: a comparison of their associations with other drug use and use disorders, affective and anxiety disorders, and psychosis. *Addiction* 96 (11):1603-1614.

Del Boca, F. K., and J. Darkes. 2007. Enhancing the validity and utility of randomized clinical trials in addictions treatment research: II. Participant samples and assessment. *Addiction* 102 (8):1194-1203.

Dempster, F. N. 1981. Memory span: Sources of individual and developmental differences. *Psychological Bulletin Psychological Bulletin* 89 (1):63-100.

Dennhardt, A. A., A. M. Yurasek, and J. G. Murphy. 2015. Change in delay discounting and substance reward value following a brief alcohol and drug use intervention. *J Exp Anal Behav* 103 (1):125-140.

Di Lemma, L. C. G., and M. Field. 2017. Cue avoidance training and inhibitory control training for the reduction of alcohol consumption: a comparison of effectiveness and investigation of their mechanisms of action. *Psychopharmacology (Berl)* 234 (16):2489-2498.

Donaldson, C. D., J. T. Siegel, and W. D. Crano. 2016. Nonmedical use of prescription stimulants in college students: Attitudes, intentions, and vested interest. *Addict Behav* 53:101-107.

Dowd, A. N., C. A. Motschman, and S. T. Tiffany. 2018. Development and Validation of the Questionnaire of Vaping Craving. *Nicotine Tob Res*.

Dunn, K. E., F. S. Barrett, E. S. Herrmann, J. G. Plebani, S. C. Sigmon, and M. W. Johnson. 2016a. Behavioral risk assessment for infectious diseases (BRAID): Self-report instrument to assess injection and noninjection risk behaviors in substance users. *Drug Alcohol Depend* 168:69-75.

Dunn, K. E., F. S. Barrett, C. Yepez-Laubach, A. C. Meyer, B. J. Hruska, S. C. Sigmon, M. Fingerhood, and G. E. Bigelow. 2016b. Brief Opioid Overdose Knowledge (BOOK): A questionnaire to assess overdose knowledge in individuals who use illicit or prescribed opioids. *J Addict Med* 10 (5):314-323.

Dunn, K. E., C. Yepez-Laubach, P. A. Nuzzo, M. Fingerhood, A. Kelly, S. Berman, and G. E. Bigelow. 2017. Randomized controlled trial of a computerized opioid overdose education intervention. *Drug Alcohol Depend* 173 Suppl 1:S39-S47.

Durazzo, T. C., S. Gazdzinski, P. Banys, and D. J. Meyerhoff. 2004. Cigarette smoking exacerbates chronic alcohol-induced brain damage: a preliminary metabolite imaging study. *Alcohol Clin Exp Res* 28 (12):1849-1860.
Dworkin, J., H. Hessel, K. Gliske, and J. H. Rudi. 2016. A Comparison of Three Online Recruitment Strategies for Engaging Parents. *Fam Relat* 65 (4):550-561.
Elman, I., S. Krause, H. C. Breiter, R. L. Gollub, J. Heintges, W. A. Baumgartner, B. R. Rosen, and D. R. Gastfriend. 2000. The validity of self-reported drug use in non-treatment seeking individuals with cocaine dependence: correlation with biochemical assays. *Am J Addict* 9 (3):216-221.
Enochson, K., and J. Culbertson. 2015. Collecting psycholinguistic response time data using Amazon mechanical Turk. *PLoS One* 10 (3):e0116946.
Epstein, L. H., J. S. Stein, R. A. Paluch, J. MacKillop, and W. K. Bickel. 2018. Binary components of food reinforcement: Amplitude and persistence. *Appetite* 120:67-74.
Estelles-Arolas, E., and F. Gonzalez-Ladron-De-Guevara. 2012. Towards an integrated crowdsourcing definition. *J. Inf. Sci.* 38 (2):189-200.
Farre, M., R. de la Torre, M. L. Gonzalez, M. T. Teran, P. N. Roset, E. Menoyo, and J. Cami. 1997. Cocaine and alcohol interactions in humans: neuroendocrine effects and cocaethylene metabolism. *J Pharmacol Exp Ther* 283 (1):164-176.
Farris, S. G., E. R. Aston, M. J. Zvolensky, A. M. Abrantes, and J. Metrik. 2017. Psychopathology and tobacco demand. *Drug Alcohol Depend* 177:59-66.
Faugier, J., and M. Sargeant. 1997. Sampling hard to reach populations. *J Adv Nurs* 26 (4):790-797.
Feitosa, J., D. L. Joseph, and D. A. Newman. 2015. Crowdsourcing and personality measurement equivalence: A warning about countries whose primary language is not English. *Personality and Individual Differences* 75:47-52.
Fernie, G., M. Peeters, M. J. Gullo, P. Christiansen, J. C. Cole, H. Sumnall, and M. Field. 2013. Multiple behavioural impulsivity tasks predict prospective alcohol involvement in adolescents. *Addiction* 108 (11):1916-1923.
Few, L. R., J. Acker, C. Murphy, and J. MacKillop. 2012. Temporal stability of a cigarette purchase task. *Nicotine Tob Res* 14 (6):761-765.
Follmer, D. J., R. A. Sperling, and H. K. Suen. 2017. The role of MTurk in education research: Advantages, issues, and future directions. *Educational Researcher* 46 (6):329-334.
Foran, H. M., and K. D. O'Leary. 2008. Alcohol and intimate partner violence: a meta-analytic review. *Clin Psychol Rev* 28 (7):1222-1234.
Forstmann, M., and C. Sagioglou. 2017. Lifetime experience with (classic) psychedelics predicts pro-environmental behavior through an increase in nature relatedness. *J Psychopharmacol* 31 (8):975-988.
Fremouw, W. J., and J. P. Brown, Jr. 1980. The reactivity of addictive behaviors to self-monitoring: a functional analysis. *Addict Behav* 5 (3):209-217.
Gearhardt, A. N., W. R. Corbin, and K. D. Brownell. 2016. Development of the Yale Food Addiction Scale Version 2.0. *Psychol Addict Behav* 30 (1):113-121.
Gentile, N. D., E. H. Librizzi, and M. P. Martinetti. 2012. Academic constraints on alcohol consumption in college students: a behavioral economic analysis. *Exp Clin Psychopharmacol* 20 (5):390-399.
Gleibs, I. H. 2017. Are all “research fields” equal? Rethinking practice for the use of data from crowdsourcing market places. *Behav Res Methods* 49 (4):1333-1342.
Goldstein, R. Z., and N. D. Volkow. 2011. Dysfunction of the prefrontal cortex in addiction: neuroimaging findings and clinical implications. *Nat Rev Neurosci* 12 (11):652-669.

Goodman, J. K., and G. Paolacci. 2017. Crowdsourcing consumer research. *Journal of Consumer Research* 44 (1):196-210.
Grant, B. F., and T. C. Harford. 1995. Comorbidity between DSM-IV alcohol use disorders and major depression: results of a national survey. *Drug Alcohol Depend* 39 (3):197-206.
Gray, J. C., and J. MacKillop. 2014. Interrelationships among individual differences in alcohol demand, impulsivity, and alcohol misuse. *Psychol Addict Behav* 28 (1):282-287.
Green, L., and J. Myerson. 2004. A discounting framework for choice with delayed and probabilistic rewards. *Psychol Bull* 130 (5):769-792.
Green, R., and L. A. Ray. 2018. Effects of varenicline on subjective craving and relative reinforcing value of cigarettes. *Drug Alcohol Depend* 188:53-59.
Greenwald, M. K. 2010. Effects of experimental Unemployment, Employment and Punishment analogs on opioid seeking and consumption in heroin-dependent volunteers. *Drug Alcohol Depend* 111 (1-2):64-73.
Greenwald, M. K., L. H. Lundahl, and C. L. Steinmiller. 2013. Yohimbine increases opioid-seeking behavior in heroin-dependent, buprenorphine-maintained individuals. *Psychopharmacology (Berl)* 225 (4):811-824.
Gul, R. B., and P. A. Ali. 2010. Clinical trials: the challenge of recruitment and retention of participants. *J Clin Nurs* 19 (1-2):227-233.
Hagman, B. T. 2017. Development and psychometric analysis of the Brief DSM-5 Alcohol Use Disorder Diagnostic Assessment: Towards effective diagnosis in college students. *Psychol Addict Behav* 31 (7):797-806.
Hallgren, K. A., B. L. Greenfield, and B. O. Ladd. 2016. Psychometric Properties of the Adolescent Reinforcement Survey Schedule-Alcohol Use Version with College Student Drinkers. *Subst Use Misuse* 51 (7):812-822.
Hamburg, M. A., and F. S. Collins. 2010. The path to personalized medicine. *New England Journal of Medicine* 363 (4):301-304.
Hansen, W. B., L. M. Collins, C. K. Malotte, C. A. Johnson, and J. E. Fielding. 1985. Attrition in prevention research. *J Behav Med* 8 (3):261-275.
Harris, M., and B. Reynolds. 2015. A pilot study of home-based smoking cessation programs for rural, appalachian, pregnant smokers. *J Obstet Gynecol Neonatal Nurs* 44 (2):236-245.
Harrison, L., and A. Hughes. 1997. Introduction--the validity of self-reported drug use: improving the accuracy of survey estimates. *NIDA Res Monogr* 167:1-16.
Hartsell, E. N., and S. D. Neupert. 2017. Chronic and Daily Stressors Along With Negative Affect Interact to Predict Daily Tiredness. *J Appl Gerontol*:733464817741684.
Hauser, D. J., and N. Schwarz. 2015. It's a trap! Instructional manipulation checks prompt systematic thinking on "tricky" tasks. *Sage Open* 5 (2):2158244015584617.
Hauser, D. J., and N. Schwarz. 2016. Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behav Res Methods* 48 (1):400-407.
Heatherton, T. F., L. T. Kozlowski, R. C. Frecker, and K. O. Fagerstrom. 1991. The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br J Addict* 86 (9):1119-1127.
Hedegaard, H., M. Warner, and A. M. Miniño. 2017. Drug overdose deaths in the United States, 1999-2016: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.
Heinz, A. J., E. N. Peters, M. T. Boden, and M. O. Bonn-Miller. 2013. A comprehensive examination of delay discounting in a clinical sample of Cannabis-dependent military

veterans making a self-guided quit attempt. *Exp Clin Psychopharmacol* 21 (1):55-65.
Henrich, J., S. J. Heine, and A. Norenzayan. 2010. Most people are not WEIRD. *Nature* 466 (7302):29.
Herrmann, E. S., P. S. Johnson, and M. W. Johnson. 2015. Examining delay discounting of condom-protected sex among men who have sex with men using crowdsourcing technology. *AIDS Behav* 19 (9):1655-1665.
Herrnstein, R. J. 1961. Relative and absolute strength of response as a function of frequency of reinforcement 1, 2. *J Exp Anal Behav* 4 (3):267-272.
Herrnstein, R. J. 1990. Rational choice theory: Necessary but not sufficient. *American Psychologist* 45 (3):356.
Heyman, G. M. 2009. *Addiction: A disorder of choice*: Harvard University Press.
Heyman, G. M. 2013. Addiction and choice: theory and new data. *Front Psychiatry* 4:31.
Hill, K. P. 2015. Medical marijuana for treatment of chronic pain and other medical and psychiatric problems: a clinical review. *JAMA* 313 (24):2474-2483.
Higgins, S. T., S. H. Heil, S. C. Sigmon, J. W. Tidey, D. E. Gaalema, M. L. Stitzer,... and M. E. Miller. 2017a. Response to varying the nicotine content of cigarettes in vulnerable populations: an initial experimental examination of acute effects. *Psychopharmacology* 234(1): 89-98.
Higgins, S. T., D. D. Reed, R. Redner, J. M. Skelly, I. A. Zvorsky, and A. N. Kurti. 2017b. Simulating demand for cigarettes among pregnant women: A Low-Risk method for studying vulnerable populations. *J Exp Anal Behav* 107(1):176-190.
Ho, Y.-Y., T. N. Vo, H. Chu, X. Luo, and C. T. Le. 2018. A Bayesian hierarchical model for demand curve analysis. *Stat Methods Med Res*:0962280216673675.
Hochster, A., J. Block-Lerner, D. R. Marks, and J. Erblich. 2018. Mindfulness buffers the effects of cue-induced craving on alcohol demand in college drinkers. *Addict Behav* 84:53-56.
Houben, K., R. C. Havermans, C. Nederkoorn, and A. Jansen. 2012. Beer a no-go: learning to stop responding to alcohol cues reduces alcohol intake via reduced affective associations rather than increased response inhibition. *Addiction* 107 (7):1280-1287.
Houben, K., C. Nederkoorn, R. W. Wiers, and A. Jansen. 2011a. Resisting temptation: decreasing alcohol-related affect and drinking behavior by training response inhibition. *Drug Alcohol Depend* 116 (1-3):132-136.
Houben, K., R. W. Wiers, and A. Jansen. 2011b. Getting a grip on drinking behavior: training working memory to reduce alcohol abuse. *Psychol Sci* 22 (7):968-975.
Howard, J., and L. Beckwith. 1996. Issues in subject recruitment and retention with pregnant and parenting substance-abusing women. *NIDA Res Monogr* 166:68-86.
Howe, J. 2006. The rise of crowdsourcing. *Wired magazine* 14 (6):1-4.
Huff, C., and D. Tingley. 2015. "Who are these people?" Evaluating the demographic characteristics and political preferences of MTurk survey respondents. *Research and Politics* 2 (3):e2053168015604648.
Huhn, A. S., A. P. Garcia-Romeu, and K. E. Dunn. 2018. Opioid overdose education for individuals prescribed opioids for pain management: Randomized comparison of two computer-based interventions. *Front Psychiatry* 9:34.
Huhn, A. S., D. A. Tompkins, and K. E. Dunn. 2017. The relationship between treatment accessibility and preference amongst out-of-treatment individuals who engage in non-medical prescription opioid use. *Drug Alcohol Depend* 180:279-285.
Hunsley, J. 1992. Development of the treatment acceptability questionnaire. *Journal of Psychopathology and Behavioral Assessment* 14 (1):55-64.

Hursh, S. R. 1980. Economic concepts for the analysis of behavior. *J Exp Anal Behav* 34 (2):219-238.
Hursh, S. R. 1984. Behavioral economics. *J Exp Anal Behav* 42 (3):435-452.
Hursh, S. R. 1991. Behavioral economics of drug self-administration and drug abuse policy. *J Exp Anal Behav* 56 (2):377-393.
Hursh, S. R., and P. G. Roma. 2013. Behavioral economics and empirical public policy. *J Exp Anal Behav* 99 (1):98-124.
Hursh, S. R., and P. G. Roma. 2016. Behavioral economics and the analysis of consumption and choice. *Managerial and Decision Economics* 37 (4-5):224-238.
Hursh, S. R., and A. Silberberg. 2008. Economic demand and essential value. *Psychological Review* 115 (1):186-198.
Ioannidis, J. P. 2005. Why most published research findings are false. *PLoS Med* 2 (8):e124.
Jacobs, E. A., and W. K. Bickel. 1999. Modeling drug consumption in the clinic using simulation procedures: demand for heroin and cigarettes in opioid-dependent outpatients. *Exp Clin Psychopharmacol* 7 (4):412-426.
Jarmolowicz, D. P., W. K. Bickel, A. E. Carter, C. T. Franck, and E. T. Mueller. 2012. Using crowdsourcing to examine relations between delay and probability discounting. *Behav Processes* 91 (3):308-312.
Jo, C. L., S. D. Golden, S. M. Noar, C. Rini, and K. M. Ribisl. 2018. Effects of e-cigarette advertising messages and cues on cessation outcomes. *Tob Regul Sci* 4 (1):562-572.
Johnson, M. W., and W. K. Bickel. 2006. Replacing relative reinforcing efficacy with behavioral economic demand curves. *J Exp Anal Behav* 85 (1):73-93.
Johnson, M. W., W. K. Bickel, and F. Baker. 2007. Moderate drug use and delay discounting: a comparison of heavy, light, and never smokers. *Exp Clin Psychopharmacol* 15 (2):187-194.
Johnson, M. W., W. K. Bickel, F. Baker, B. A. Moore, G. J. Badger, and A. J. Budney. 2010. Delay discounting in current and former marijuana-dependent individuals. *Exp Clin Psychopharmacol* 18 (1):99-107.
Johnson, M. W., and N. R. Bruner. 2012. The Sexual Discounting Task: HIV risk behavior and the discounting of delayed sexual rewards in cocaine dependence. *Drug Alcohol Depend* 123 (1-3):15-21.
Johnson, M. W., A. Garcia-Romeu, P. S. Johnson, and R. R. Griffiths. 2017a. An online survey of tobacco smoking cessation associated with naturalistic psychedelic use. *J Psychopharmacol* 31 (7):841-850.
Johnson, M. W., P. S. Johnson, O. Rass, and L. R. Pacek. 2017b. Behavioral economic substitutability of e-cigarettes, tobacco cigarettes, and nicotine gum. *J Psychopharmacol* 31 (7):851-860.
Johnson, P. S., E. S. Herrmann, and M. W. Johnson. 2015. Opportunity costs of reward delays and the discounting of hypothetical money and cigarettes. *J Exp Anal Behav* 103 (1):87-107.
Johnson, P. S., and M. W. Johnson. 2014. Investigation of "bath salts" use patterns within an online sample of users in the United States. *J Psychoactive Drugs* 46 (5):369-378.
Jones, A., and M. Field. 2013. The effects of cue-specific inhibition training on alcohol consumption in heavy social drinkers. *Exp Clin Psychopharmacol* 21 (1):8-16.
Kahneman, D., and A. Tversky. 2013. Choices, values, and frames. In *Handbook of the Fundamentals of Financial Decision Making: Part I*: World Scientific, 269-278.

Kampov-Polevoy, A. B., J. C. Garbutt, and D. S. Janowsky. 1999. Association between preference for sweets and excessive alcohol intake: a review of animal and human studies. *Alcohol Alcohol* 34 (3):386-395.
Kampov-Polevoy, A. B., C. Eick, G. Boland, E. Khalitov, and F. T. Crews. 2004. Sweet liking, novelty seeking, and gender predict alcoholic status. *Alcoholism: clinical and experimental research* 28 (9):1291-1298.
Kaplan, B. A., R. N. Foster, D. D. Reed, M. Amlung, J. G. Murphy, and J. MacKillop. 2018. Understanding alcohol motivation using the Alcohol Purchase Task: A methodological systematic review. *Drug Alcohol Depend*.
Kaplan, B. A., and D. D. Reed. 2018. Happy hour drink specials in the Alcohol Purchase Task. *Exp Clin Psychopharmacol* 26 (2):156-167.
Kaplan, B. A., D. D. Reed, J. G. Murphy, A. J. Henley, F. D. D. Reed, P. G. Roma, and S. R. Hursh. 2017. Time constraints in the alcohol purchase task. *Exp Clin Psychopharmacol* 25 (3):186-197.
Kardefelt-Winther, D., A. Heeren, A. Schimmenti, A. van Rooij, P. Maurage, M. Carras, J. Edman, A. Blaszczynski, Y. Khazaal, and J. Billieux. 2017. How can we conceptualize behavioural addiction without pathologizing common behaviours? *Addiction* 112 (10):1709-1715.
Keith, M. G., L. Tay, and P. D. Harms. 2017. Systems Perspective of Amazon Mechanical Turk for organizational research: Review and recommendations. *Front Psychol* 8:1359.
Kessler, R. C., C. B. Nelson, K. A. McGonagle, M. J. Edlund, R. G. Frank, and P. J. Leaf. 1996. The epidemiology of co-occurring addictive and mental disorders: implications for prevention and service utilization. *Am J Orthopsychiatry* 66 (1):17-31.
Khatib, F., S. Cooper, M. D. Tyka, K. Xu, I. Makedon, Z. Popovic, D. Baker, and F. Players. 2011. Algorithm discovery by protein folding game players. *Proc Natl Acad Sci U S A* 108 (47):18949-18953.
Kiluk, B. D., J. A. Dreifuss, R. D. Weiss, J. Morgenstern, and K. M. Carroll. 2013. The Short Inventory of Problems - revised (SIP-R): psychometric properties within a large, diverse sample of substance use disorder treatment seekers. *Psychol Addict Behav* 27 (1):307-314.
Kim, H. S., and D. C. Hodgins. 2017. Reliability and validity of data obtained from alcohol, cannabis, and gambling populations on Amazon's Mechanical Turk. *Psychol Addict Behav* 31 (1):85-94.
Kirby, K. N., and N. M. Petry. 2004. Heroin and cocaine abusers have higher discount rates for delayed rewards than alcoholics or non-drug-using controls. *Addiction* 99 (4):461-471.
Kirby, K. N., N. M. Petry, and W. K. Bickel. 1999. Heroin addicts have higher discount rates for delayed rewards than non-drug-using controls. *Journal of Experimental psychology: general* 128 (1):78.
Kiselica, A. M., T. A. Webber, and M. A. Bornovalova. 2016. Validity of the alcohol purchase task: a meta-analysis. *Addiction* 111 (5):806-816.
Koffarnus, M. N., and W. K. Bickel. 2014. A 5-trial adjusting delay discounting task: accurate discount rates in less than one minute. *Exp Clin Psychopharmacol* 22 (3):222-228.
Koffarnus, M. N., C. T. Franck, J. S. Stein, and W. K. Bickel. 2015. A modified exponential behavioral economic demand model to better describe consumption data. *Exp Clin Psychopharmacol* 23 (6):504-512.

Koffarnus, M. N., D. P. Jarmolowicz, E. T. Mueller, and W. K. Bickel. 2013. Changing delay discounting in the light of the competing neurobehavioral decision systems theory: a review. *J Exp Anal Behav* 99 (1):32-57.
Koffarnus, M. N., and B. A. Kaplan. 2017. Clinical models of decision making in addiction. *Pharmacol Biochem Behav*.
Kokkevi, A., C. Richardson, B. Palermou, and V. Leventakou. 1997. Reliability of drug dependents' self-reports. *Drug Alcohol Depend* 45 (1-2):55-61.
Kollins, S. H. 2003. Delay discounting is associated with substance use in college students. *Addict Behav* 28 (6):1167-1173.
Koob, G. F., M. Rocio, A. Carrera, L. H. Gold, C. J. Heyser, C. Maldonado-Irizarry, A. Markou, L. H. Parsons, A. J. Roberts, and G. Schulteis. 1998. Substance dependence as a compulsive behavior. *J Psychopharmacol* 12 (1):39-48.
Kozlowski, L. T., R. E. Mann, D. Wilkinson, and C. X. Poulos. 1989. “Cravings” are amibiguous: Ask about urges or desires. *Addict Behav* 14 (4):443-445.
Kranzler, H. R., K. A. Sandstrom, and J. Van Kirk. 2001. Sweet taste preference as a risk factor for alcohol dependence. *American Journal of Psychiatry* 158 (5):813-815.
Kranzler, H. R., and M. Soyka. 2018. Diagnosis and pharmacotherapy of alcohol use disorder: a review. *JAMA* 320 (8):815-824.
Kranzler, H. R., H. Tennen, S. Armeli, G. Chan, J. Covault, A. Arias, and C. Oncken. 2009. Targeted naltrexone for problem drinkers. *J Clin Psychopharmacol* 29 (4):350.
Krishnan-Sarin, S., B. Reynolds, A. M. Duhig, A. Smith, T. Liss, A. McFetridge, D. A. Cavallo, K. M. Carroll, and M. N. Potenza. 2007. Behavioral impulsivity predicts treatment outcome in a smoking cessation program for adolescent smokers. *Drug Alcohol Depend* 88 (1):79-82.
Kuerbis, A., F. J. Muench, R. Lee, J. Pena, and L. Hail. 2016. An exploratory pilot study of mechanisms of action within normative feedback for adult drinkers. *PeerJ* 4:e2114.
Kuerbis, A. N., L. Hail, A. A. Moore, and F. J. Muench. 2017. A pilot study of online feedback for adult drinkers 50 and older: Feasibility, efficacy, and preferences for intervention. *J Subst Abuse Treat* 77:126-132.
Kufahl, P. R., L. R. Watterson, and M. F. Olive. 2014. The development of acamprosate as a treatment against alcohol relapse. *Expert Opin Drug Discov* 9 (11):1355-1369.
Kurti, A. N., D. R. Davis, R. Redner, B. P. Jarvis, I. Zvorsky, D. R. Keith, H. A. Bolivar, T. J. White, P. Rippberger, C. Markesich, G. Atwood, and S. T. Higgins. 2016. A Review of the Literature on Remote Monitoring Technology in Incentive-Based Interventions for Health-Related Behavior Change. *Transl Issues Psychol Sci* 2 (2):128-152.
Kushnir, V., and J. A. Cunningham. 2014. Event-specific drinking in the general population. *J Stud Alcohol Drugs* 75 (6):968-972.
Lac, A., and D. E. Berger. 2013. Development and validation of the Alcohol Myopia Scale. *Psychol Assess* 25 (3):738-747.
Lakens, D. 2013. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology* 4:863.
Lamb, R. J., and B. C. Ginsburg. 2017. Addiction as a BAD, a Behavioral Allocation Disorder. *Pharmacol Biochem Behav*.
Lanaj, K., R. E. Johnson, and C. M. Barnes. 2014. Beginning the workday yet already depleted? Consequences of late-night smartphone use and sleep. *Organizational Behavior and Human Decision Processes* 124 (1):11-23.
Landers, R. N., and T. S. Behrend. 2015. An inconvenient truth: Arbitrary distinctions between organizational, Mechanical Turk, and other convenience samples. *Industrial and Organizational Psychology* 8 (2):142-164.

Lau-Barraco, C., A. L. Braitman, A. N. Linden-Carmichael, and A. L. Stamates. 2016. Differences in weekday versus weekend drinking among nonstudent emerging adults. *Exp Clin Psychopharmacol* 24 (2):100-109.
Lazard, A. J., A. Schmidt, H. Vu, M. J. Byron, E. Peters, M. H. Boynton, and N. T. Brewer. 2017. Icons for health effects of cigarette smoke: a test of semiotic type. *J Behav Med* 40 (4):641-650.
Leas, E. C., J. P. Pierce, C. V. Dimofte, A. Villasenor, and D. R. Strong. 2017. US adult smokers' perceptions' of Australia's cigarette warning labels: variance by warning content and consistency across sociodemographic subsegments. *Tob Control* 26 (4):485-486.
Lee, D. C., B. S. Crosier, J. T. Borodovsky, J. D. Sargent, and A. J. Budney. 2016. Online survey characterizing vaporizer use among cannabis users. *Drug Alcohol Depend* 159:227-233.
Lee, Y. J., J. A. Arida, and H. S. Donovan. 2017. The application of crowdsourcing approaches to cancer research: a systematic review. *Cancer Med* 6 (11):2595-2605.
Levy, M. S. 2008. Listening to our clients: the prevention of relapse. *J Psychoactive Drugs* 40 (2):167-172.
Liao, W., X. Luo, C. T. Le, H. Chu, L. H. Epstein, J. Yu, J. S. Ahluwalia, and J. L. Thomas. 2013. Analysis of cigarette purchase task instrument data with a left-censored mixed effects model. *Exp Clin Psychopharmacol* 21 (2):124.
Litman, L., J. Robinson, and T. Abberbock. 2017. TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behav Res Methods* 49 (2):433-442.
Litt, M. D., N. L. Cooney, and P. Morse. 1998. Ecological momentary assessment (EMA) with treated alcoholics: methodological problems and potential solutions. *Health Psychol* 17 (1):48-52.
Litten, R. Z., B. B. Wilford, D. E. Falk, M. L. Ryan, and J. B. Fertig. 2016. Potential medications for the treatment of alcohol use disorder: An evaluation of clinical efficacy and safety. *Subst Abus* 37 (2):286-298.
Lucas, P. 2012. Cannabis as an adjunct to or substitute for opiates in the treatment of chronic pain. *J Psychoactive Drugs* 44 (2):125-133.
Mack, K. A., C. M. Jones, and M. F. Ballesteros. 2017. Illicit drug use, illicit drug use disorders, and drug overdose deaths in metropolitan and nonmetropolitan areas—United States. *American journal of transplantation* 17 (12):3241-3252.
MacKillop, J. 2016. The behavioral economics and neuroeconomics of alcohol use disorders. *Alcohol Clin Exp Res* 40 (4):672-685.
MacKillop, J., M. T. Amlung, L. R. Few, L. A. Ray, L. H. Sweet, and M. R. Munafo. 2011. Delayed reward discounting and addictive behavior: a meta-analysis. *Psychopharmacology (Berl)* 216 (3):305-321.
MacKillop, J., C. L. Brown, M. K. Stojek, C. M. Murphy, L. Sweet, and R. S. Niaura. 2012. Behavioral economic analysis of withdrawal- and cue-elicited craving for tobacco: an initial investigation. *Nicotine Tob Res* 14 (12):1426-1434.
MacKillop, J., and C. W. Kahler. 2009. Delayed reward discounting predicts treatment response for heavy drinkers receiving smoking cessation treatment. *Drug Alcohol Depend* 104 (3):197-203.
MacKillop, J., R. Miranda, Jr., P. M. Monti, L. A. Ray, J. G. Murphy, D. J. Rohsenow, J. E. McGeary, R. M. Swift, J. W. Tidey, and C. J. Gwaltney. 2010a. Alcohol demand, delayed reward discounting, and craving in relation to drinking and alcohol use disorders. *J Abnorm Psychol* 119 (1):106-114.
MacKillop, J., and J. G. Murphy. 2007. A behavioral economic measure of demand for alcohol predicts brief intervention outcomes. *Drug Alcohol Depend* 89 (2-3):227-233.

MacKillop, J., J. G. Murphy, L. A. Ray, D. T. Eisenberg, S. A. Lisman, J. K. Lum, and D. S. Wilson. 2008. Further validation of a cigarette purchase task for assessing the relative reinforcing efficacy of nicotine in college smokers. *Exp Clin Psychopharmacol* 16 (1):57-65.

Mackillop, J., J. G. Murphy, J. W. Tidey, C. W. Kahler, L. A. Ray, and W. K. Bickel. 2009. Latent structure of facets of alcohol reinforcement from a behavioral economic demand curve. *Psychopharmacology (Berl)* 203 (1):33-40.

MacKillop, J., S. O'Hagen, S. A. Lisman, J. G. Murphy, L. A. Ray, J. W. Tidey, J. E. McGeary, and P. M. Monti. 2010b. Behavioral economic analysis of cue-elicited craving for alcohol. *Addiction* 105 (9):1599-1607.

Madden, G. J., and D. Kalman. 2010. Effects of bupropion on simulated demand for cigarettes and the subjective effects of smoking. *Nicotine Tob Res* 12 (4):416-422.

Madras, B. K., W. M. Compton, D. Avula, T. Stegbauer, J. B. Stein, and H. W. Clark. 2009. Screening, brief interventions, referral to treatment (SBIRT) for illicit drug and alcohol use at multiple healthcare sites: comparison at intake and 6 months later. *Drug Alcohol Depend* 99 (1-3):280-295.

Maisel, N. C., J. C. Blodgett, P. L. Wilbourne, K. Humphreys, and J. W. Finney. 2013. Meta-analysis of naltrexone and acamprosate for treating alcohol use disorders: when are these medications most helpful? *Addiction* 108 (2):275-293.

Mason, W., and S. Suri. 2012. Conducting behavioral research on Amazon's Mechanical Turk. *Behav Res Methods* 44 (1):1-23.

Matusiewicz, A. K., A. E. Carter, R. D. Landes, and R. Yi. 2013. Statistical equivalence and test-retest reliability of delay and probability discounting using real and hypothetical rewards. *Behav Processes* 100:116-122.

McClure, E. A., S. P. Acquavita, E. Harding, and M. L. Stitzer. 2013a. Utilization of communication technology by patients enrolled in substance abuse treatment. *Drug Alcohol Depend* 129 (1-2):145-150.

McClure, E. A., R. G. Vandrey, M. W. Johnson, and M. L. Stitzer. 2013b. Effects of varenicline on abstinence and smoking reward following a programmed lapse. *Nicotine Tob Res* 15 (1):139-148.

McClure, S. M., and W. K. Bickel. 2014. A dual-systems perspective on addiction: contributions from neuroimaging and cognitive training. *Ann N Y Acad Sci* 1327:62-78.

McGregor, T. J., and J. C. Batis. 2016. A novel method for assessing caffeine dependence. *Journal of Caffeine Research* 6 (1):26-33.

McKay, J. R., A. I. Alterman, M. J. Rutherford, J. S. Cacciola, and A. T. McLellan. 1999. The relationship of alcohol use to cocaine relapse in cocaine dependent patients in an aftercare study. *J Stud Alcohol* 60 (2):176-180.

McKerchar, T. L., and J. E. Mazur. 2016. Human choices between variable and fixed rewards in hypothetical variable-delay and double-reward discounting procedures. *J Exp Anal Behav* 106 (1):1-21.

Melby-Lervag, M., and C. Hulme. 2013. Is working memory training effective? A meta-analytic review. *Dev Psychol* 49 (2):270-291.

Mellis, A. M., S. E. Snider, and W. K. Bickel. 2018. Narrative theory: II. Self-generated and experimenter-provided negative income shock narratives increase delay discounting. *Exp Clin Psychopharmacol*.

Mellis, A. M., A. E. Woodford, J. S. Stein, and W. K. Bickel. 2017. A second type of magnitude effect: Reinforcer magnitude differentiates delay discounting between substance users and controls. *J Exp Anal Behav* 107 (1):151-160.

Mendoza, T., T. Mayne, D. Rublee, and C. Cleeland. 2006. Reliability and validity of a modified Brief Pain Inventory short form in patients with osteoarthritis. *European Journal of Pain* 10 (4):353-353.
Meneses-Gaya, C. d., A. W. Zuardi, S. R. Loureiro, and J. A. S. Crippa. 2009. Alcohol Use Disorders Identification Test (AUDIT): an updated systematic review of psychometric properties. *Psychology & Neuroscience* 2 (1):83-97.
Metrik, J., E. R. Aston, C. W. Kahler, D. J. Rohsenow, J. E. McGeary, V. S. Knopik, and J. MacKillop. 2016. Cue-elicited increases in incentive salience for marijuana: Craving, demand, and attentional bias. *Drug Alcohol Depend* 167:82-88.
Metrik, J., C. W. Kahler, B. Reynolds, J. E. McGeary, P. M. Monti, M. Haney, H. de Wit, and D. J. Rohsenow. 2012. Balanced placebo design with marijuana: pharmacological and expectancy effects on impulsivity and risk taking. *Psychopharmacology (Berl)* 223 (4):489-499.
Metrik, J., D. J. Rohsenow, P. M. Monti, J. McGeary, T. A. Cook, H. de Wit, M. Haney, and C. W. Kahler. 2009. Effectiveness of a marijuana expectancy manipulation: Piloting the balanced-placebo design for marijuana. *Exp Clin Psychopharmacol* 17 (4):217-225.
Miller, J., R. Schaffer, and S. A. Hackley. 1991. Effects of preliminary information in a Go versus No-go task. *Acta Psychol (Amst)* 76 (3):241-292.
Miller, J. D., M. Crowe, B. Weiss, J. L. Maples-Keller, and D. R. Lynam. 2017. Using online, crowdsourcing platforms for data collection in personality disorder research: The example of Amazon's Mechanical Turk. *Personal Disord* 8 (1):26-34.
Morean, M. E., N. Lipshie, M. Josephson, and D. Foster. 2017. Predictors of adult e-cigarette users vaporizing cannabis using e-cigarettes and vape-pens. *Subst Use Misuse* 52 (8):974-981.
Morris, V., M. Amlung, B. A. Kaplan, D. D. Reed, T. Petker, and J. MacKillop. 2017. Using crowdsourcing to examine behavioral economic measures of alcohol value and proportionate alcohol reinforcement. *Exp Clin Psychopharmacol* 25 (4):314-321.
Morrison, P., and P. Morrison. 1997. Illusions. *Scientific American* 277:99-101.
Moyer, A., J. W. Finney, C. E. Swearingen, and P. Vergun. 2002. Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. *Addiction* 97 (3):279-292.
Muench, F., M. Hayes, A. Kuerbis, and S. Shao. 2015. The independent relationship between trouble controlling Facebook use, time spent on the site and distress. *J Behav Addict* 4 (3):163-169.
Mullinix, K. J., T. J. Leeper, J. N. Druckman, and J. Freese. 2015. The generalizability of survey experiments. *Journal of Experimental Political Science* 2 (2):109-138.
Murphy, C. M., J. MacKillop, R. A. Martin, J. W. Tidey, S. M. Colby, and D. J. Rohsenow. 2017. Effects of varenicline versus transdermal nicotine replacement therapy on cigarette demand on quit day in individuals with substance use disorders. *Psychopharmacology (Berl)* 234 (16):2443-2452.
Murphy, C. M., M. M. Owens, L. H. Sweet, and J. MacKillop. 2016. The substitutability of cigarettes and food: A behavioral economic comparison in normal weight and overweight or obese smokers. *Psychol Addict Behav* 30 (8):857-867.
Murphy, J. G., C. J. Correia, S. M. Colby, and R. E. Vuchinich. 2005. Using behavioral theories of choice to predict drinking outcomes following a brief intervention. *Exp Clin Psychopharmacol* 13 (2):93-101.
Murphy, J. G., A. A. Dennhardt, J. R. Skidmore, B. Borsari, N. P. Barnett, S. M. Colby, and M. P. Martens. 2012a. A randomized controlled trial of a behavioral economic supplement to brief motivational interventions for college drinking. *J Consult Clin Psychol* 80 (5):876.

Murphy, J. G., A. A. Dennhardt, A. M. Yurasek, J. R. Skidmore, M. P. Martens, J. MacKillop, and M. E. McDevitt-Murphy. 2015. Behavioral economic predictors of brief alcohol intervention outcomes. *J Consult Clin Psychol* 83 (6):1033-1043.

Murphy, J. G., and J. MacKillop. 2006. Relative reinforcing efficacy of alcohol among college student drinkers. *Exp Clin Psychopharmacol* 14 (2):219-227.

Murphy, J. G., J. MacKillop, J. R. Skidmore, and A. A. Pederson. 2009. Reliability and validity of a demand curve measure of alcohol reinforcement. *Exp Clin Psychopharmacol* 17 (6):396-404.

Murphy, J. G., J. MacKillop, J. W. Tidey, L. A. Brazil, and S. M. Colby. 2011. Validity of a demand curve measure of nicotine reinforcement with adolescent smokers. *Drug Alcohol Depend* 113 (2-3):207-214.

Murphy, J. G., J. R. Skidmore, A. A. Dennhardt, M. P. Martens, B. Borsari, N. P. Barnett, and S. M. Colby. 2012b. A behavioral economic supplement to brief motivational interventions for college drinking. *Addiction Research & Theory* 20 (6):456-465.

Murphy, J. G., A. M. Yurasek, L. Z. Meshesha, A. A. Dennhardt, J. Mackillop, J. R. Skidmore, and M. P. Martens. 2014. Family history of problem drinking is associated with less sensitivity of alcohol demand to a next-day responsibility. *J Stud Alcohol Drugs* 75 (4):653-663.

Napper, L. E., D. G. Fisher, M. E. Johnson, and M. M. Wood. 2010. The reliability and validity of drug users' self reports of amphetamine use among primarily heroin and cocaine users. *Addict Behav* 35 (4):350-354.

National Drug Intelligence Center. 2011. The economic impact of illicit drug use on American society: United States Department of Justice Washington, DC.

National Institute on Alcohol Abuse Alcoholism. 2007. *Helping Patients who Drink Too Much: A Clinician's Guide: Updated 2005 Edition*: US Department of Health and Human Services, National Institutes of Health, National Institute on Alcohol Abuse and Alcoholism.

Necka, E. A., S. Cacioppo, G. J. Norman, and J. T. Cacioppo. 2016. Measuring the prevalence of problematic respondent nehaviors among MTurk, campus, and community participants. *PLoS One* 11 (6):e0157732.

Nosek, B. A., G. Alter, G. C. Banks, D. Borsboom, S. D. Bowman, S. J. Breckler, S. Buck, C. D. Chambers, G. Chin, G. Christensen, M. Contestabile, A. Dafoe, E. Eich, J. Freese, R. Glennerster, D. Goroff, D. P. Green, B. Hesse, M. Humphreys, J. Ishiyama, D. Karlan, A. Kraut, A. Lupia, P. Mabry, T. A. Madon, N. Malhotra, E. Mayo-Wilson, M. McNutt, E. Miguel, E. L. Paluck, U. Simonsohn, C. Soderberg, B. A. Spellman, J. Turitto, G. VandenBos, S. Vazire, E. J. Wagenmakers, R. Wilson, and T. Yarkoni. 2015. SCIENTIFIC STANDARDS. Promoting an open research culture. *Science* 348 (6242):1422-1425.

Noyes, E. T., and R. C. Schlauch. 2018. Examination of approach and avoidance inclinations on the reinforcing value of alcohol. *Addict Behav* 79:61-67.

Odum, A. L. 2011. Delay discounting: I'm a k, you're a k. *J Exp Anal Behav* 96 (3):427-439.

Ohmura, Y., T. Takahashi, N. Kitamura, and P. Wehr. 2006. Three-month stability of delay and probability discounting measures. *Exp Clin Psychopharmacol* 14 (3):318-328.

Oleson, E. B., J. M. Richardson, and D. C. Roberts. 2011. A novel IV cocaine self-administration procedure in rats: differential effects of dopamine, serotonin, and GABA drug pre-treatments on cocaine consumption and maximal price paid. *Psychopharmacology (Berl)* 214 (2):567-577.

Owens, M. M., L. A. Ray, and J. MacKillop. 2015. Behavioral economic analysis of stress effects on acute motivation for alcohol. *J Exp Anal Behav* 103 (1):77-86.

Pacek, L. R., O. Rass, and M. W. Johnson. 2017a. Knowledge about nicotine among HIV-positive smokers: Implications for tobacco regulatory science policy. *Addict Behav* 65:81-86.
Pacek, L. R., O. Rass, and M. W. Johnson. 2017b. Positive smoking cessation-related interactions with HIV care providers increase the likelihood of interest in cessation among HIV-positive cigarette smokers. *AIDS Care* 29 (10):1309-1314.
Paolacci, G., and J. Chandler. 2014. Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science* 23 (3):184-188.
Pearson, J. L., A. Richardson, S. P. Feirman, A. C. Villanti, J. Cantrell, A. Cohn, M. Tacelosky, and T. R. Kirchner. 2016. American Spirit pack descriptors and perceptions of harm: A crowdsourced comparison of modified packs. *Nicotine Tob Res* 18 (8):1749-1756.
Peer, E., J. Vosgerau, and A. Acquisti. 2014. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behav Res Methods* 46 (4):1023-1031.
Peitz, G. W., J. C. Strickland, E. G. Pitts, M. Foley, S. Tonidandel, and M. A. Smith. 2013. Peer influences on drug self-administration: an econometric analysis in socially housed rats. *Behav Pharmacol* 24 (2):114-123.
Perry, J. L., and M. E. Carrol. 2008. The role of impulsive behavior in drug abuse. *Psychopharmacology (Berl)* 200 (1):1-26.
Peters, E. N., N. M. Petry, D. M. Lapaglia, B. Reynolds, and K. M. Carroll. 2013. Delay discounting in adults receiving treatment for marijuana dependence. *Exp Clin Psychopharmacol* 21 (1):46-54.
Peters, E. N., Z. R. Rosenberry, G. L. Schauer, K. E. O'Grady, and P. S. Johnson. 2017. Marijuana and tobacco cigarettes: Estimating their behavioral economic relationship using purchasing tasks. *Exp Clin Psychopharmacol* 25 (3):208-215.
Petry, N. M. 2001. Delay discounting of money and alcohol in actively using alcoholics, currently abstinent alcoholics, and controls. *Psychopharmacology (Berl)* 154 (3):243-250.
Petry, N. M., and W. K. Bickel. 1998. Polydrug abuse in heroin addicts: a behavioral economic analysis. *Addiction* 93 (3):321-335.
Pickover, A. M., B. G. Messina, C. J. Correia, K. B. Garza, and J. G. Murphy. 2016. A behavioral economic analysis of the nonmedical use of prescription drugs among young adults. *Exp Clin Psychopharmacol* 24 (1):38-47.
Pilowsky, D. J., and L. T. Wu. 2013. Screening instruments for substance use and brief interventions targeting adolescents in primary care: a literature review. *Addict Behav* 38 (5):2146-2153.
Pinheiro, J., D. Bates, S. DebRoy, D. Sarkar, and R Core Team. R package version 3.1-137. *nlme: Linear and nonlinear mixed effects models* 2018 [cited R package version 3.1-137]. Available from https://cran.r-project.org/package=nlme.
Pitts, R. C. 2014. Reconsidering the concept of behavioral mechanisms of drug action. *J Exp Anal Behav* 101 (3):422-441.
Porter-Stransky, K. A., B. S. Bentzley, and G. Aston-Jones. 2015. Individual differences in orexin-I receptor modulation of motivation for the opioid remifentanil. *Addiction Biology*.
Powell, R. A., P. L. Honey, and D. G. Symbaluk. 2016. *Introduction to learning and behavior*: Cengage Learning.
Priester, M. A., T. Browne, A. Iachini, S. Clone, D. DeHart, and K. D. Seay. 2016. Treatment access barriers and disparities among individuals with co-occurring mental health and substance use disorders: an integrative literature review. *J Subst Abuse Treat* 61:47-59.

Quigley, B. M., and K. E. Leonard. 2000. Alcohol and the continuation of early marital aggression. *Alcohol Clin Exp Res* 24 (7):1003-1010.
Quisenberry, A. J., M. N. Koffarnus, L. H. Epstein, and W. K. Bickel. 2017. The Experimental Tobacco Marketplace II: Substitutability and sex effects in dual electronic cigarette and conventional cigarette users. *Drug Alcohol Depend* 178:551-555.
Rachlin, H. 1974. Self-control. *Behaviorism* 2 (1):94-107.
Rachlin, H. 1980. Economics and behavioral psychology. *Limits to action: The allocation of individual behavior*:205-236.
Rachlin, H., 1992. Diminishing marginal value as delay discounting. *J Exp Anal Behav 57*(3):407-415.
Rachlin, H. 2006. Notes on discounting. *J Exp Anal Behav* 85 (3):425-435.
Rachlin, H., and L. Green. 1972. Commitment, choice and self-control. *J Exp Anal Behav* 17 (1):15-22.
Rachlin, H., L. Green, J. H. Kagel, and R. C. Battalio. 1976. Economic demand theory and psychological studies of choice. In *Psychology of Learning and Motivation*: Elsevier, 129-154.
Raiff, B. R., B. P. Jarvis, M. Turturici, and J. Dallery. 2013. Acceptability of an Internet-based contingency management intervention for smoking cessation: views of smokers, nonsmokers, and healthcare professionals. *Exp Clin Psychopharmacol* 21 (3):204-213.
Raineri, A. and H. Rachlin. 1993. The effect of temporal constraints on the value of money and other commodities. *J Behav Decis Mak* 6(2):77-94.
Rainey, J. C., C. R. Furman, and A. N. Gearhardt. 2018. Food addiction among sexual minorities. *Appetite* 120:16-22.
Raleigh, M. D., P. R. Pentel, and M. G. LeSage. 2014. Pharmacokinetic correlates of the effects of a heroin vaccine on heroin self-administration in rats. *PLoS One* 9 (12):e115696.
Ramo, D. E., and S. A. Brown. 2008. Classes of substance abuse relapse situations: a comparison of adolescents and adults. *Psychol Addict Behav* 22 (3):372-379.
Ramo, D. E., M. Kaur, E. S. Corpuz, D. D. Satre, K. Delucchi, S. A. Brown, and J. J. Prochaska. 2018. Using Facebook to address smoking and heavy drinking in young adults: Protocol for a randomized, controlled trial. *Contemp Clin Trials* 68:52-60.
Ramo, D. E., and J. J. Prochaska. 2012. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. *J Med Internet Res* 14 (1).
Rand, D. G. 2012. The promise of Mechanical Turk: how online labor markets can help theorists run behavioral experiments. *J Theor Biol* 299:172-179.
Rasmussen, E. B., S. R. Lawyer, and W. Reilly. 2010. Percent body fat is related to delay and probability discounting for food in humans. *Behav Processes* 83 (1):23-30.
Rass, O., L. R. Pacek, P. S. Johnson, and M. W. Johnson. 2015a. Characterizing use patterns and perceptions of relative harm in dual users of electronic and tobacco cigarettes. *Exp Clin Psychopharmacol* 23 (6):494-503.
Rass, O., R. L. Schacht, K. Buckheit, M. W. Johnson, E. C. Strain, and M. Z. Mintzer. 2015b. A randomized controlled trial of the effects of working memory training in methadone maintenance patients. *Drug Alcohol Depend* 156:38-46.
Reed, D. D., A. Becirevic, P. Atchley, B. A. Kaplan, and B. S. Liese. 2016. Validation of a novel delay discounting of text messaging questionnaire. *The Psychological Record* 66 (2):253-261.

Reed, D. D., C. R. Niileksela, and B. A. Kaplan. 2013. Behavioral economics: a tutorial for behavior analysts in practice. *Behav Anal Pract* 6 (1):34-54.
Regier, D. A., M. E. Farmer, D. S. Rae, B. Z. Locke, S. J. Keith, L. L. Judd, and F. K. Goodwin. 1990. Comorbidity of mental disorders with alcohol and other drug abuse. Results from the Epidemiologic Catchment Area (ECA) Study. *JAMA* 264 (19):2511-2518.
Rehm, J. 2011. The risks associated with alcohol use and alcoholism. *Alcohol Res Health* 34 (2):135-143.
Reinert, D. F., and J. P. Allen. 2007. The alcohol use disorders identification test: an update of research findings. *Alcohol Clin Exp Res* 31 (2):185-199.
Reynolds, B., M. Harris, S. A. Slone, B. J. Shelton, J. Dallery, W. Stoops, and R. Lewis. 2015. A feasibility study of home-based contingency management with adolescent smokers of rural Appalachia. *Exp Clin Psychopharmacol* 23 (6):486-493.
Robinson, T. E., and K. C. Berridge. 1993. The neural basis of drug craving: an incentive-sensitization theory of addiction. *Brain Res Brain Res Rev* 18 (3):247-291.
Roma, P. G., S. R. Hursh, and S. Hudja. 2016. Hypothetical purchase task questionnaires for behavioral economic assessments of value and motivation. *Managerial and Decision Economics* 37 (4-5):306-323.
Rosner, S., A. Hackl-Herrwerth, S. Leucht, S. Vecchi, M. Srisurapanont, and M. Soyka. 2010. Opioid antagonists for alcohol dependence. *Cochrane Database Syst Rev* (12):CD001867.
Rudski, J. 2016. Public perspectives on expanding naloxone access to reverse opioid overdoses. *Subst Use Misuse* 51 (13):1771-1780.
Sacco, P., C. A. Smith, D. Harrington, D. V. Svoboda, and B. Resnick. 2016. Feasibility and Utility of Experience Sampling to Assess Alcohol Consumption Among Older Adults. *J Appl Gerontol* 35 (1):106-120.
Sacks, J. J., K. R. Gonzales, E. E. Bouchery, L. E. Tomedi, and R. D. Brewer. 2015. 2010 National and State Costs of Excessive Alcohol Consumption. *Am J Prev Med* 49 (5):e73-79.
Saunders, J. B., O. G. Aasland, T. F. Babor, J. R. de la Fuente, and M. Grant. 1993. Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption--II. *Addiction* 88 (6):791-804.
Sayette, M. A., S. Shiffman, S. T. Tiffany, R. S. Niaura, C. S. Martin, and W. G. Schadel. 2000. The measurement of drug craving. *Addiction* 95 (8s2):189-210.
Schlienz, N. J., L. W. Hawk, Jr., S. T. Tiffany, R. J. O'Connor, and M. C. Mahoney. 2014. The impact of pre-cessation varenicline on behavioral economic indices of smoking reinforcement. *Addict Behav* 39 (10):1484-1490.
Schranz, A. J., J. Barrett, C. B. Hurt, C. Malvestutto, and W. C. Miller. 2018. Challenges facing a rural opioid epidemic: Treatment and prevention of HIV and Hepatitis C. *Current HIV/AIDS Reports* 15 (3):245-254.
Schulte, E. M., and A. N. Gearhardt. 2017. Development of the Modified Yale Food Addiction Scale Version 2.0. *Eur Eat Disord Rev* 25 (4):302-308.
Schulte, M. H. J., R. W. Wiers, W. J. Boendermaker, A. E. Goudriaan, W. van den Brink, D. S. van Deursen, M. Friese, E. Brede, and A. J. Waters. 2018. The effect of N-acetylcysteine and working memory training on cocaine use, craving and inhibition in regular cocaine users: correspondence of lab assessments and Ecological Momentary Assessment. *Addict Behav* 79:24-31.
Seithe, M., J. Morina, and A. Glockner. 2016. Bonn eXperimental System (BoXS): An open-source platform for interactive experiments in psychology and economics. *Behav Res Methods* 48 (4):1454-1475.

Shapiro, D. N., J. Chandler, and P. A. Mueller. 2013. Using Mechanical Turk to study clinical populations. *Clinical Psychological Science* 1 (2):213-220.
Sheehan, D. V., Y. Lecrubier, K. H. Sheehan, P. Amorim, J. Janavs, E. Weiller, T. Hergueta, R. Baker, and G. C. Dunbar. 1998. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 59 Suppl 20:22-33;quiz 34-57.
Shi, Z., A. L. Wang, L. F. Emery, K. M. Sheerin, and D. Romer. 2017. The importance of relevant emotional arousal in the efficacy of pictorial health warnings for cigarettes. *Nicotine Tob Res* 19 (6):750-755.
Simpson, C. A., and R. E. Vuchinich. 2000. Reliability of a measure of temporal discounting. *The Psychological Record* 50 (1):3-16.
Singer, J. D., and J. B. Willett. 2003. *Applied longitudinal data analysis: Modeling change and event occurrence*: Oxford university press.
Sinha, R. 2008. Chronic stress, drug use, and vulnerability to addiction. *Ann N Y Acad Sci* 1141:105-130.
Skidmore, J. R., and J. G. Murphy. 2011. The effect of drink price and next-day responsibilities on college student drinking: a behavioral economic analysis. *Psychol Addict Behav* 25 (1):57-68.
Smith, A. E., M. P. Martens, J. G. Murphy, J. Buscemi, A. M. Yurasek, and J. Skidmore. 2010. Reinforcing efficacy moderates the relationship between impulsivity-related traits and alcohol use. *Exp Clin Psychopharmacol* 18 (6):521-529.
Smith, J. L., N. J. Dash, S. J. Johnstone, K. Houben, and M. Field. 2017. Current forms of inhibitory training produce no greater reduction in drinking than simple assessment: A preliminary study. *Drug Alcohol Depend* 173:47-58.
Snider, S. E., K. M. Cummings, and W. K. Bickel. 2017. Behavioral economic substitution between conventional cigarettes and e-cigarettes differs as a function of the frequency of e-cigarette use. *Drug Alcohol Depend* 177:14-22.
Snider, S. E., H. U. Deshpande, J. M. Lisinski, M. N. Koffarnus, S. M. LaConte, and W. K. Bickel. 2018. Working Memory Training Improves Alcohol Users' Episodic Future Thinking: A Rate-Dependent Analysis. *Biol Psychiatry Cogn Neurosci Neuroimaging* 3 (2):160-167.
Snider, S. E., S. M. LaConte, and W. K. Bickel. 2016. Episodic future thinking: Expansion of the temporal window in individuals with alcohol dependence. *Alcohol Clin Exp Res* 40 (7):1558-1566.
Sobell, L. C., J. Brown, G. I. Leo, and M. B. Sobell. 1996. The reliability of the Alcohol Timeline Followback when administered by telephone and by computer. *Drug Alcohol Depend* 42 (1):49-54.
Sobell, L. C., M. B. Sobell, G. J. Connors, and S. Agrawal. 2003. Assessing drinking outcomes in alcohol treatment efficacy studies: selecting a yardstick of success. *Alcohol Clin Exp Res* 27 (10):1661-1666.
Solinas, M., L. V. Panlilio, and S. R. Goldberg. 2004. Exposure to delta-9-tetrahydrocannabinol (THC) increases subsequent heroin taking but not heroin's reinforcing efficacy: a self-administration study in rats. *Neuropsychopharmacology* 29 (7):1301-1311.
Spillman, D. 1990. Survey of food and vitamin intake responses reported by university students experiencing stress. *Psychol Rep* 66 (2):499-502.
Stairs, D. J., and M. T. Bardo. 2009. Neurobehavioral effects of environmental enrichment and drug abuse vulnerability. *Pharmacol Biochem Behav* 92 (3):377-382.

Stanger, C., S. R. Ryan, H. Fu, R. D. Landes, B. A. Jones, W. K. Bickel, and A. J. Budney. 2012. Delay discounting predicts adolescent substance abuse treatment outcome. *Exp Clin Psychopharmacol* 20 (3):205-212.
Stead, L. F., and T. Lancaster. 2012. Combined pharmacotherapy and behavioural interventions for smoking cessation. *Cochrane Database Syst Rev* 10:CD008286.
Stein, J. S., M. N. Koffarnus, S. E. Snider, A. J. Quisenberry, and W. K. Bickel. 2015. Identification and management of nonsystematic purchase task data: Toward best practice. *Exp Clin Psychopharmacol* 23 (5):377-386.
Stein, J. S., M. N. Koffarnus, I. Stepanov, D. K. Hatsukami, and W. K. Bickel. 2018a. Cigarette and e-liquid demand and substitution in e-cigarette-naïve smokers. *Exp Clin Psychopharmacol* 26 (3):233.
Stein, J. S., Y. Y. Sze, L. Athamneh, M. N. Koffarnus, L. H. Epstein, and W. K. Bickel. 2017. Think fast: rapid assessment of the effects of episodic future thinking on delay discounting in overweight/obese participants. *J Behav Med*.
Stein, J. S., A. N. Tegge, J. K. Turner, and W. K. Bickel. 2018b. Episodic future thinking reduces delay discounting and cigarette demand: an investigation of the good-subject effect. *J Behav Med* 41 (2):269-276.
Stein, M. D. 1999. Medical consequences of substance abuse. *Psychiatr Clin North Am* 22 (2):351-370.
Stoet, G. 2017. PsyToolkit a novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology* 44 (1):24-31.
Stoops, W. W., J. Dallery, N. M. Fields, P. A. Nuzzo, N. E. Schoenberg, C. A. Martin, B. Casey, and C. J. Wong. 2009. An internet-based abstinence reinforcement smoking cessation intervention in rural smokers. *Drug Alcohol Depend* 105 (1-2):56-62.
Stoops, W. W., and C. R. Rush. 2014. Combination pharmacotherapies for stimulant use disorder: a review of clinical findings and recommendations for future research. *Expert Rev Clin Pharmacol* 7 (3):363-374.
Stoops, W. W., J. C. Strickland, L. R. Hays, A. O. Rayapati, J. A. Lile, and C. R. Rush. 2016. Safety and tolerability of intranasal cocaine during phendimetrazine maintenance. *Psychopharmacology (Berl)* 233 (11):2055-2063.
Strathdee, S. A., and C. Beyrer. 2015. Threading the Needle--How to Stop the HIV Outbreak in Rural Indiana. *N Engl J Med* 373 (5):397-399.
Strickland, J.C., J. S. Beckmann, C. R. Rush, and W. W. Stoops. 2017a. A pilot study of loss aversion for drug and non-drug commodities in cocaine users. *Drug Alcohol Depend* 180:.223-226.
Strickland, J. C., B. L. Bolin, J. A. Lile, C. R. Rush, and W. W. Stoops. 2016a. Differential sensitivity to learning from positive and negative outcomes in cocaine users. *Drug Alcohol Depend* 166:61-68.
Strickland, J. C., J. C. Hill, W. W. Stoops, and C. R. Rush. in press. Feasibility, acceptability, and initial efficacy of delivering alcohol use cognitive interventions via crowdsourcing. *Alcohol Clin Exp Res* epub ahead of print.
Strickland, J. C., J. A. Lile, C. R. Rush, and W. W. Stoops. 2016b. Comparing exponential and exponentiated models of drug demand in cocaine users. *Exp Clin Psychopharmacol* 24 (6):447-455.
Strickland, J. C., J. A. Lile, and W. W. Stoops. 2017b. Unique prediction of cannabis use severity and behaviors by delay discounting and behavioral economic demand. *Behav Processes* 140:33-40.
Strickland, J. C., A. R. Reynolds, and W. W. Stoops. 2016c. Regulation of cocaine craving by cognitive strategies in an online sample of cocaine users. *Psychol Addict Behav*.

Strickland, J. C., and W. W. Stoops. 2015. Perceptions of research risk and undue influence: Implications for ethics of research conducted with cocaine users. *Drug Alcohol Depend* 156:304-310.
Strickland, J. C., and W. W. Stoops. 2017. Stimulus selectivity of drug purchase tasks: A preliminary study evaluating alcohol and cigarette demand. *Exp Clin Psychopharmacol* 25 (3):198-207.
Strickland, J. C., and W. W. Stoops. 2018a. Evaluating autonomy, beneficence, and justice with substance-using populations: Implications for clinical research participation. *Psychology of Addictive Behaviors* 32 (5):552.
Strickland, J. C., and W. W. Stoops. 2018b. Feasibility, acceptability, and validity of crowdsourcing for collecting longitudinal alcohol use data. *J Exp Anal Behav* 110 (1):136-153.
Strickland, J. C., and W. W. Stoops. 2019. The use of crowdsourcing in addiction science research: Amazon Mechanical Turk. *Exp Clin Psychopharmacol*:Epub Ahead of Print.
Strickland, J. C., F. P. Wagner, W. W. Stoops, and C. R. Rush. 2015. Profile of internet access in active cocaine users. *The American Journal on Addictions* 24 (7):582-585.
Sze, Y. Y., J. S. Stein, W. K. Bickel, R. A. Paluch, and L. H. Epstein. 2017. Bleak present, bright future: Online episodic future thinking, scarcity, delay discounting, and food demand. *Clin Psychol Sci* 5 (4):683-697.
Thompson, T., and C. Schuster. 1968. Behavioral pharmacology. *Englewood Cliffs*.
Thorn, D. A., L. Jing, Y. Qiu, A. M. Gancarz-Kausch, C. M. Galuska, D. M. Dietz, Y. Zhang, and J. X. Li. 2014. Effects of the trace amine-associated receptor 1 agonist RO5263397 on abuse-related effects of cocaine in rats. *Neuropsychopharmacology* 39 (10):2309-2316.
Tompkins, D. A., A. S. Huhn, P. S. Johnson, M. T. Smith, E. C. Strain, R. R. Edwards, and M. W. Johnson. 2017. To take or not to take: the association between perceived addiction risk, expected analgesic response and likelihood of trying novel pain relievers in self-identified chronic pain patients. *Addiction*.
Tompkins, D. A., P. S. Johnson, M. T. Smith, E. C. Strain, R. R. Edwards, and M. W. Johnson. 2016. Temporal preference in individuals reporting chronic pain: discounting of delayed pain-related and monetary outcomes. *Pain* 157 (8):1724-1732.
Tremblay, P. F., K. Graham, S. Wells, R. Harris, R. Pulford, and S. E. Roberts. 2010. When do first-year college students drink most during the academic year? An internet-based study of daily and weekly drinking. *J Am Coll Health* 58 (5):401-411.
Tsukayama, E., and A. L. Duckworth. 2010. Domain-specific temporal discounting and temptation. *Judgment and Decision Making* 5:72-82.
Tucker, M. R., M. Laugesen, and R. C. Grace. 2017. Estimating Demand and Cross-Price Elasticity for Very Low Nicotine Content (VLNC) Cigarettes Using a Simulated Demand Task. *Nicotine Tob Res*.
Turner, C. F., L. Ku, S. M. Rogers, L. D. Lindberg, J. H. Pleck, and F. L. Sonenstein. 1998. Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science* 280 (5365):867-873.
Uchalik, D. C. 1979. A comparison of questionnaire and self-monitored reports of alcohol intake in a nonalcoholic population. *Addict Behav* 4 (4):409-413.
US Department of Health Human Services. 2014. The health consequences of smoking—50 years of progress: a report of the Surgeon General. *Atlanta, GA: US 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* 17.

Van Handel, M., C. E. Rose, E. J. Hallisey, J. L. Kolling, J. E. Zibbell, B. Lewis, M. K. Bohm, C. M. Jones, B. E. Flanagan, and A.-E. Siddiqi. 2016. County-level vulnerability assessment for rapid dissemination of HIV or HCV infections among persons who inject drugs, United States. *Journal of acquired immune deficiency syndromes (1999)* 73 (3):323-331.

Veilleux, J. C., K. D. Skinner, E. D. Reese, and J. A. Shaver. 2014. Negative affect intensity influences drinking to cope through facets of emotion dysregulation. *Personality and Individual Differences* 59:96-101.

Verdejo-Garcia, A. 2016. Cognitive training for substance use disorders: Neuroscientific mechanisms. *Neurosci Biobehav Rev* 68:270-281.

Von Neumann, J., and O. Morgenstern. 2007. *Theory of games and economic behavior (commemorative edition)*: Princeton university press.

Walters, K., D. A. Christakis, and D. R. Wright. 2018. Are Mechanical Turk worker samples representative of health status and health behaviors in the US? *PLoS One* 13 (6):e0198835.

Wang, B., Y. Shaham, D. Zitzman, S. Azari, R. A. Wise, and Z. B. You. 2005. Cocaine experience establishes control of midbrain glutamate and dopamine by corticotropin-releasing factor: a role in stress-induced relapse to drug seeking. *J Neurosci* 25 (22):5389-5396.

Wang, B., Z. B. You, K. C. Rice, and R. A. Wise. 2007. Stress-induced relapse to cocaine seeking: roles for the CRF(2) receptor and CRF-binding protein in the ventral tegmental area of the rat. *Psychopharmacology (Berl)* 193 (2):283-294.

Wanmaker, S., S. M. J. Leijdesdorff, E. Geraerts, B. J. M. van de Wetering, P. J. Renkema, and I. H. A. Franken. 2018. The efficacy of a working memory training in substance use patients: A randomized double-blind placebo-controlled clinical trial. *J Clin Exp Neuropsychol* 40 (5):473-486.

Warlick, C., K. P. Richter, D. Catley, B. J. Gajewski, L. E. Martin, and L. M. Mussulman. 2018. Two brief valid measures of therapeutic alliance in counseling for tobacco dependence. *J Subst Abuse Treat* 86:60-64.

Washio, Y., S. T. Higgins, S. H. Heil, T. L. McKerchar, G. J. Badger, J. M. Skelly, and R. L. Dantona. 2011. Delay discounting is associated with treatment response among cocaine-dependent outpatients. *Exp Clin Psychopharmacol* 19 (3):243-248.

Weafer, J., M. J. Baggott, and H. de Wit. 2013. Test-retest reliability of behavioral measures of impulsive choice, impulsive action, and inattention. *Exp Clin Psychopharmacol* 21 (6):475-481.

Weafer, J., and M. T. Fillmore. 2012. Alcohol-related stimuli reduce inhibitory control of behavior in drinkers. *Psychopharmacology (Berl)* 222 (3):489-498.

Wen, X., S. T. Higgins, C. Xie, and L. H. Epstein. 2016. Improving public acceptability of using financial incentives for smoking cessation during pregnancy: A randomized controlled experiment. *Nicotine Tob Res* 18 (5):913-918.

Wilson, A. G., C. T. Franck, M. N. Koffarnus, and W. K. Bickel. 2016. Behavioral Economics of Cigarette Purchase Tasks: Within-Subject Comparison of Real, Potentially Real, and Hypothetical Cigarettes. *Nicotine Tob Res* 18 (5):524-530.

Woods, A. T., C. Velasco, C. A. Levitan, X. Wan, and C. Spence. 2015. Conducting perception research over the internet: a tutorial review. *PeerJ* 3:e1058.

World Health Organization. 2018. *Global status report on alcohol and health, 2018*: World Health Organization.

Yoon, J. H., S. T. Higgins, S. H. Heil, R. J. Sugarbaker, C. S. Thomas, and G. J. Badger. 2007. Delay discounting predicts postpartum relapse to cigarette smoking among pregnant women. *Exp Clin Psychopharmacol* 15 (2):176-186.

Young, A. M., and J. R. Havens. 2012. Transition from first illicit drug use to first injection drug use among rural Appalachian drug users: a cross-sectional comparison and retrospective survival analysis. *Addiction* 107 (3):587-596.

Yu, J., L. Liu, R. L. Collins, P. C. Vincent, and L. H. Epstein. 2014. Analytical problems and suggestions in the analysis of behavioral economic demand curves. *Multivariate Behav Res* 49 (2):178-192.

Yurasek, A. M., A. A. Dennhardt, and J. G. Murphy. 2015. A randomized controlled trial of a behavioral economic intervention for alcohol and marijuana use. *Exp Clin Psychopharmacol* 23 (5):332.

Yurasek, A. M., J. G. Murphy, A. H. Clawson, A. A. Dennhardt, and J. MacKillop. 2013. Smokers report greater demand for alcohol on a behavioral economic purchase task. *J Stud Alcohol Drugs* 74 (4):626-634.

Zeisser, C., K. Thompson, T. Stockwell, C. Duff, C. Chow, K. Vallance, A. Ivsins, W. Michelow, D. Marsh, and P. Lucas. 2012. A 'standard joint'? The role of quantity in predicting cannabis-related problems. *Addiction Research & Theory* 20 (1):82-92.

Zhao, T., X. Luo, H. Chu, C. T. Le, L. H. Epstein, and J. L. Thomas. 2016. A two-part mixed effects model for cigarette purchase task data. *J Exp Anal Behav* 106 (3):242-253.

VITA

JUSTIN CHARLES STRICKLAND

EDUCATION

2017 **Graduate Certificate, Biostatistics**
University of Kentucky, Lexington, KY

2016 **Master of Science, Experimental Psychology**
University of Kentucky, Lexington, KY

2014 **Bachelor of Science, Psychology & Biology**
Summa Cum Laude; First Honor; High Honors in Psychology
Davidson College, Davidson, NC

TRAINING AND PROFESSIONAL EXPERIENCE

2014-2019 **Graduate Research Fellow**
Laboratory of Human Behavioral Pharmacology
University of Kentucky, Lexington, KY

2017 **Guest Associate Editor**
Special Issue: Marijuana
Translational Issues in Psychological Science

2017 **Graduate Teaching Assistant**
Psychology 312: Brain & Behavior (Honors Section)
University of Kentucky, Lexington, KY

2011-2014 **Undergraduate Research Assistant**
Behavioral Pharmacology Research Laboratory
Davidson College, Davidson, NC

2012-2014 **Undergraduate Research Assistant**
Developmental Biology Research Laboratory
Davidson College, Davidson, NC

2013-2014 **Undergraduate Research Assistant**
Herpetology Research Laboratory
Davidson College, Davidson, NC

SCHOLASTIC AND PROFESSIONAL HONORS

Awards and Honors

2019 ASPET Travel Award, American Society for Pharmacology and Experimental Therapeutics

2019 Stimulus Properties of Drugs Behavioral Pharmacology Society Meeting Scholarship, Behavioral Pharmacology Society

2018	Outstanding Oral Presentation, Behavior, Biology, and Chemistry Conference
2018	Student Travel Award, Behavior, Biology, and Chemistry Conference
2018	Student Travel Award, American Psychological Association
2017	William L. Woolverton Travel Award, International Study Group Investigating Drugs as Reinforcers
2017	NIDA Women & Sex/Gender Differences Junior Investigator Travel Award, College on Problems of Drug Dependence
2016	Student Travel Award, American Psychological Association
2016	Student Travel Award, Behavior, Biology, and Chemistry Conference
2015	Student Travel Award, European Behavioural Pharmacology Society
2015	Student Travel Award, American Psychological Association
2015	Early Career Travel Award, American Psychological Association Divisions 28 & 50
2015	Student Travel Award, Behavior, Biology, and Chemistry Conference
2015	Outstanding Poster Presentation, Behavior, Biology, and Chemistry Conference
2014	William Gatewood Workman Award, Davidson College
2014	Sigma Xi Research Award, Davidson College
2014	Tom Daggy Biology Award for Academic Achievement, Service, and Love of Nature, Davidson College
2013	Barry Goldwater Scholarship, Barry Goldwater Scholarship and Excellence in Education Foundation
2011	Alumni Association Award, Davidson College

Research Grants and Fellowships

2018	Psi Chi Graduate Research Grant
2018	University of Kentucky Behavioral Science Pilot Grant
2018	University of Kentucky College of Social Work Collaboration Award Pilot Funds (Co-I)
2017	Society of Addiction Psychology Student Research Grant
2017	University of Kentucky Center on Drug and Alcohol Research Petite Research Grant
2015	Psi Chi Graduate Research Grant
2014	University of Kentucky Behavioral Science Pilot Grant
2014	Davidson College George L. Abernethy Endowment Award

Fellowships

2014-2019	National Science Foundation Graduate Research Fellowship
2014	Multiyear Fellowship, University of Kentucky
2014	Lipman Fellowship, University of Kentucky
2014	Royster Award of Special Distinction, University of Kentucky
2012	Davidson Research Initiative Summer Research Program

PROFESSIONAL PUBLICATIONS

Peer-Reviewed Journal Articles

48. **Strickland JC**, Alcorn III JL, and Stoops WW (in press) Using behavioral economic variables to predict future alcohol use in a crowdsourced sample. *Journal of Psychopharmacology*
47. **Strickland JC**, Lile JA, and Stoops WW (in press) Contribution of cannabis-related cues to concurrent reinforcer choice in humans. *Drug and Alcohol Dependence*
46. **Strickland JC**, Hill JC, Stoops WW, and Rush CR (in press) Feasibility, acceptability, and initial efficacy of delivering alcohol use cognitive interventions via crowdsourcing. *Alcoholism: Clinical & Experimental Research*
45. Lacy RT, Austin BP, and **Strickland JC** (in press) The influence of sex and estrous cyclicity on cocaine and remifentanil demand in rats. *Addiction Biology*
44. Tillson M, Staton M, **Strickland JC**, and Pangburn K (in press) An examination of the age of substance use onset and adult severity of use among offenders entering treatment. *Journal of Drug Issues*
43. Stoops WW, Johnson MF, **Strickland JC**, Knudsen HK, Gilbert GH, Ray MN, Reynolds G, and Studts JL (in press) Feasibility of collecting saliva for biological verification of smoking status from dental clinics and patients' homes: Results from the National Dental PBRN. *Community Dental Health*
42. **Strickland JC** and Stoops WW (2019) The use of crowdsourcing in addiction science research: Amazon Mechanical Turk. *Experimental and Clinical Psychopharmacology, 27*, 1-18
41. Campbell EM and **Strickland JC** (2019) Reliability and validity of the Brief DSM-5 Alcohol Use Disorder Diagnostic Assessment: A systematic replication in a crowdsourced sample. *Addictive Behaviors, 92*, 194-198
40. **Strickland JC**, Stoops WW, Kincer MA, and Rush CR (2019) The impact of financial strain on medication non-adherence: Influence of psychiatric medication use. *Psychiatry Research*, *271*, 389-395
39. Harvanko AM, **Strickland JC**, Sloane SA, Shelton BJ, and Reynolds BA (2019) Dimensions of impulsive behavior: Predicting contingency management treatment outcomes for adolescent smokers. *Addictive Behaviors*, *90*, 334-340
38. Staton M, **Strickland JC**, Webster JM, Leukefeld C, Oser CB, and Pike E (2018) HIV prevention in rural Appalachian jails: Implications for re-entry risk reduction among women who use drugs. *AIDS and Behavior*, *22*, 4009-4018
37. Tosun NL, Allen SS, Eberly LE, Yao M, Stoops WW, **Strickland JC**, Harrison KA, al Absi M, and Carroll ME (2018) Association of exercise with smoking-related symptomology, smoking behavior and impulsivity in men and women. *Drug and Alcohol Dependence*, *192*, 29-37
36. **Strickland JC**, Marks KR, Beckmann JS, Lile JA, Rush CR, and Stoops WW (2018) Contribution of cocaine-related cues to concurrent monetary choice in humans. *Psychopharmacology*, *235*, 2871-2881
35. **Strickland JC**, Chen I, Wang C, and Fardo DW (2018) Longitudinal data methods for evaluating genome by epigenome interactions in families. *BMC Genetics*, *19*, *e82*

34. **Strickland JC** and Stoops WW (2018) Evaluating autonomy, beneficence, and justice with substance-using populations: Implications for clinical research participation. *Psychology of Addictive Behaviors*, *32*, 552-563

33. **Strickland JC** and Stoops WW (2018) Feasibility, acceptability, and validity of crowdsourcing for collecting intensive longitudinal alcohol use data. *Journal of the Experimental Analysis of Behavior*, *110*, 136-153

32. **Strickland JC**, Staton M, Leukefeld CG, Oser CB, and Webster JM (2018) Hepatitis C antibody reactivity among high-risk rural women: Opportunities for services and treatment in the criminal justice system. *International Journal of Prisoner Health*, *14*, 89-100

31. Staton M, **Strickland JC**, Tillson M, Leukefeld C, Webster M, and Oser C (2017) Partner relationships and high-risk practices among rural Appalachian women who inject. *Women's Health Issues*, *27*, 652-659

30. Reynolds AR, **Strickland JC**, Stoops WW, Lile JA, and Rush CR (2017) Buspirone maintenance does not alter the reinforcing, subjective, and cardiovascular effects of intranasal methamphetamine. *Drug and Alcohol Dependence*, *181*, 25-20

29. **Strickland JC**, Beckmann JS, Rush CR, and Stoops WW (2017) A pilot study of loss aversion for drug and non-drug commodities in cocaine users. *Drug and Alcohol Dependence*, *180*, 223-226

28. Tillson M, **Strickland JC**, and Staton M (2017) Age of first arrest, sex, and drug use as correlates of adult risk behaviors among rural women in jails. *Women & Criminal Justice*, *27*, 287-301

27. **Strickland JC** and Stoops WW (2017) Stimulus-selectivity of drug purchase tasks: A preliminary study of alcohol and cigarette demand. *Experimental and Clinical Psychopharmacology*, *25*, 198-207

26. **Strickland JC**, Lile JA, and Stoops WW (2017) Unique prediction of cannabis use severity and behaviors by delay discounting and behavioral economic demand. *Behavioural Processes*, *140*, 33-40

25. **Strickland JC**, Bolin BL, Romanelli MR, Rush CR, and Stoops WW (2017) Effects of acute buspirone administration on inhibitory control and sexual discounting in cocaine users. *Human Psychopharmacology: Clinical and Experimental*, *32*, e2567

24. Smith MA and **Strickland JC** (2017) Modeling the impact of social contact on substance use. *Neuropsychopharmacology Reviews*, *42*, 364-365

23. **Strickland JC**, Lile JA, Rush CR, and Stoops WW (2016) Comparing exponential and exponentiated models of drug demand in cocaine users. *Experimental and Clinical Psychopharmacology*, *24*, 447-455

22. **Strickland JC** and Smith MA (2016) Animal models of resistance exercise and their application to neuroscience research. *Journal of Neuroscience Methods*, *273*, 191-200

21. **Strickland JC**, Reynolds AR, and Stoops WW (2016) Regulation of craving by cognitive strategies in an online sample of cocaine users. *Psychology of Addictive Behaviors*, *30*, 607-612

20. **Strickland JC**, Bolin BL, Lile JA, Rush CR, and Stoops WW (2016) Differential sensitivity to learning from positive and negative outcomes in cocaine users. *Drug and Alcohol Dependence*, *166*, 61-68

19. Lacy RT, **Strickland JC**, Feinstein MA, Robinson AM, and Smith MA (2016) The effects of sex, estrous cycle, and social contact on cocaine and heroin self-administration. *Psychopharmacology*, *233*, 3201-3210

18. Robinson AM, Lacy RT, **Strickland JC**, and Smith MA (2016) The effects of social contact on cocaine intake under extended-access conditions in male rats. *Experimental and Clinical Psychopharmacology*, *24*, 285-296

17. **Strickland JC**, Abel JM, Lacy RT, Beckmann JS, Witte MA, Lynch WJ, and Smith MA (2016) The effects of resistance exercise on cocaine self-administration, muscle hypertrophy, and BDNF expression in the nucleus accumbens. *Drug and Alcohol Dependence*, *163*, 186-194

16. **Strickland JC**, Pinheiro AP, Cecala KK, and Dorcas ME (2016) Relationship between behavioral thermoregulation and physiological function in larval stream salamanders. *Journal of Herpetology*, *50*, 239-244

15. Stoops WW, **Strickland JC**, Hays LR, Rayapati AO, Lile JA, and Rush CR (2016) Safety and tolerability of intranasal cocaine during phendimetrazine maintenance. *Psychopharmacology*, *233*, 2055-2063

14. **Strickland JC**, Feinstein MA, Lacy RT, and Smith MA (2016) The effects of physical activity on impulsive choice: Influence of sensitivity to reinforcement amount and delay. *Behavioural Processes*, *126*, 36-45

13. **Strickland JC** and Stoops WW (2015) Perceptions of research risk and undue influence: Implications for ethics of research conducted with cocaine users. *Drug and Alcohol Dependence*, *156*, 304-310

12. **Strickland JC** and Smith MA (2015) Animal models of social contact and drug self-administration. *Pharmacology, Biochemistry, and Behavior*, *136*, 47-54

11. **Strickland JC**, Wagner FP, Stoops WW, and Rush CR (2015) Profile of Internet access in active cocaine users. *The American Journal on Addictions*, *24*, 582-585

10. Smith MA, **Strickland JC**, Bills SE, and Lacy RT (2015) The effects of a shared history of drug exposure on social choice. *Behavioral Pharmacology*, *26*, 631-635

9. **Strickland JC**, Rush CR, and Stoops WW (2015) Mu opioid mediated discriminative-stimulus effects of tramadol: An individual subjects analysis. *Journal of the Experimental Analysis of Behavior*, *103*, 361-374

8. **Strickland JC**, Bahram CH, Harden L, Pittman SE, Kern MM, and Dorcas ME (2015) Life-history costs of reproductive behaviors in a wetland-breeding amphibian. *Journal of Freshwater Ecology*, *30*, 435-444

7. Lacy RT, **Strickland JC**, Brophy MK, Witte MA, and Smith MA (2014) Exercise decreases speedball self-administration. *Life Sciences*, *114*, 86-92

6. Lacy RT, **Strickland JC**, and Smith MA (2014) Cocaine self-administration in social dyads using custom-built operant conditioning chambers. *Journal of Neuroscience Methods*, *236*, 11-16

5. **Strickland JC**, Lile JA, Rush CR, and Stoops WW (2014) Relationship between intranasal cocaine self-administration and subject-rated effects: Predictors of

cocaine taking on progressive ratio schedules. *Human Psychopharmacology: Clinical and Experimental*, *29*, 342-350

4. **Strickland JC** and Smith MA (2014) The anxiolytic effects of resistance exercise. *Frontiers in Psychology: Movement Science and Sport Psychology*, *e753*
3. Smith MA, Lacy RT, and **Strickland JC** (2014) The effects of social learning on the acquisition of drug self-administration. *Drug and Alcohol Dependence*, *141*, 1-8
2. **Strickland JC** and Smith MA (2014) The effects of social contact on drug use: Behavioral mechanisms controlling drug intake. *Experimental and Clinical Psychopharmacology*, *22*, 23-34
1. Peitz GW, **Strickland JC**, Pitts EG, Foley M, Tonidandel S, and Smith MA (2013) Peer influences on drug self-administration: An econometric analysis in socially housed rats. *Behavioural Pharmacology*, *24*, 114-123

Other Publications:

4. **Strickland JC** and Stoops WW (in press) The prevention and treatment of adolescent stimulant and methamphetamine use. In: Gullota T and Leukefeld C (Eds.), *Adolescent Substance Abuse: Evidence-Based Approaches to Prevention and Treatment (2nd Edition)*
3. **Strickland JC**, Cloutier RM, Ecker AH, & Buckner JD (2018) Advances in psychological research on marijuana. *Translational Issues in Psychological Science, 4*, 1-5 [editorial introduction to special issue]
2. **Strickland JC** (2017) Nonparametric statistics. *The SAGE Encyclopedia of Industrial and Organizational Psychology, 2nd edition*, 1022-1023
1. **Strickland JC** (2014) Textbook review: Guide to research techniques in neuroscience. *The Journal of Undergraduate Neuroscience Education, 13,* R1-R2

Chaired Conference Symposia:

4. **Strickland JC** and Huhn AS (August 2018) The Internet laboratory: A guide to conducting psychological research on Amazon Mechanical Turk. Symposium at the 126th annual meeting of the American Psychological Association: San Francisco, CA
3. Dunn KE and **Strickland JC** (co-chair) (August 2017) Addiction and behavioral science data blast. Symposium at the 125th annual meeting of the American Psychological Association: Washington, DC
2. **Strickland JC** (August 2017) Enhancing presentation skills using emerging technologies. Symposium at the 125th annual meeting of the American Psychological Association: Washington, DC
1. **Strickland JC** and Panicker S (August 2016) Emerging reforms in psychological science: How the changing face of scientific research may Influence your research. Symposium at the 124th annual meeting of the American Psychological Association: Denver, CO

Conference Oral Presentations:

19. **Strickland JC** (August 2018) Utilizing Amazon Mechanical Turk: A live demonstration. Oral presentation at the 126th annual meeting of the American Psychological Association: San Francisco, CA

18. Stoops WW, Alcorn JL, **Strickland JC**, Hays LR, Rayapati AO, Lile JA, and Rush CR (June 2018) Influence of phendimetrazine maintenance on the reinforcing, subjective, and physiological effects of intranasal cocaine. Oral presentation at the 80th annual meeting of the College on Problems of Drug Dependence: San Diego, CA
17. **Strickland JC** and Rush CR (April 2018) Utilizing crowdsourcing to evaluate correlates of psychiatric medication adherence. Oral presentation at the 1st annual Eastern State Hospital Research and Practice Initiatives Day (RAPID): Lexington, KY
16. **Strickland JC***, Marks KR, Beckmann JS, Lile JA, Rush CR, and Stoops WW (March 2018) Contribution of cocaine-related cues to concurrent monetary choice in humans. Oral presentation at the 10th annual Behavior, Biology, and Chemistry: Translational Research in Addiction Conference: San Antonio, TX

***Awarded Outstanding Oral Presentation**

15. **Strickland JC** and Stoops WW (June 2017) Stimulus-selectivity of alcohol and cigarette purchase tasks. Oral presentation at the satellite meeting of the International Study Group Investigating Drugs as Reinforcers: Montreal, Canada
14. **Strickland JC**, Beckmann JS, Rush CR, and Stoops WW (June 2017) Loss aversion in cocaine users: Role of risk and commodity type. Oral presentation at the 79th annual meeting of the College on Problems of Drug Dependence: Montreal, Canada
13. Stoops WW and **Strickland JC** (June 2017) Relationship between loss aversion and delay discounting in an online drug-using sample. Oral presentation at the 79th annual meeting of the College on Problems of Drug Dependence: Montreal, Canada
12. **Strickland JC**, Chen I, Wang C, and Fardo DW (March 2017) Longitudinal data methods for evaluating genome by epigenome interactions in families. Oral presentation at the 20th annual Genetic Analysis Workshop: San Diego, California
11. **Strickland JC**, Rush CR, and Stoops WW (June 2016) Influence of cocaine cues on monetary choice in cocaine users. Oral presentation at the 78th annual meeting of the College on Problems of Drug Dependence: Palm Springs, CA
10. **Strickland JC** (June 2016) Using Amazon's Mechanical Turk (mTurk) to sample substance using populations. Invited workshop presentation at the 78th annual meeting of the College on Problems of Drug Dependence: Palm Springs, CA
9. Harvanko AM, **Strickland JC**, and Reynolds BA (June 2016) Predicting contingency management treatment efficacy by using measures of impulsivity. Oral presentation at the 78th annual meeting of the College on Problems of Drug Dependence: Palm Springs, CA
8. Robinson AM, Lacy RT, **Strickland JC**, Magee CP, and Smith MA (June 2016) The effects of social contact on "binge" cocaine self-administration. Oral presentation at the 78th annual meeting of the College on Problems of Drug Dependence: Palm Springs, CA
7. Putka DJ, **Strickland JC**, and Tonidandel S (April 2016) Estimating relative weights in the face of model selection uncertainty. Oral presentation at the 31st annual

conference of the Society for Industrial and Organizational Psychology: Anaheim, CA

6. **Strickland JC**, Rush CR, and Stoops WW (June 2015) Contribution of conditioned drug action to cocaine self-administration. Oral presentation at the satellite meeting of the International Study Group Investigating Drugs as Reinforcers: Phoenix, AZ
5. Smith MA, **Strickland JC**, Lacy RT, Witte MA, Abel JM, and Lynch WJ (June 2015) The effects of strength training on the positive reinforcing effects of cocaine. Oral presentation at the 59th annual meeting of the Behavioral Pharmacology Society: Boston, MA
4. Lacy RT, **Strickland JC**, and Smith MA (June 2014) The effects of social learning on the acquisition of drug self-administration. Oral presentation at the satellite meeting of the International Study Group Investigating Drugs as Reinforcers: San Juan, Puerto Rico
3. **Strickland JC***, Pinheiro AP, Cecala KK, and Dorcas ME (March 2014) Physiological constraints to respond to climate change: Insights from the effects of temperature on standard metabolic rate in larval salamanders. Oral presentation at the 75th annual meeting of the Association of Southeastern Biologists: Spartanburg, SC

***Awarded Student Research Award**

2. **Strickland JC** (September 2013) Effects of resistance exercise on the positive reinforcing effects of cocaine. Oral presentation at the 2013 Wake Forest-Emory Lab Exchange: Atlanta, GA
1. **Strickland JC**, Nyein, KP, White TE, and Good JJ (April 2013) The effects of Good Samaritan law awareness on helping behavior. Oral presentation at the 38th annual Carolina's Psychology Conference: Raleigh, NC

Conference Poster Presentations:

31. Qalbani SH, Meadows AL, Rush CR, and **Strickland JC** (December 2018) Comparing controlled substances prescribing trends between dentist, physicians, and APRNs using KASPER data from 2011-2017. Poster to be presented at the 29th annual meeting of the American Academy of Addiction Psychiatry: Bonita Springs, FL
30. **Strickland JC** and Stoops WW (August 2018) Perceptions of autonomy, beneficence, and justice by individuals reporting substance use. Poster presentation at the 126th annual meeting of the American Psychological Association: San Francisco, CA
29. **Strickland JC** and Stoops WW (June 2018) Feasibility and validity of collecting intensive longitudinal alcohol use data with Mechanical Turk. Poster presentation at the 80th annual meeting of the College on Problems of Drug Dependence: San Diego, CA
28. Reynolds AR, **Strickland JC**, Stoops WW, Lile JA, and Rush CR (June 2018) Varenicline for smoking cessation in patients with cocaine-use disorder: A proof-of-concept pilot trial. Poster presentation at the 80th annual meeting of the College on Problems of Drug Dependence: San Diego, CA
27. Staton M, **Strickland JC**, Webster JM, Leukefeld C, Oser CB, and Pike E (June 2018) HIV prevention in rural Appalachian jails: Implications for re-entry risk

reduction among women who use drugs. Poster presentation at the 80th annual meeting of the College on Problems of Drug Dependence: San Diego, CA

26. **Strickland JC** and Stoops WW (November 2017) Utilizing the hypothetical purchase task to evaluate cocaine demand during phendimetrazine maintenance. Poster presentation at the 47th annual meeting of the Society for Neuroscience: Washington, DC
25. Staton-Tindall M, **Strickland JC**, Havens JR, and Webster JM (June 2017) Correlates of hepatitis C seropositivity among high-risk rural women: Opportunities for treatment and services in the criminal justice system. Poster presentation at the 79th annual meeting of the College on Problems of Drug Dependence: Montreal, Canada
24. Marks KM, **Strickland JC**, Leukefeld CG, Oser CB, and Staton-Tindall M (June 2017) Strengths can decrease likelihood of drug use among high-risk rural women following brief intervention. Poster presentation at the 79th annual meeting of the College on Problems of Drug Dependence: Montreal, Canada
23. Lacy RT, Austin BP, and **Strickland JC** (June 2017) Sex differences and the role of estrous cyclicity assessing cocaine and remifentanil demand in rats. Poster presentation at the 79th annual meeting of the College on Problems of Drug Dependence: Montreal, Canada
22. Alcorn JL III, **Strickland JC**, Lile JA, Stoops WW, and Rush CR (March 2017) Acute methylphenidate administration reduces attentional bias in cocaine users, but does not change behavioral inhibition. Poster presentation at the 9th annual Behavior, Biology, and Chemistry: Translational Research in Addiction Conference: San Antonio, TX
21. **Strickland JC** and Stoops WW (August 2016) Latent factor structure of cocaine demand in an online sample of cocaine users. Poster presentation at the 124th annual meeting of the American Psychological Association: Denver, CO
20. Wagner FP, Romanelli MR, **Strickland JC**, Lile JA, Stoops WW, and Rush CR (August 2016) Relationship between age of drug use initiation and self-reported ADHD symptoms in cocaine users. Poster presentation at the 124th annual meeting of the American Psychological Association: Denver, CO
19. **Strickland JC**, Lile JA, Rush CR, and Stoops WW (March 2016). Sensitivity to reinforcement and punishment learning in active cocaine users. Poster presentation at the 8th annual Behavior, Biology, and Chemistry: Translational Research in Addiction Conference: San Antonio, TX
18. Magee CP, Lacy RT, Robinson AM, **Strickland JC**, and Smith MA (November 2015) The effects of social contact on “binge” cocaine self-administration. Poster presentation at the 23rd annual Faculty for Undergraduate Neuroscience Satellite Event at the Society for Neuroscience: Chicago, IL’
17. **Strickland JC** and Stoops WW (September 2015) Perceptions of research risk and undue influence in an online sample of cocaine users. Poster presentation at the 16th annual meeting of the European Behavioral Pharmacology Society: Verona, Italy
16. **Strickland JC**, Stoops WW, and Rush CR (August 2015) The association between intranasal methamphetamine self-administration and subject-rated effects. Poster presentation at the 123rd annual meeting of the American Psychological Association: Toronto, Canada

15. Lacy RT, **Strickland JC**, Bills SE, and Smith MA (August 2015) A shared history of drug exposure influences social preference. Poster presentation at the 123rd annual meeting of the American Psychological Association: Toronto, Canada
14. **Strickland JC**, Stoops WW, and Rush CR (June 2015) The relationship between methamphetamine self-administration and subject-rated effects. Poster presentation at the 77th annual meeting of the College on Problems of Drug Dependence: Phoenix, AZ
13. Wagner F, **Strickland JC**, Stoops WW, and Rush CR (June 2015) Feasibility of web-based treatment delivery for cocaine use disorder: Profile of Internet access by active cocaine users. Poster presentation at the 77th annual meeting of the College on Problems of Drug Dependence: Phoenix, AZ
12. Lacy RT, Feinstein MA, **Strickland JC**, and Smith MA (June 2015) The effects of estrous cycling on cocaine self-administration in socially housed male-female dyads. Poster presentation at the 77th annual meeting of the College on Problems of Drug Dependence: Phoenix, AZ
11. **Strickland JC***, Rush CR, and Stoops WW (March 2015) Discriminative-stimulus effects of tramadol: An individual subjects analysis of mu opioid-receptor mediated effects. Poster presentation at the 7th annual Behavior, Biology, and Chemistry: Translational Research in Addiction Conference: San Antonio, TX

***Awarded Outstanding Poster Presentation**

10. **Strickland JC**, Lacy RT, Brophy MK, Witte MA, and Smith MA (June 2014) Aerobic exercise decreases speedball self-administration in female rats. Poster presentation at the 76th annual meeting of the College on Problems of Drug Dependence: San Juan, Puerto Rico
9. Smith MA, Lacy RT, and **Strickland JC** (June 2014) The effects of social learning on the acquisition of cocaine self-administration. Poster presentation at the 76th annual meeting of the College on Problems of Drug Dependence in San Juan, Puerto Rico
8. Bahram CH, **Strickland JC**, Harden LA, Pittman SE, Kern MM, and Dorcas ME (March 2014) Influence of sex and migration behavior on reproductive cost of spotted salamander. (*Ambystoma maculatum*). Poster presentation at the 75th annual meeting of the Association of Southeastern Biologists: Spartanburg, SC
7. Smith MA, **Strickland JC**, and Witte MA (June 2013) The effects of strength training on cocaine self-administration. Poster presentation at the 75th annual meeting of the College on Problems of Drug Dependence: San Diego, CA
6. Smith MA, **Strickland JC**, Pitts EG and Witte MA (May 2013) The effects of forced running procedures on the self-administration of cocaine. Poster presentation at the 1st annual meeting of Collaborative Perspectives on Addiction: Atlanta, GA
5. **Strickland JC**, Witte MA, and Smith MA (March 2013) The effects of exercise on cocaine self-administration: Role of strength and resistance training. Poster presentation at the 7th annual Symposium for Young Neuroscientists and Professors of the SouthEast: Columbia, SC
4. **Strickland JC** and Smith MA (October 2012) The effects of forced running procedures on the self-administration of cocaine. Poster presentation at the 20th annual Faculty for Undergraduate Neuroscience Satellite Event at the Society for Neuroscience: New Orleans, LA

3. Smith MA, Pietz GW, **Strickland JC**, Pitts EG, Tonidandel S, and Foley MC (October 2012) Peer influences on drug self-administration: An econometric analysis in socially housed rats. Poster presentation at the 42nd annual meeting of the Society for Neuroscience: New Orleans, LA
2. **Strickland JC**, Pitts EG, and Smith MA (March 2012) The relationship between exercise duration and the positive reinforcing effects of cocaine. Poster presentation at the 6th annual Symposium for Young Neuroscientists and Professors of the SouthEast: Columbia, SC
1. **Strickland JC** and Smith MA (November 2011) The effects of aerobic exercise on cocaine self-administration: Importance of exercise output. Poster presentation at the 19th annual Faculty for Undergraduate Neuroscience Satellite Event at the Society for Neuroscience: Washington, DC