Can Computers Assist Treatment? Virtual Reality as a Possible Cue Exposure Technique With Adolescent Substance Abusers

by

Jacqueline Renée Hersh

Department of Psychology & Neuroscience
Duke University

Date:_____

Approved:

_____
John F. Curry, Supervisor

_____
M. Zachary Rosenthal

_____
Timothy J. Strauman

_____
John G. Looney

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Psychology & Neuroscience in the Graduate School
of Duke University

2014

ABSTRACT

Can Computers Assist Treatment? Virtual Reality as a Possible Cue Exposure Technique With Adolescent Substance Abusers

by

Jacqueline Renée Hersh

Department of Psychology & Neuroscience
Duke University

Date:_____

Approved:

_____
John F. Curry, Supervisor

_____
M. Zachary Rosenthal

_____
Timothy J. Strauman

_____
John G. Looney

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology & Neuroscience in the Graduate School of Duke University

2014

Abstract

Substance use disorders are one of the most common psychiatric diagnoses among adolescents; marijuana is the illicit drug used most frequently by youth. Treatment dropout and relapse following treatment are common; innovative strategies are needed to improve treatment outcomes for youth substance abusers. The aim of this study was to develop and evaluate the feasibility of a virtual reality (VR) cue reactivity paradigm for adolescent cannabis abusers and to compare it to a video cue reactivity paradigm. Forty-two treatment-seeking youth with a cannabis use disorder completed the study, which incorporated three parts. During Part 1, drug and neutral video clips were shown to 11 youth and five substance-abuse experts who provided craving/usefulness ratings for each video clip. During Part 2, five youth met in a focus group and then individually to provide input on the development of the VR paradigm. During Part 3, 26 youth completed a laboratory procedure involving neutral and drug-related video clips and VR presentations. Heart rate, skin conductance, and skin temperature were measured as well as craving. Higher levels of craving and skin conductance were observed during drug-related presentations. The presentations did not significantly differ in their ability to elicit craving and arousal. Results suggest that youth can experience subjective and physiological reactivity to VR drug cues warranting further study with a larger, more diverse sample. Implications are discussed.

Dedication

I dedicate my dissertation to my family who has supported me throughout life, no matter how difficult the circumstances were. I would like to express particular gratitude to my loving parents, Margaret and Howard Hersh, for their encouragement, constant willingness to prioritize my well-being and success, and faith in my ability to accomplish anything and everything. I love and appreciate you more than I could ever express. I am also grateful to my siblings, David and Lisa Hersh, who since childhood have helped raise me as though I was their own child, and my brother- and sister-in-law, Michael Casaren and Sung Yun, who have always supported and treated me as though I was their own sibling. Without the guidance and support of my family, I never would have been able to complete graduate school and this dissertation. Thank you to you all.

Contents

Abstract iv

List of Tables ix

List of Figures x

Acknowledgements xi

1. Introduction 1

    1.1. What is a Craving? 3

    1.2. What is a Cue? 4

    1.3. How Craving and Cues Are Related to Drug-Seeking Behavior 5

    1.4. Craving Among SUD-Adolescents 7

    1.5. Cue-Exposure Paradigms 9

        1.5.1. Cue-exposure paradigms among youth 10

    1.6. Virtual Reality 11

    1.7. Virtual Reality with Substance Abusers 12

2. Current Study 18

    2.1. Hypotheses 18

3. Methods 20

    3.1. Study Part 1: Video Clip Paradigm Development 21

        3.1.1. Participants 21

        3.1.2. Materials 22

            3.1.2.1. YouTube videos 22

3.1.3. Procedure 23
3.2. Study Part 2: Virtual Reality Paradigm Development 25
3.2.1. Participants 25
3.2.2. Materials and measures 25
3.2.2.1. Virtual reality software 25
3.2.2.2. Immersion 27
3.2.2.3. Tolerability and likeability 28
3.2.2.4. Simulator sickness 28
3.2.3. Procedure 29
3.3. Study Part 3: Virtual Reality Paradigm versus Video Paradigm 33
3.3.1. Participants 34
3.3.2. Materials and measures 37
3.3.2.1. Cue stimuli 37
3.3.2.2. Physiological measures 39
3.3.2.3. Craving, emotional arousal, and emotional valence 40
3.3.2.4. Immersion and presence 42
3.3.2.5. Simulator sickness questionnaire 43
3.3.2.6. Additional participant information 43
3.3.2.7. Additional questions 46
3.3.3. Procedure 47
4. Results 51
4.1. Study Part 1: Video Clip Paradigm Development 51

4.1.1. Youth ratings 52
4.1.2. Expert ratings 53
4.1.3. Video selection 54
4.2. Study Part 2: Virtual Reality Paradigm Development 56
4.2.1. Scene selection 56
4.2.2. Presence 59
4.2.3. Additional questions 60
4.3. Study Part 3: Paradigm Testing 61
4.3.1. Data analytic plan 62
4.3.2. Normality 63
4.3.3. Preliminary analyses 65
4.3.4. Presence 71
4.3.5. Differences in craving 72
4.3.5.1. Exploratory analysis of craving by scene 74
4.3.6. Differences in arousal 75
4.3.7. Differences between the paradigms 77
5. Discussion 87
Appendix A 95
Appendix B 98
References 103
Biography 110

List of Tables

Table 1: Description of when measures were administered during Part 3 of the study .. 50

Table 2: Descriptions of the videos 51

Table 3: Comparison of contents of the eight drug and four neutral videos..... 52

Table 4: Intercorrelations among the dependent variables..... 66

Table 5: Significant correlations among craving variables and potential covariates..... 67

Table 6: Significant correlations among physiological variables and potential demographic covariates..... 68

Table 7: Significant correlations among select physiological variables and additional potential covariates 69

Table 8: Summary of analyses following significant multivariate effect of cue in the prediction of MCQ change scores and single-item craving change scores..... 73

Table 9: Summary of analyses following significant multivariate effect of cue in the prediction of physiological change scores 77

Table 10: Participant responses to which paradigm they prefer and why 85

List of Figures

Figure 1: Screenshot of the VR neighborhood with the middle-class and abandoned house. 30

Figure 2: Average responses to questions regarding each paradigm. 78

Figure 3: Participant responses to statements regarding how interesting the presentations were. 79

Figure 4: Participant responses to statements regarding whether they had a positive reaction to the presentations. 80

Figure 5: Participant responses to statements regarding how enjoyable the presentations were. 80

Figure 6: Participant responses to statements regarding acceptability of the presentations for creating craving/urges to use. 81

Figure 7: Participant responses to statements regarding willingness to use the presentations as part of treatment. 81

Figure 8: Participant responses to statements regarding whether the presentations would increase interest in treatment and/or likelihood of staying in treatment. 82

Figure 9: Participant responses to statements regarding the effectiveness of the presentations in creating cravings/urges to use 82

Figure 10: Participant responses to statements regarding the presentations *not* having bad side effects. 83

Figure 11: Participant responses to statements regarding how realistic the presentations were. 83

Figure 12: Participant responses to statements regarding whether the participant would want to continue interacting with the presentations. 84

Acknowledgements

I would like to thank Drs. John Curry, M. Zachary Rosenthal, Timothy Strauman, Philip Costanzo, and John Looney for their guidance and support throughout the course of this research project. I am especially grateful to Dr. John Curry for his enduring support both personally and professionally throughout my graduate school career. I am also especially grateful to Dr. Zachary Rosenthal for allowing use of his lab and materials and to his lab team for their invaluable assistance in training me on usage of virtual reality and physiological measures. This study could not have been completed without the assistance of three substance abuse counselors: Elise Alexander, Roxanne Ellington, and Gary Sauls whose dedication to improving the lives of youth substance abusers is inspiring. I would also like to thank Eva Bahnuk for assistance in collecting, entering, and quality-checking data for this project. I would additionally like to thank consultants with the Social Science Research Institute at Duke University and Dr. Mark Leary for their consultation regarding the analyses used in this study. I would also like to thank the many YouTube video owners who granted permission for the use of their videos in this study, including (but not limited to) the following YouTube subscribers who requested acknowledgment: smokingJOEnl, ThePafy, bayerdv82, CincinnatiGifts, Linguaspectrum, and jjnature.

1. Introduction

Substance use and substance use disorders are significant public health concerns in the United States, including among the youth population. Research has demonstrated that substance use disorders (SUDs) comprise one of the leading diagnoses in adolescence (Copeland, Shanahan, Costello, & Angold; 2011; Merikangas et al., 2010). The diagnosis of a SUD reflects chronic and heavy use with impairment in functioning or distress. According to the 2012 National Survey on Drug Use and Health (NSDUH), it was estimated that 6.1% of youth aged 12-17 would receive a diagnosis of substance abuse or dependence (SAMHSA, 2013). The 2013 Monitoring the Future (MTF) national survey estimated a lifetime prevalence rate for any illicit drug use to be 35.8% among youth surveyed (those in grades 8, 10, and 12). The annual prevalence rate was estimated to be 28.4% and the 30-day prevalence rate was estimated to be 17.4% (Johnston, O'Malley, Miech, Bachman, & Schulenberg, 2014).

Youth reports across these two national surveys indicate that cannabis was the illicit drug used most frequently by 14-17 year olds (Johnston et al., 2014; SAMHSA, 2013). According to the 2012 NSDUH survey, among 12-13 year olds, prescription drugs used nonmedically (predominantly pain relievers) were the most frequent illicit drug (1.7%) followed by cannabis (1.2%). Cannabis was the illicit drug with the highest rate of past year dependence or abuse in 2012 for people aged 12 or older (4.3 million, which represents 58.9% of all those classified with illicit drug dependence or abuse)(SAMHSA,

2013). While the more recent surveys do not provide estimates of specific substance use disorder diagnoses for youth aged 12-17, according to the 2007 NSDUH survey, it was estimated that 4.3% would receive a diagnosis of cannabis abuse or dependence (Chung & Martin, 2011). Importantly, the 2013 MTF survey found that while rates of alcohol and cigarette use decreased among 8th, 10th, and 12th graders, there was an increase in lifetime cannabis prevalence rates for all three grades surveyed. Further, annual prevalence of cannabis use among 8th and 10th graders increased (to 12.7% and 29.8%, respectively) as did 30-day prevalence rates (7% and 18%, respectively). Rates among 12th graders remained relatively stable. For the three grades combined, annual, 30-day, and daily prevalence rates of cannabis use all increased compared to the 2012 survey (Johnston et al., 2014).

Early treatment is imperative given the consequences of use among adolescents. Acutely, ingesting large amounts of cannabis can lead to panic attacks and in some instances, even to acute psychosis (Brown & Coupey, 1993). More generally, cannabis use has also been related to impairment in short-term memory and attention, motor skills, and reaction time (Hall & Solowij, 1998). These impairments in attention and reaction time can be detrimental in risky situations, such as driving a motor vehicle. Further, adolescent cannabis use has been shown to be associated both with structural changes in the brain and with changes in cognitive functioning such as learning and problem-solving (e.g., Paulus & Tapert, 2010).

With prolonged abstinence, at least some of the effects of adolescent substance use can diminish (Jacobus, Bava, Cohen-Zion, Mahmood, & Tapert, 2009). This highlights the need to treat adolescent substance use early to prevent or reverse potentially damaging effects of use. This is a complicated task given low retention in treatment, with dropout rates ranging from 20% to 50% among substance abusing adolescents (Winters, 1999). Further, these youth commonly experience poor long-term outcomes with findings indicating 70% of adolescents who have completed treatment programs relapse (Florsheim, Heavin, Tiffany, Colvin, & Hiraoka, 2008). Hence, it is important to develop additional engaging and effective treatment components that target the specific difficulties youth face when quitting such as craving. Virtual reality could be one mechanism of targeting craving and enhancing treatment engagement and effectiveness among adolescents. Section 1.6 below includes more information on virtual reality.

1.1. What is a Craving?

Among both humans and animals, drug effects and cues have been associated with drug urges/craving (with the terms "urges" and "cravings" being used interchangeably). While some researchers debate how to define the terms "craving" and "urges", Tiffany and Conklin (2000) summarized usage of the phrases by noting that:

> Across almost all theories, drug urges are assumed to be subjective, emotional-motivational states. They are viewed as subjective in the sense that they refer to the phenomenological experience of the individual, emotional in that the subjective experience of urge has some hedonic quality, and motivational in the

> sense that the subjective urge state presumably activates drug-seeking behavior. (p. 148)

The authors suggest that drug urges and cravings are unique to the individual and motivate drug-seeking behavior. Kozlowski and Wilkinson have highlighted the strong perceived need for a drug as part of the craving process (as cited in Florsheim et al., 2008, p. 1206).

1.2. What is a Cue?

A cue can be defined as any stimulus that is associated with the ingestion of drugs (Drummond & Glautier, 1994). Examples of cues include the sight and smell of a drug or the image of someone ingesting the drug. Research has shown that when paired with repeated drug use, cues can become associated with the act of using (Drummond & Glautier, 1994). This is related to the observation that drug abusers are more likely to relapse and use drugs when they are in environments associated with past drug use (Carter & Tiffany, 1999). Drug abusers are surrounded by several cues on a daily basis, some of which are unavoidable and make it more difficult to remain abstinent despite a desire to quit using drugs. For instance, drug abusers may share a home with others who use drugs, have to continue living in a neighborhood where drug use is common and see or smell the substance they used, or attend school or work with people they have used with, any of which can constitute triggers and also, can lead to increased pressure from others to use.

1.3. How Craving and Cues Are Related to Drug-Seeking Behavior

A large number of research studies evidence that urges and cravings have in turn been associated with drug-seeking behavior and relapse (Bradizza, Stasiewicz, & Maisto, 1994; Drummond, 2001; Jellinek, et al., 1955). Even early on, researchers posited that craving is a defining aspect of addiction and can be used to explain initiation, maintenance, and relapse associated with drug use (Jellinek, et al., 1955).

Classical conditioning and social learning theories have largely been used to explain the association between cues and drug use or relapse (Bradizza, et al., 1994). From a classical conditioning perspective, environmental cues (e.g., a beer bottle) can become conditioned stimuli (CS) that elicit conditioned drug responses (e.g., craving) if these cues have been repeatedly paired with the effects of a drug (e.g., intoxication). These conditioned responses (CR) (e.g., desire for the drug, affective changes, physiological responses) act as motivators of drug use and drug-seeking behaviors (Bradizza, et al., 1994).

Within the classical conditioning framework, models of drug urges and cravings have been attributed either to drug withdrawal (negative reinforcement whereby craving reflects the anticipation and desire for the relief from withdrawal effects (Ludwig & Wikler, 1974)) or to appetitive (positive reinforcing) effects of drugs. There appears to be more experimental support for the conditioned appetitive model than for the conditioned withdrawal model (Rohsenow & Monti, 1999). Repeatedly research

evidences that non-addicted rats (Wise & Bozarth, 1987) and even humans (Gilman, Ramchandani, Davis, Bjork, & Hommer, 2008) will self-administer drugs (e.g., alcohol, cocaine, opiates, amphetamines) because of the appetitive effects or positively reinforcing properties. Proponents of this approach argue that craving reflects a desire for the excitatory or euphoric effects of using a drug (Mcauliffe & Gordon, 1974); indeed evidence exists illustrating that memories for the positive effects of a drug relate to craving (Wise, 1988). Additionally, neuroscientists have proposed that at least some drugs (e.g., opiates and stimulants) act on common neurochemical systems of the brain and produce positive appetitive states. These positive appetitive states not only maintain drug use but also lead conditioned stimuli (e.g., cannabis rolling papers or an empty beer bottle) to arouse these neural states when not in the presence of the drug (Stewart, Dewit, & Eikelboom, 1984). This can lead the cues to increase the number of drug-related thoughts and hence the possibility of relapse.

Social cognitive learning theory posits that there are many learned cognitions and behaviors associated with relapse such as conditioned responses, positive expectations of drug effects, and unsuccessful coping strategies that include assuming that the only way to handle high relapse-risk situations is to use the drug (Marlatt, 1985; Niaura et al., 1988). This theory enhances the classical conditioning model by adding that CSs or stressful situations can increase urges to use or craving, especially when combined with such cognitive variables as positive drug expectations.

Both classical conditioning and social cognitive learning theories highlight the importance of drug-related cues and their association with relapse. Because a strong relationship exists between cues and relapse and has been shown across studies and across theoretical models, drug-related cues can serve an important function during treatment. Since many of the cues will be unavoidable (e.g., people, places, activities that were associated with one's use), utilizing cue-exposure in treatment could assist recovering addicts in preparing to face the risky situations and likely future cravings they will have, while they are still receiving therapeutic support. Additionally, some researchers argue that becoming aware of drug cues and one's reaction to them can protect against using the drug (Rohsenow & Monti, 1999). Cue-exposures can allow recovering addicts to become more aware of their own reactions to cues and if this occurs during treatment, the treatment team can assist addicts in recognizing their feelings and physical reactions.

1.4. Craving Among SUD-Adolescents

In order for cue exposure to be useful in treatment, it is important to know if adolescents with substance use disorders exhibit craving similar to that of adults. The limited research to date suggests that they do. Indeed, one study addressing withdrawal symptoms in 21 youth (aged 13–19 years old) with cannabis dependence indicated that "thoughts of and cravings for cannabis" were one of the symptoms endorsed with the greatest severity in week one of SUD treatment. Furthermore, the authors found that

youth ratings of overall severity of withdrawal were the most significantly correlated with thoughts of (r = .86) and craving for (r = .69) cannabis (Milin, Manion, Dare, & Walker, 2008). Similarly, in another study investigating adolescent cannabis withdrawal, among the 72 adolescents (aged 14-19 years old) included, one of the most commonly reported symptoms was craving for cannabis, which was also one of the few symptoms that no less than 30% of the sample rated as being of at least moderate severity (Vandrey, Budney, Kamon, & Stanger, 2005). Although craving exists among youths with SUDs, the authors highlight that the incidence and magnitude of craving were lower among adolescents when compared to their adult sample (Vandrey et al., 2005).

Further considering the impact of alcohol-related cues on SUD-youth, neuroscientists have found that adolescents with alcohol use disorders exhibit significantly greater brain activation in response to images of alcoholic beverages than do controls, with the greatest activation amongst those who drank the most and expressed the strongest desire to drink (Tapert et al., 2003). This was principally in the left anterior, limbic, and visual system areas, which are regions that have been associated with reward, desire, positive affect, and episodic recall (possibly supporting the appetitive model of craving among youth)(Tapert et al., 2003). Together, these studies suggest that craving can be elicited in youth and hence, paradigms that target craving could be feasible in this population and are worth investigation.

1.5. Cue-Exposure Paradigms

Cue-exposure for use in treatments can be understood from a classical conditioning perspective. Specifically, cue-exposure can be useful for facilitating extinction, done by repeatedly presenting the conditioned stimulus (CS) without the drug or drug effects. This is thought to lead to the extinction of the conditioned response or rather, to reduce the motivation to use the drug (Rankin, Hodgson, & Stockwell, 1983). For instance, both early (Rankin et al., 1983) and recent studies (George et al., 2008) have used paradigms that include participant's holding and smelling their choice drink (the conditioned stimulus) without drinking it, which blocks drug effects (response).

Several cue-exposure paradigms exist, which typically use imagery scripts (describing a drug-related scene), drug-related images or videos, *in vivo* cues (e.g., holding cannabis), or virtual reality cues. An early study with adult alcoholics evidenced that *in vivo* cues, which included viewing, holding, and smelling their choice beverage produced a significantly greater impact on behavioral and subjective ratings of desire to drink and difficulty to resist alcohol than imaginal cues, which produced trivial changes on these variables (Rankin et al., 1983). Other studies suggest that using drug-related images and recall of autobiographical memories of craving are as effective as *in vivo* exposures for eliciting craving (Weinstein, Lingford-Hughes, Martinez-Raga, & Marshall, 1998).

1.5.1. Cue-exposure paradigms among youth

Limited studies have investigated cue-exposure among adolescents. One study used a laboratory-based cue-exposure paradigm to assess the link between cannabis craving and cues (Gray, LaRowe, & Upadhyaya, 2008). Consistent with the adult cue-exposure studies, the authors used multiple types of cues to explore whether the type of cue influenced the results. This included: 1) auditory imagery scripts (whereby youth imagined themselves in the scene being read to them) that consisted of either a neutral beach script or a cannabis script that described sitting in a room surrounded by friends who were smoking cannabis; 2) video cues that showed clips of either a neutral nature scene with adolescents drinking water or cannabis-related clips that showed adolescents preparing or smoking cannabis; and 3) *in vivo* cues that involved the adolescents touching neutral (e.g., pencil, eraser) or cannabis-related objects (e.g., cannabis cigarette, lighter).

The researchers obtained self-report measures of craving as well as physiological measures of heart rate and skin conductance. Results indicated that youth exhibited greater craving and increased skin conductance during the presentation of cannabis cues than during neutral cues. Consistent with the adult literature (e.g., Niaura et al., 1998; Staiger & White, 1991), youth participants exhibited a greater skin conductance response to the *in vivo* cues than to the imaginal or video cues. This study is particularly

important because it not only indicates that youth experience cue-reactivity but also demonstrates that the reactivity can vary based on the paradigm being utilized.

1.6. Virtual Reality

Past research suggests that virtual reality (VR) paradigms are equally or more effective than other cue-exposure paradigms at eliciting craving in adults (Bordnick, et al., 2008; Kuntze, et al., 2001; Lee, et al., 2003). These paradigms utilize a virtual environment comprised of computer-simulated interactive spaces that attempt to provide the user with the look and feel of a real setting. This is accomplished by use of enhanced visual experiences with 3D environments projected either on a desktop computer screen or through the use of a head-mounted display (HMD). The HMD has a small display optic in front of one eye (monocular HMD) or both eyes (binocular HMD) to display the virtual images. To further enhance the realistic feel of virtual reality paradigms, the user can be given the opportunity to navigate through a virtual space and interact with digital versions of humans, referred to as characters or avatars. These characters can talk, gesture, and move around. Characters and the virtual scenes that users view can be modified and tailored to the developer and user needs.

Both the realistic feel of VR and the ability to tailor the visual experience provide a uniquely advantageous use of VR as a cue-exposure paradigm for substance abusers. Stimuli can be chosen and strategically placed as cues to induce craving (e.g., a character offering the user the opportunity to smoke cannabis); further, the amount and intensity

of cues can be manipulated in a virtual environment. Some studies have shown that there are differences in craving based on whether the virtual reality paradigm includes characters or not: among adults with alcohol use disorders exposed to virtual environments, craving was higher in situations with avatars (social pressure) than without (no social pressure)(Cho, et al., 2008).

1.7. Virtual Reality with Substance Abusers

Virtual reality (VR) is increasingly being used as a cue-exposure paradigm for substance abusers. Studies have demonstrated that virtual reality cues can elicit self-reported craving in adult cigarette smokers, drinkers, and cannabis smokers (e.g., Baumann & Sayette, 2006; Bordnick et al., 2004; Bordnick et al., 2009; Bordnick et al., 2008).

Limited studies exist examining the use of virtual reality with adolescents. One study, conducted by Lee and colleagues (2004) exposed 16 older adolescent males (M = 17 years old) who smoked at least 10 cigarettes a day to a virtual reality environment. The VR world included craving environments (e.g., virtual bar) and objects (an alcoholic drink, a packet of cigarettes, a lighter, an ashtray, a glass of beer, and advertising posters) as well as a virtual person smoking (an avatar), and sound designed to replicate the noise and music of a restaurant. The experimenters evidenced that with six sessions of exposure to the virtual environment, the craving for cigarettes was gradually reduced.

This study demonstrates that virtual reality can be used to elicit craving among adolescents and further that it can be used in treatment to reduce cravings.

A pivotal study by Bordnick and colleagues (2009) exposed 20 predominantly African American non-treatment-seeking adults who met criteria for cannabis abuse or dependence to a VR cue-exposure paradigm. When participants arrived at the research center, they provided a saliva sample for rapid drug screen testing and then completed an initial 15-minute VR session in a VR world with no drug-related stimuli. This initial session allowed participants to become familiar with the VR technology and the visual analog scale (VAS) used throughout the experiment. Participants then underwent a mandatory 20-minute break. After returning, the participants completed a five minute introduction period followed by a 12 minute VR session consisting of three minute exposures to each of four VR environments: 1) neutral room 1; 2) paraphernalia room; 3) social interaction party room; and 4) neutral room 2. The neutral rooms were identical and resembled an art gallery with no cannabis-related cues; participants were randomized into viewing either the social room or paraphernalia room first. The paraphernalia room had no characters present and included several inanimate cannabis cues such as bongs, joints, rolling papers, and cannabis plants. The party room included indoor and outdoor areas of a home in which characters were eating food, drinking alcohol, and smoking cannabis. Additionally, at times the characters interacted with the participant by offering the participant to "take a hit". After each room, participants were

asked to measure their current level of craving using a VAS scale that was displayed in the VR system as a way to maintain immersion in the VR world. At this time, participants were also asked to rate their attention to the sight, smell (since olfactory cues were presented), and thoughts related to smoking cannabis on additional VAS scales.

In this study, Bordnick and colleagues (2009) found neither main effects nor interactions related to the presentation order of the rooms and no significant differences between the paraphernalia and party rooms. They found that participants experienced significantly more craving, attention to the sight and smell of cannabis, as well as more cannabis-related thoughts in the rooms that included cannabis-related cues compared to the neutral rooms. This study was vital for showing that virtual reality can be used to elicit strong craving in cannabis abusers (with large Cohen's *d* effect sizes such as 1.07 for cannabis rooms compared to neutral rooms). There remain several questions that need to be researched including whether tailoring of stimuli influences craving, whether the results are generalizable to other populations and age groups, whether other related measures yield consistent findings, and how VR compares to other cue-exposure paradigms. The present study will partially address gaps in the literature by investigating if these findings can be replicated and extended to a treatment-seeking adolescent population, utilizing objective physiological measures that are less susceptible to potential prompting, bias, or participant expectation effects, as well as

comparing the VR system to another non-VR approach, which is necessary for arguing that VR is a useful alternative cue-exposure paradigm.

There is currently a need in the field to more fully examine the use of virtual reality with adolescents and specifically for the range of substances that they commonly use. The most recent national surveys of adolescent drug use suggest youth most often use and abuse cannabis and alcohol (Johnston et al., 2014; SAMHSA, 2013). Given the rates of these disorders among youth, research needs to examine virtual reality cues with these two specific substances.

Specifically, virtual reality is a useful paradigm to test since it can offer advantages beyond those of other cue-exposure paradigms. As Baumann and Sayette (2006) highlighted, although the cravings they induced using virtual reality were comparable to those elicited by *in vivo* exposure in other studies, virtual reality possibly offers "…the advantage of simulating more naturalistic and complex settings in a controlled environment" (p. 484). In a research setting, virtual reality allows researchers to have control of the timing and presentation of stimuli, which limits variation across participants and/or exposure sessions. Further, it can allow for easy tailoring of stimuli, creating a more personal and realistic cue-exposure. In a clinical setting, the more naturalistic and complex stimuli used in virtual reality likely promote a more realistic feel than that of other cue-exposure paradigms; this might enhance clinicians' abilities to assist substance abusers to experience extinction with craving-inducing cues.

Additionally as a benefit for both research and clinical settings, virtual reality uniquely provides scenes that would be difficult to create realistically in an office (e.g., a party scene) and that would have ethical risks for safety and maintenance of confidentiality if real world exposures were attempted.

Although virtual reality can offer several potential benefits beyond that of traditional cue-exposure paradigms, no studies have directly compared likeability and effectiveness between virtual reality and alternative cue-exposure paradigms, indicating a gap in this promising field. In order for virtual reality to be utilized as a cue-exposure paradigm in research and clinical settings, it has to be able to elicit craving comparable to or stronger than other traditional cue-exposure methods. If this is found to be true, then virtual reality can be used as an alternative providing the additional aforementioned benefits.

A further limitation of past research is that typically studies include only one way of measuring craving, usually self-report. Self-report is an important mechanism for assessing craving but can be influenced by possible bias or participant expectations. For instance, participants know that these studies include exposing them to drug-related cues and assessing their level of craving and hence, participants may feel pressure to respond in a way they think the experimenter wants or expects. It is advantageous to utilize multimodal ways of obtaining data regarding craving, including more objective measures such as physiological measures that are less susceptible to possible bias. This

is paramount given that verbal report, overt behavior, and somatovisceral responses are not necessarily strongly correlated (Rohsenow & Monti, 1999) but can all be related to craving and subsequent substance use.

2. Current Study

The current study adds to existing literature in the following six ways: 1) developing a virtual reality paradigm specifically for adolescents by obtaining their feedback during a focus group rather than assuming a paradigm designed for adults is adequate; 2) utilizing a sample of adolescents and targeting craving for cannabis, a commonly used substance among youth; 3) obtaining not only multiple self-report measures of craving but also more objective somatovisceral responses to cues; 4) directly comparing virtual reality with another cue-exposure paradigm on amount of craving elicited as well as acceptability and enjoyment; 5) directly comparing neutral cues to drug-related cues for each cue-exposure paradigm as well as across paradigms; 6) using a within-subjects sample, which maximizes power and minimizes statistical noise created by individual differences.

2.1. Hypotheses

Based on previous research, it was hypothesized that 1) for both cue-exposure paradigms (VR and video), drug-related cues would elicit more arousal (as indicated by increased skin conductance, increased heart rate, and decreased skin temperature) and self-reported craving than would neutral cues; 2) while both drug conditions would elicit increased arousal, the VR paradigm would elicit more arousal than the video paradigm due to increased immersion with a multi-sensory approach that has visual

and auditory components and realistic-looking environments; 3) youth would find the VR paradigm more enjoyable and acceptable than the video cue-exposure paradigm.

3. Methods

The data for the study was collected within Dr. M. Zachary Rosenthal's lab, which allowed for use of virtual reality technology and mechanisms of measuring physiological responses. Participants were cannabis abusing adolescents, aged 14 to 21 years old, who were receiving substance abuse treatment with one of three medical center-based providers. This study had three separate components, involving three separate samples and IRB-approved methods. Participants for all three parts were recruited from the Center for Adolescent Substance Use Treatment (CAST) located in the Duke Child Development and Behavioral Health Clinic (CDBH). As compensation for their time and energy, each adolescent received $10 for every hour spent completing research procedures and reimbursement for parking if driven to the lab. Gift cards were used rather than cash to reduce the potential risk that study funds could be used for the purchase of drugs or any illegal activity. This research was partly funded by the Aleane Webb Dissertation Fellowship, which awarded me $350 for the purpose of compensating participants in the study and the American Psychological Association (APA) Dissertation Research Award, which awarded me $1,000 for study-related expenses.

For each part of the study, youth were first told of the study by a CAST counselor who generally described the study to his/her clients and determined if they were interested in participating. The counselor provided the youth (and parents when youth were under 18) a copy of the consent form and the experimenter's phone number

to contact with any questions. Prior to/following a regular treatment session at CDBH, I met with the adolescent (and parents in the case of minors) to explain the study to them and discuss any questions. To reduce the potential for feeling pressured to participate and to ensure that youth (and parents when relevant) had the opportunity to make a thoughtful decision about participation, a week was given before re-approaching the youth about the study. At that time for those who wished to participate, using IRB-approved procedures, written informed consent was obtained for participants age 18 and older. For those under 18 years of age, written parental consent was obtained in addition to adolescent assent.

3.1. Study Part 1: Video Clip Paradigm Development

The first part of the study focused on testing the ability of video clips to elicit craving for cannabis and determining which video clips to use during Part 3 of the study when comparing VR to video. A within-subjects design was used and included showing youth (N=11) 12 video clips, which were montages of videos obtained from YouTube (with permission from the video owners).

3.1.1. Participants

A power analysis was conducted to determine the minimum sample size required to adequately test if the drug videos elicited more craving than the neutral videos. A study involving video cues for this population was used to obtain an estimated effect size of .68 for craving elicited by video neutral cues versus video drug

cues (Gray, LaRowe, & Upadhyaya, 2008). Power analysis based on this estimated effect size of .68 was conducted using G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007); results indicated that only four people would be needed to have .95 power when running a repeated-measures ANOVA. Eleven youth were recruited to strive for variation in age, educational level (high school, college, no school), race/ethnicity, and gender. The sample included 10 males who identified as Caucasian/White and one female who identified as Hispanic/Latino. Youth ages ranged from 14-20 years old. All youth were diagnosed with a cannabis use disorder by a licensed substance abuse counselor. Seven of the participants were attending high school, three were in college, and one was not attending school.

3.1.2. Materials

3.1.2.1. YouTube videos

Several databases were searched to find studies that included cannabis-related cue-exposure videos. Attempts were made to obtain the video exposures used in prior studies but one person stated that his lab no longer had the videos and the videos provided by another study researcher were too short to be the only videos used in this study (35 seconds each)(Eastwood, Bradley, Mogg, Tyler, & Field, 2010). The Duke University Department of Psychology and Neuroscience research librarian also could not locate research-related videos with cannabis cues and hence, she suggested that YouTube videos be used. Using YouTube videos offered a potential advantage beyond

that afforded by past study videos such as having videos with the same themes (e.g., smoking out of a bong, playground scene, having varied ethnicities shown) and timing (e.g., duration of scenes) as the VR paradigm, which provides a stronger comparison condition. Consultants in the Duke Multimedia Project Studio and an IT analyst within the Duke Department of Psychology and Neuroscience assisted with providing software to convert and edit the YouTube videos in order for several of them to be combined and altered to be a certain length.

Each video clip was three minutes long in order to match the duration of scenes in the VR paradigm to which they were later compared. Eight of the clips were drug-related (depicting cannabis paraphernalia, people preparing to use cannabis, and/or people using cannabis) and four were neutral clips (showing images of fruit, drawing, and walking).

3.1.3. Procedure

All 11 youth individually viewed all of the 12 video clips. The video clips were shown across four separate sessions to reduce the likelihood of carryover craving. At the beginning of each session a baseline craving rating was obtained. Following this baseline rating, the participant viewed two drug video clips separated by a neutral clip. The drug clip presentation order was randomized using an online random numbers generator to avoid the possibility that the order of paradigms could make any difference in the results. Hence, each participant viewed a unique order of video clips. At the end of each

video clip, participants were asked to rate on a 7-point Likert scale their current level of craving to use cannabis. The term "marijuana" was used throughout the study with the adolescents rather than other terms such as "cannabis" based on a literature review of scales targeting youth drug use and consultation with a substance abuse counselor who specializes in adolescent substance abuse treatment. After viewing each video, the youth were also asked what their highest craving rating during the video had been. Craving ratings were used to determine if the drug clips were statistically more craving-inducing than the neutral clips. After each session, participants met with a licensed substance abuse counselor until their craving level returned to baseline.

Separately, a group of five substance abuse experts rated the video clips for acceptability and usefulness for inducing craving. Two of the experts were substance abuse researchers within Duke University Medical Center, two were CAST substance abuse counselors, and one was a counselor from outside of Duke who specialized in adolescent substance abuse. The group of five experts were gathered together and shown the 12 videos presented in a randomized order based on an online random numbers generator. They individually rated each video on a 7-point Likert scale for how useful/relevant the video would be to elicit craving among adolescent cannabis-abusers. Discussion of videos was postponed until after all ratings were made to avoid creating bias among raters.

3.2. Study Part 2: Virtual Reality Paradigm Development

The second part of the study included a focus group and pilot participants to obtain qualitative information regarding tailoring drug-related stimuli and measures of virtual reality tolerability.

3.2.1. Participants

A group of five youth were recruited from the CAST program. Attempts were made to recruit youth of various ages, educational levels, gender, and race/ethnicity. The youth ranged in age from 14-18 years old and included one female. Despite efforts to recruit ethnically diverse participants, all youth included in this portion of the study were Caucasian/White due to the makeup of CAST at the time. Four of the participants were in high school and the other youth was not currently enrolled in school. A college undergraduate student was also originally included to represent this population of CAST but the day before the focus group was scheduled to be held, he stated that he could not participate and hence, no college students were included in the focus group.

3.2.2. Materials and measures

3.2.2.1. Virtual reality software

The VR system was developed for a previous National Institute of Drug Abuse (NIDA)-funded project conducted within Dr. Rosenthal's lab. The system is connected to speakers and the program is equipped with sounds spanning both the environment (e.g., cars driving by) and characters (e.g., conversation when approached). The VR

program has the participant begin standing on a sidewalk with the ability to turn and walk down one of two possible streets. Although the program allows the participant to be able to navigate through the VR world and several specific scenes using a joystick, for this study participants viewed a fixed path rather than navigating on their own. While using a fixed path may reduce immersion and engagement, it limits potential confounds such as navigation errors. For instance, when navigating in a virtual reality paradigm, participants can experience aimless movement such that they move around the same position without being able to move through the virtual scene, or they may experience inability to move whereby participants become stuck between virtual objects or next to walls. Participant-controlled VR runs and navigation errors can also introduce the potential for participants to have substantial variation in their time being spent exposed to the virtual reality world and each scene. This would become an added factor to consider when interpreting the data and could influence the comparison to video conditions since the length of the videos was matched to control for the time spent being exposed to VR cues.

In addition to there being both indoor (e.g., bar/restaurant, apartment) and outdoor (e.g., sidewalk, alleys, playground) scenes, there are over 30 modifiable stimuli and characters that can be placed throughout the VR world. The scenes, potential stimuli, characters, and dialogue were developed by Psychology Software Tools, Inc. in collaboration with Dr. Rosenthal and his colleagues.

3.2.2.2. Immersion

Witmer, Jerome, and Singer (2005) developed the Presence Questionnaire, Version 4 (PQ), a 33-item self-report measure of responses to the VR world. The PQ includes factors thought to contribute to a sense of being present in the VR world including factors of: involvement, sensory fidelity, adaptation/immersion, and interface quality. The PQ was administered after completing the VR paradigm. The PQ has been described as having high internal consistency reliability with a Cronbach's alpha of .91 (Witmer et al., 2005). The scale is widely used in VR research, particularly in VR paradigm development (e.g., Vora et al., 2002). Total scores range from 33- 231 with higher scores reflecting a greater sense of presence in the virtual environment. Ranges for the factor subscales are: Involvement from 12-84, Sensory Fidelity from 6-42, Adaptation/Immersion from 8-56, and Interface Quality from 3-21.

Based on data from running over 1,500 participants (including adolescents) through two VR worlds, Bangay and Preston (1998) determined that prior experience with VR, excitement, comfort, control, desire to repeat the experience, and quality of sounds and images are associated with immersion. Because, of these variables, the PQ directly addresses only excitement and control, the scale was supplemented by an additional four questions to address: 1) desire to continue interacting with the program, 2) quality of sounds, 3) quality of images, and 4) comfort/acceptability of the VR program.

3.2.2.3. Tolerability and likeability

No existing relevant scales could be found to measure tolerability and likeability. Hence, 12 questions were created to assess these desired domains and as already noted, to further assess immersion. Participants were asked to select for each question their level of agreement with the statement choosing from four options: Strongly Disagree, Disagree, Agree, and Strongly Agree (see Part 2 Additional Questions in Appendix A).

3.2.2.4. Simulator sickness

The Simulator Sickness Questionnaire (SSQ) is a 29-item self-report scale developed by Kennedy, Lane, Berbaum, and Lilienthal (1993) and most recently updated in 2001. It measures three sets of possible symptoms people may experience after being exposed to VR including: Oculomotor (e.g., eyestrain, blurry vision, headache), Disorientation (e.g., dizziness), and Nausea (e.g., nausea, stomach ache, increased salivation). These clusters were identified by the scale developers through use of factor analysis. Respondents rate how much each of the symptoms is presently affecting them on a four-point scale (0=none, 1=slight, 2=moderate, 3=severe). The Total Severity score uses all of the symptoms. The SSQ has been described as having high split-half correlation (r= .80) as well as good internal consistency reliability with a Cronbach's alpha of .78 (Kennedy et al., 2003). Further, the SSQ appears to be widely used with no psychometrically superior comparable scale available. Four paired samples t-tests were conducted to compare pre- and post- scores on the SSQ three subscale scores and total

score. A Bonferroni-corrected p level of 0.0125 (0.05/4) was used to protect against inflation of Type I errors produced by multiple comparisons. There were no significant differences in any of the pre to post SSQ scores implying that the VR paradigm did not cause sickness or any of the measured side effects.

3.2.3. Procedure

Participants in this portion of the study were first gathered to be part of a focus group to provide feedback on the empty virtual reality paradigm. Subsequently, these same youth were invited to individually visit the lab within the Duke University Medical Center (DUMC) where the virtual reality (VR) system is located to provide feedback on the feasibility and tolerability of the developed VR world.

The five person pilot group was used to assist in determining what stimuli would be important and useful to have in a cue-exposure paradigm for adolescent cannabis-abusers. Before beginning the group, each youth signed an IRB-approved confidentiality agreement that explained the importance of confidentiality in this study. Specifically, the form highlighted the importance of refraining from sharing with anyone group member names or identifying information as well as any information about what was about to be viewed. This was important for all participants to feel maximally comfortable with participating and speaking freely and to limit potential biasing of others in treatment with group participants who may complete Part 3 of the study.

After signing the confidentiality agreement, the group of youth were shown a video of the empty VR world depicting its various scenes including: a park/playground, alleyway, an apartment located inside an apartment building, an apartment with an exterior entrance, bar/restaurant, a motel, and a middle-class house as well as an abandoned house situated in a neighborhood of lower income houses (e.g., Figure 1).

Figure 1: Screenshot of the VR neighborhood with the middle-class and abandoned house.

After viewing the empty VR world, the youth were asked to rank (in order) the 8 scenes for "how relevant you consider it as a place you or others your age would use marijuana". This information was used to determine which scenes were to be included in the VR presentation shown in Part 3 of the study. After individually ranking the

scenes, as a group the youth were asked for feedback, suggestions, and what stimuli should be included in the scenes to enhance craving and relevance (e.g., drug-related paraphernalia, specific characters). Next, the youth were given a form with a pre-generated list of possible stimuli and asked to rate on a 7-point Likert scale, "how important or useful do you think it is to include this item to create craving/ urges to use marijuana among adolescent marijuana users?". Blank spaces were left on the list for participants to add any additional stimuli or fill in any stimuli the youth may have generated during their discussion that were not already on the list; participants were asked to similarly rate these stimuli. Each participant added 10-12 stimuli.

Participants were escorted from CDBH to the lab in the building where the VR system is located and the group was held; they were then escorted back to CDBH to ensure that youth were in a controlled environment at all times and able to speak to licensed substance abuse counselors to assist in managing the craving after the group. Approximately one week later, four of the pilot group youth were individually escorted back to the lab. One pilot group participant could not complete this part of the study due to returning to Wilderness treatment for a higher level of care a few days after the pilot group was held (for use that was occurring prior to the focus group rather than being triggered by the group).

During the individual visit, each youth completed a run through the VR drug world that was developed based on the pilot group's suggestions, completed

questionnaires, and gave further feedback. The SSQ was administered both before and after the VR run. Obtaining SSQ data before the VR run provided comparative baseline information for interpreting findings (e.g., if participants were nauseous before the VR run, it can be assumed the VR paradigm did not cause the nausea). The PQ was completed after the VR run. Based on results from the pilot group and uncertainty about which scenes to include, also following the VR presentation, youth were asked to choose between two scenes shown (abandoned house and apartment). The run took approximately 8 minutes and completing questionnaires as well as providing feedback to the experimenter took approximately 22 minutes. The visit occurred before or during the youth's treatment visit to ensure that if the VR paradigm elicited craving, the youth would be returning directly to treatment and have the ability to discuss and process the craving with a licensed substance abuse counselor.

Piloting the VR presentation served the purpose of obtaining information about immersion and tolerability. Merriam-Webster defines immersion as "absorbing involvement" ("Immerse," n.d.). Immersion within a VR world has several components including mental, emotional, visual, and perceptual immersion (Robertson, Czeminski, & van Dantzich, 1997). Immersion is important to measure and attempt to improve because it has been postulated that immersion is one of the benefits of VR (Robertson et al., 1997), which makes it more captivating and potentially useful as a cue-exposure

paradigm. After completing the VR run, participants were asked to complete the PQ assessing their immersion in the virtual environment.

Additionally, for VR to be useful as a cue-exposure paradigm, participants must find it feasible and tolerable. After completing the VR run, participants were asked to respond to questions directly asking about tolerability and enjoyment. In regard to tolerability, as described in a meta-analysis of VR studies for anxiety exposures, one of the reported ill effects of VR is simulator sickness, which can include symptoms of nausea, disorientation, eyestrain, and headaches (Gregg & Tarrier, 2007). It is important to have a measure assessing these symptoms because results from the meta-analysis further indicated that experiencing simulator sickness can influence behavior and performance during the VR task (Gregg & Tarrier, 2007). Pilot participants were asked to fill out the SSQ, a measure of simulator sickness both before (as a control) and after completion of the VR run. Past research indicates that simulator symptoms are typically mild and short lived (e.g., Cobb, Nichols, Ramsey, & Wilson, 1999). It was planned that if participants reported notable symptoms after completing the experiment, they would be asked to sit in a separate room until the symptoms subside prior to leaving the lab building. However, no participants required this procedure.

3.3. Study Part 3: Virtual Reality Paradigm versus Video Paradigm

There were two primary aims of this portion of the study. First was to test the ability of the VR cue-exposure system to elicit craving (compared to a VR system with

only neutral cues) among a population of adolescent cannabis abusers. Given the anticipated hope that virtual reality could eventually be used to augment standard drug treatments, eliciting craving is an essential step toward using the system as a therapeutic exposure technique. Second, to make further development of VR paradigms for substance abuse worthwhile, it is important to evidence that VR can elicit at least as much craving as another more traditional cue-exposure paradigm that uses video clips to elicit craving. It was expected that VR would have added benefits above and beyond those afforded by less immersive cue-exposure paradigms, such as increased enjoyment and increased arousal (possibly due to increased immersion), both of which could make using VR in treatment beneficial.

3.3.1. Participants

For this third portion of the study, power analyses based on effect sizes found in relevant prior research were conducted using G*Power 3 (Faul et al., 2007) to determine the target sample size. Power analyses were conducted for both comparisons: 1) virtual reality neutral cues versus virtual reality drug cues and 2) virtual reality drug cues versus video clip drug cues. In a previous study examining the comparison of neutral versus drug cues in virtual reality with adult cannabis abusers, researchers found Cohen's *d* effect sizes ranging from .85 to 1.29 (Bordnick et al., 2009). Results from G*Power 3 for a repeated measures, within factors MANOVA determined that based on the lowest effect size of .85 for the comparison of neutral versus drug cues for 20

measurements (accounting for all of the dependent variables measured over time) , 22 people would be needed to obtain a .95 estimate of power.

Studies directly comparing virtual reality and video craving elicitation for adolescent cannabis abusers could not be found. However, a study involving video cues for this population was used to obtain an effect size of .68 for craving elicited by video neutral cues versus video drug cues (Gray et al., 2008). Per the advice of a statistician consultant at Duke University, this video effect size of .68 was subtracted from the average virtual reality effect size ([1.29 - .85] / 2 = 1.07) to yield an estimated effect size for use in a power analysis of the comparison of craving in virtual reality versus video. Calculating an estimated effect size this way allowed for both VR and Video effect sizes to be accounted for and ensured that the subsequent power estimates were being made based on a conservative estimate of effect size. Results from G*Power 3 determined that based on this effect size of .39 (based on 1.07 - .68) for the comparison of virtual reality versus video paradigms using a within factors MANOVA, 24 people would be needed to obtain a .95 estimate of power. Based on these power analyses, it was determined that recruiting 26 youth for the third part of the experiment would provide sufficient power for the proposed analyses.

In order to be included in the study, adolescents had to be receiving treatment for substance abuse with a licensed mental health provider or substance abuse counselor. Youth were excluded if: (1) their substance use was severe enough to warrant

immediate hospitalization; (2) their substance use disorder did not include cannabis as a primary substance; (3) they were experiencing an acute untreated psychiatric or medical condition that required immediate attention or would inhibit full participation in the study; (4) they exhibited insufficient mental capacity to comprehend the consent form or participate in the study; or (5) they had a history of self-reported motion sickness. In an effort to increase generalizability, adolescents were not excluded for receiving concomitant medication.

Attempts were made to recruit youth of various ages, educational levels, gender, and race/ethnicity. The youth ranged in age from 15-20 years old (*M*= 16.92) and included three females. Approximately 73% of youth were enrolled in high school at the time of the study; five participants were enrolled in college and three participants were not currently enrolled in school. The sample was comprised of predominantly White/Caucasian (76.9%) youth with 19.2% identifying as Black/African-American and 3.8% identifying as more than one race. Hollingshead socioeconomic status scores ranged from 22-61 (*M*=50.13) out of a possible range of 8-66; the most frequent scores were 50.5 and 58 with 11.5% of the sample reporting each of these scores implying that the sample was comprised of predominantly upper-class youth. Age of first cannabis use (based on participant self-report) ranged from 9-18 years old and recency of cannabis use ranged from two days prior to the study to 10 months prior to the study

(per participant self-report). Approximately 42% of the youth had last used cannabis within the past month prior to completing the study.

3.3.2. Materials and measures

3.3.2.1. Cue stimuli

The experiment included two paradigms (VR and Video) and each paradigm had two cue types (Neutral and Drug) yielding four tasks. All participants completed all four tasks. In the VR neutral task, the participant only saw neutral stimuli such as a tape measure and bananas. The VR drug task included drug-related stimuli in each scene. For instance, this task included avatars smoking a joint outside and a party in a house complete with loud music, small talk with characters, and characters using cannabis. The determination of stimuli as neutral or drug was made based on the focus group ratings from the previously described second part of the study. Given that stimuli had been rated on a 0-6 scale with 0 being "not at all" and 6 being "the most it could be", items with averaged ratings of less than 2 were considered neutral whereas items with averaged ratings of 3 or greater were considered drug stimuli.

Based on past studies (e.g., Bordnick et al., 2009), youth were exposed to the various scenes for three minutes each and a total of three scenes were included in each of the VR tasks. The same three scenes were used in both the neutral task and drug task; the scenes only differed in which stimuli were included. Attempts were made to match the number of stimuli in each scene as well as their usefulness ratings. Similarly, the

approximate number of stimuli in the neutral task was matched to the number of stimuli in the drug task to control for the possible effect number of stimuli included could have on arousal and craving ratings. The location of stimuli across both the neutral and drug VR tasks was also matched (e.g., if an object was placed on the kitchen counter in the drug task, an object was also placed on the kitchen counter in the neutral task). A total of approximately 10 minutes was spent completing each VR task (three minutes per scene plus time for walking between scenes). Consistent with other studies, the video task included brief video clips. The three neutral clips excluded drug references (e.g., depicted people walking or drawing) whereas the three drug-related clips depicted paraphernalia and/or people preparing and using cannabis. To control for the time spent being exposed to VR cues, video clips were also shown for three minutes each. A 20 second blank screen was shown in-between the video scenes to control for the 20 seconds spent walking between scenes in the VR tasks.

In order to avoid the possibility that the order of paradigms could make any difference in the results, the paradigm order was counterbalanced and randomized using an online random numbers generator: half of the participants completed the Video paradigm first and then the VR paradigm, and the other half in the reverse order. Consistent with past research and procedures used by Gray and colleagues (2008), tasks with neutral stimuli were presented first in order to prevent possible carryover effects from presenting drug-related stimuli first (Monti et al., 1987).

3.3.2.2. Physiological measures

Heart rate, skin conductance, and skin temperature were measured during the study using Ag-AgCl electrodes and tape. Two heart rate electrodes embedded in adhesive pads were attached under the right clavicle and the left lowest rib. This allowed for ECG measurement, which measures the rate and regularity of heartbeats and can be transformed to yield data showing beats per minute as it is commonly analyzed. Two skin conductance electrodes were placed on the middle phalanx of the middle and ring fingers of the participant's non-dominant hand using Signa Electrode Gel (for enhanced signal) and were secured with tape to ensure that random movements did not create noise in the data. Skin conductance electrodes allowed for galvanic skin response (GSR) measurement, which is expressed in microSiemens units. This was useful to measure since GSR varies with moisture level and the amount of sweat produced can represent an indicator of psychological and physiological stress. Hence, skin conductance was expected to increase when the cannabis-abusing youth were shown drug-related stimuli that were expected to produce craving. The skin temperature electrode was placed on the wrist of the participant's dominant hand using Signa Electrode Gel and tape. Physiological data was acquired using the Acq*Knowledge* software program distributed by BIOPAC Systems Inc. This program allows for real-time observation of the data acquisition, which means any abnormalities noticed during

baseline measurement (e.g., unexpected spikes possibly due to a loose electrode) could be addressed prior to further data collection.

3.3.2.3. Craving, emotional arousal, and emotional valence

Self-reported craving was measured in two ways: using the Marijuana Craving Questionnaire (MCQ) and a single-item craving rating. The MCQ is a validated 12-item assessment to characterize cannabis use and craving (based on the longer 47-item MCQ; Heishman et al., 2001), that takes less than five minutes to complete. Each item is rated on a 7-point Likert scale. The MCQ has four craving subscales: Compulsivity, Emotionality, Expectancy, and Purposefulness.

Self-reported arousal and pleasure were assessed through use of the arousal and pleasure dimensions of the Self-Assessment Manikin (Bradley & Lang, 1994). Both dimensions have human representations/manikins exhibiting variation on a 1-9 scale; the arousal scale shows a mass increasing in size and sharpness in the center of the manikin as the numbers increase and the pleasure scale shows a change in the facial expression of the manikin as the numbers increase (starting with a frown and increasing to a wide smile). On the arousal dimension, respondents were asked to mark on a "calm-excited scale" how emotionally aroused they feel right now with "1" corresponding to completely calm, sleepy, or unaroused and "9" corresponding to completely excited, jittery, or aroused; they were asked to select "5" if neither excited nor calm. Similarly, on the pleasure dimension, respondents were asked to indicate on a "unhappy-happy

scale" how happy or unhappy they feel right now with "1" corresponding with completely unhappy, bored, or annoyed and "9" corresponding to completely happy, pleased, or satisfied; participants were asked to select "5" if feeling completely neutral. Obtaining this information allowed for consideration of feelings of arousal or emotional mood on subjective ratings of craving or physiological data.

In order to have multiple measures of self-reported craving, participants were asked to respond to a single-item analog scale assessing "how strong is your urge to use marijuana right now?" on a scale of 0-10, with '0' being absolutely no urge and '10' being the strongest urge they have ever experienced. This 0-10 craving scale is what the CAST counselors use with the youth in treatment and hence, was the scale the participants were most familiar with and easily able to respond to during a quick assessment. Using just this one item has been shown to be highly sensitive to craving manipulations and is easier to administer repeatedly than are longer scales (Juliano & Brandon, 1998; Sayette, Martin, Hull, Wertz, & Perrott, 2003). For both VR and Video tasks, this question was verbally posed to participants and they were asked to then respond verbally. This technique enabled the participant to remain engaged in the task without removing his/her eyes from the images being presented whereas a paper and pen method would have required more attention away from the task.

3.3.2.4. Immersion and presence

Witmer and Singer (1998) developed The Immersive Tendencies Questionnaire, Version 3.01 (ITQ), which is a 34-item self-report measure of "the capability or tendency of individuals to be involved or immersed" (p. 230). The items inquire about a person's involvement in several common activities. Participants are instructed to rate each response for how frequently they do or how well they do on each of the situations described on a 7-point Likert scale ranging from doing "not at all" (1) to "very well" (7) or "never" (1) to "often" (7). An example from the ITQ is "When playing sports, do you become so involved in the game that you lose track of time?" Choosing a higher number indicating greater frequency would be a response consistent with immersion. The scale has a total summary score as well as subscale scores for Focus, Involvement, and Games. The Games subscale is comprised of two items which ask about frequency of playing video games and how often the respondent feels he/she is inside the game rather than using a joystick and watching the screen. The already described PQ was also included in this portion of the study. Both the PQ and ITQ have been described as having high internal consistency reliability with Cronbach's alphas of .91 and .75, respectively (Witmer et al., 2005; Witmer & Singer, 1998). The scales are widely used in VR research, particularly in VR paradigm development (e.g., Vora et al., 2002).

3.3.2.5. Simulator sickness questionnaire

As already described in Part 2 of the study, the Simulator Sickness Questionnaire (SSQ) is a 27-item scale assessing side effects that may occur after exposure to virtual reality. The scale includes three subscales (Nausea, Oculomotor, and Disorientation) as well as a total score. Four paired samples t-tests were conducted to compare pre- and post- scores on the SSQ three subscale scores and total score. A Bonferroni-corrected *p* level of 0.0125 (0.05/4) was used to protect against inflation of Type I errors produced by multiple comparisons. There were no significant differences in any of the pre- to post-SSQ scores implying that the VR paradigm did not cause sickness or any of the measured side effects.

3.3.2.6. Additional participant information

To obtain data to control for potential confounds, participants were also asked for demographic and descriptive information including: age, years of education, gender, race/ethnicity, socioeconomic status (SES), when they first used cannabis, and when they last used cannabis. A CAST counselor provided consultation in the development of this questionnaire and what potential confounds should be included. A total SES score was computed based on a four-factor scale of mother and father (or guardian) occupations and levels of education (Hollingshead, 1975); information used to calculate SES was obtained by the participant's parent or guardian if youth still lived at home. Participant responses were corroborated with the youth's substance abuse counselor.

To additionally obtain information on comorbid conditions as potential confounds, The Achenbach System of Empirically Based Assessment (ASEBA) scales were used to examine comorbidity. The Child Behavior Checklist (CBCL/6-18) was used for youth under 18 years old (n= 17) whereas the Adult Behavior Checklist (ABCL/18-59) was used for youth over 17 years old (n=9). Both scales ask the informant to complete a checklist of items that describe a person now or within the past six months and to indicate if the item is very true/often true of the person, somewhat/sometimes true of the person, or not true of the person. The items listed inquire about potential behavioral, emotional, and social problems. When possible, the form was completed by the youth's parent or guardian (n=16). For college students whose parents were not involved in treatment, the CDBH counselor completed the form (n=8). There was minimal missing data; parents of two participants did not return their scales (one CBCL and one ABCL). The CBCL and ABCL yield syndromes scales as well as DSM-oriented scales consistent with DSM-IV categories in addition to summary scores including an internalizing score and externalizing score. The CBCL has been shown to have good test-retest reliability (intraclass correlation coefficients of at least .95) and internal consistency (Cronbach's alpha ranged from .72 to .97 for subscales) (Achenbach & Rescorla, 2001). Similarly, The ABCL has been found to have good test-retest reliability (intraclass correlation coefficients of at least .73) and internal consistency (Cronbach's alpha ranged from .71 to .97 for subscales) (Achenbach & Rescorla, 2003).

The Brief Sensation Seeking Scale (BSSS; Hoyle, Stephenson, Palmgreen, Lorch, & Donohew, 2002) was also administered to assess sensation seeking as a potential confound. The BSSS asks respondents to indicate on a 1-5 point scale how much they agree or disagree to each of eight items with 1 being "strongly disagree", 3 being "neither disagree or agree", and 5 being "strongly agree". The BSSS yields a total score as well as four content domain scores including: experience seeking, boredom susceptibility, thrill and adventure seeking, and disinhibition. As described by Hoyle and colleagues (2002), in a study of more than 7,000 adolescents, the authors found that the BSSS yielded good item characteristics and internal consistency (alpha= .76). Further, in a second study of over 6,000 adolescents, they found that, "[s]cores on the full BSSS correlated inversely with negative attitudes toward drug use and positively with drug use" (p. 401). The BSSS scores were also reliably and significantly positively correlated with other risk factors and negatively correlated with protective factors.

Further, to ensure participants were not intoxicated at the time of the experiment, similar to procedures used by Bordnick and colleagues (2009), a saliva sample was obtained from participants to undergo a rapid alcohol and drug screen prior to completion of study procedures. Alco-Screen 02 saliva alcohol test strips (Department of Transportation, DOT-approved test) were used to detect alcohol and SalivaConfirm-4 Four Panel Saliva Drug Tests were used to detect four common drugs of abuse: cannabis (THC), cocaine, opiates, and methamphetamines. Results for both sets of tests are visibly

obtained within ten minutes. Saliva samples are preferable because they are quick and do not require the supervision that would be needed to ensure a urine sample was unadulterated. If the results indicated that participants had any of these drugs in their system (excluding prescribed medications, which were verified with the substance abuse counselor or parent beforehand), they were to withdraw from the study, re-consent, and re-enroll to participate in the study at least two weeks later; they were not to be reimbursed for time spent completing the drug test. Participants were made aware of these study procedures prior to consenting to participate in the study. No participants had to be withdrawn due to a failed drug or alcohol test. Two participants had positive drug tests due to prescribed medications; one for a stimulant medication and one for an opiate he was prescribed by the emergency room following a car accident two days before the study was conducted.

3.3.2.7. Additional questions

As shown in Appendix B, to assess tolerability and likeability as well as to compare the paradigms on several domains, participants were administered additional questions. Ten questions inquired about the two paradigms. Each of the 10 questions included a statement regarding the Video paradigm as well as a statement regarding the VR paradigm and asked participants to rate their level of agreement with the statement. The questions asked about interest, positive reaction, enjoyment, acceptability, willingness to use in treatment, whether use of the paradigm would increase interest

and/or likelihood of staying in treatment, effectiveness at eliciting craving, side effects, how realistic the paradigms were, and interest in continuing with the paradigm. To reduce the number of questions participants would have to answer at the end of the study and increase the likelihood that participants would carefully read the questions, the first three questions included on the Part 3 Additional Questions form were given directly after viewing the presentation rather than with the other questions at the end (see Part 3 Additional Questions [after each task example] in Appendix B). The additional questions form also asked participants how often they play video games due to the possibility that frequency of video game playing could influence the results. On this form, participants were also asked which paradigm they prefer and why, as well as whether they have any suggestions on how to improve the VR program with space for additional comments.

3.3.3. Procedure

Youth for this portion of the study were also recruited from the CAST program and were referred by their substance abuse counselor. As already described, youth were consented using IRB-approved procedures at least one week after first having the study described to them. Adolescents were told that their answers and responses throughout the study were kept confidential to ensure accurate reporting of substance cravings. Those who volunteered were scheduled to meet at the Virtual Reality Lab in the

Cognitive Behavioral Research and Treatment Center for a single cue-exposure study of approximately 1 hour and 45 minutes just prior to a counseling session.

Upon first arriving to the lab, youth were asked to provide a saliva sample to complete the rapid alcohol and drug tests. While waiting the ten minutes for the results, youth had the procedures for the experiment described to them and were then asked to complete scales including the BSSS, ITQ, initial MCQ, and additional information form. Youth completed the SSQ just prior to completing the VR tasks; hence, half of the youth completed the scale at this time whereas the other half completed the scale after viewing the nature slideshow. Subsequently, I connected participants to the physiological measures. Physiological indicators of heartbeat, skin temperature, and skin conductance were recorded. A baseline for the three physiological measures was obtained for five minutes while the participant sat quietly with eyes open directly in front of a large wide-screened monitor. Participants then completed the previously described four tasks during which they made several ratings of their current level of craving. During the experiment, I sat off to the side and behind participants in order to be able to verbally pose the single-item craving question throughout the tasks. Throughout the study, physiological recording was linked to the VR simulation by creating markers, in order to temporally associate physiological responses with exposure to cannabis cues.

In order to minimize carryover effects, participants were shown five minutes of a nature slide show in-between the two paradigms (VR and Video). If after the five

minutes, participants were still experiencing craving greater than ‘2’ (on a 0-10 scale with 0 being no craving and 10 being the most ever), five additional minutes were shown. This duration of time was chosen because during the first two parts of the study, most of the youth had their cravings dissipate within five minutes of being exposed to triggering cues and all youth had their cravings subside within 10 minutes. The level ‘2’ was chosen instead of ‘0’ due to the fact that several youth in the prior two phases reported that their constant "baseline" craving is a ‘2’ rather than a ‘0’. If participants continued to experience craving above a level ‘2’; they were led through a deep breathing exercise. These procedures allowed the participants to rest after being exposed to drug-related stimuli before viewing the next set of neutral cues.

After completing the VR tasks, the participants completed the PQ and SSQ. At the end of the experiment, the participants answered questions regarding the two paradigms (VR and Video; see Additional Part 3 Questions in Appendix B).

Participants were then thanked for their participation and given their gift card as compensation. This experimental part of the study took approximately 1.75 hours. Afterwards, participants were escorted to meet with their licensed substance abuse counselor at CDBH. The counselor assisted participants in processing any residual cravings until their craving levels returned to baseline. Table 1 outlines when during the experiment each of these measures were administered.

Table 1: Description of when measures were administered during Part 3 of the study

	Timepoint										
Measure	At start of the expt/ Before Task 1	During All Tasks	Post-Task 1	Pre-Task 2	Post-Task 2	Nature Slideshow	Pre-Task 3	Post-Task 3	Pre-Task 4	Post-Task 4	End of expt
ASEBA scale	X										
Saliva tests	X										
Additional Information	X										
BSSS	X										
ITQ	X										
SSQ	X (pre-VR)									X (post-VR)	
SAM	X		X		X		X	X		X	
MCQ	X		X		X		X	X		X	
Craving	X	X	X	X	X		X	X	X	X	X
Physio	X	X	X	X	X		X	X	X	X	
Additional Questions										X	
Presence Questionnaire										X (post-VR)	

Note: Expt= Experiment, Post= After, Pre= Before, Saliva tests= Saliva drug and alcohol tests, BSSS= Brief Sensation Seeking Scale, ITQ= Immersive Tendencies Questionnaire, SSQ= Simulator Sickness Questionnaire, SAM= Self-Assessment Manikin, MCQ= Marijuana Craving Questionnaire, Craving= Single-item craving measure, Physio= Physiological measures.

4. Results

4.1. Study Part 1: Video Clip Paradigm Development

Part 1 included a sample of 11 youth who rated the 12 videos for their ability to produce craving. This portion of the study also included a separate sample of five substance abuse experts who rated the videos for their usefulness and relevance in eliciting craving in adolescent cannabis abusers. Three of the experts were youth substance abuse counselors (one male and two females) and two experts were substance abuse researchers (both male). Tables 2 and 3 include information regarding the 12 videos.

Table 2: Descriptions of the videos

Video	Description
Drug 1	Paraphernalia only, no drug use
Drug 2	Playground/outdoor scenes. Includes skateboarding as well as sitting and using
Drug 3 Drug 5 Drug 6 **Drug 7**	Indoor (couch) use
Drug 4	Couch use and use in an abandoned house
Drug 8	Using in an alley
Neutral 1	Teaching the names of fruit, only fruit is shown
Neutral 2	Drawing a zendala, only hands are shown
Neutral 3	Walking through an upper-class Midwestern neighborhood
Neutral 4	Walking through a middle-class town

Note: Bolded videos represent the videos used in Part 3 of the study.

Table 3: Comparison of contents of the eight drug and four neutral videos

Video	MJ Use Prep	MJ	Bowl	Blunt Or Joint	Bong	Talking Alone	Talking To Others	Music	Group	F	M	In	Out
D1	X	X	X	X	X	X		X				X	
D2				X			X		X	X	X		X
D3				X	X		X	X	X	X	X	X	
D4	X	X	X	X	X			X	X		X	X	
D5		X		X	X		X	X	X		X	X	
D6	X	X			X			X		X	X	X	
D7	X	X		X	X			X	X	X	X	X	
D8				X				X	X	X	X		X
N1						X	X						
N2								X					
N3							X				X		X
N4													X

Note: MJ= Marijuana, Prep= Preparation, F= Female, M=Male, In= Indoors, Out=Outdoors, D= Drug, N= Neutral.

4.1.1. Youth ratings

A paired samples t-test was conducted to compare craving induction in drug video and neutral video conditions, using single-item craving scores. First, to account for baseline craving, change scores for each participant were calculated by subtracting the craving rating made immediately before viewing each video from the craving rating made after viewing that video. Next, an average of the craving change scores for the eight drug videos was calculated (creating an average drug craving change score) as was

an average of the craving change scores for the four neutral videos (creating an average neutral craving change score); these scores were used in the paired samples t-test. There was a significant difference in the craving change scores for drug (*M*= 1.78, *SD*=1.17) and neutral (*M*= -1.57, *SD*= 1.05) conditions, t(10)= 5.09, *p*< .001. These results indicated that the neutral videos were associated with lessened craving, whereas the drug videos were associated with increased craving; the drug videos induced significantly more craving than the neutral videos.

4.1.2. Expert ratings

The expert panel members were asked to rate all eight drug videos and all four neutral videos on a 0-6 Likert scale with '0' representing 'not at all' and '6' representing 'the most it could be' for each video's ability to elicit craving among adolescent cannabis abusers. A paired samples t-test was conducted to compare usefulness/relevance ratings in drug video and neutral video conditions. An average of the usefulness/relevance scores for the eight drug videos was calculated (creating an average drug usefulness/relevance score) as was an average of the usefulness/relevance scores for the four neutral videos (creating an average neutral usefulness/relevance score); these scores were used in the paired samples t-test. There was a significant difference in the usefulness/relevance scores for drug (*M*= 4.68, *SD*=0.66) and neutral (*M*= 0.80, *SD*= 1.02) conditions, t(4)= 17.97, *p*< .001. These results indicate that the drug videos were rated significantly more useful/relevant for eliciting craving than the neutral videos.

4.1.3. Video selection

Because the goal was to choose drug videos that would elicit the most craving and neutral videos that would elicit the least craving, each youth participant's ratings of the videos were ordered into ranks. Two Friedman tests (one for drug and one for neutral videos) were used as a non-parametric, one-way repeated measures analysis of variance by ranks. Results of the Friedman tests indicated that there was not a statistically significant difference among neutral video rankings, $\chi 2(3) = 3.75$, $p = 0.290$ but that there was a significant difference among drug video rankings, $\chi 2(7) = 14.27$, $p < .05$. Post-hoc analysis with Wilcoxon signed-rank tests was conducted on the eight drug videos. A Bonferroni correction was applied, resulting in a significance level set at $p < 0.006$. At this significance level, there were no significant differences between the eight videos. Even if a p level of .05 had been used, the only significant differences were between Video 8 and six of the other videos, with Video 8 receiving the lowest mean rank compared to the other videos.

Similarly, each expert panel member's ratings were ordered into ranks. Two Friedman tests (one for drug and one for neutral videos) were used as a non-parametric, one-way repeated measures analysis of variance by ranks. Results of the Friedman tests indicated that there was not a statistically significant difference among drug video rankings, $\chi 2(7) = 8.23$, $p = 0.313$ nor among neutral video rankings, $\chi 2(3) = 2.54$, $p = 0.468$.

Because results from the Friedman tests were inconclusive, youth and expert ratings were considered when determining which three drug videos and which three neutral videos should be selected for use in Part 3 of the study. Additionally, consideration was given to which videos best matched the VR stimuli when determining which videos to use. Neutral Videos 1, 2, and 4 were chosen for inclusion in Part 3. Neutral Video 1 included images of fruit and hence, providing a good match for the inclusion of fruit in the Neutral VR presentation. Similarly, Neutral Video 2 showed only hands drawing, similar to the VR presentations showing the participant's virtual hands manipulating objects. Neutral Video 3 and Neutral Video 4 both showed a person walking. Participant ratings indicated that Neutral Video 4 had a lower average craving change score (M= -2.05) than Neutral Video 3 (M= -1.55). Similarly, the expert panel rated Neutral Video 3 as more useful for inducing craving (M= 4.7) than Neutral Video 4 (M= 1). Hence, Neutral Video 4 was chosen for inclusion in Part 3 of the study.

Drug Videos 1, 2, and 7 were chosen for inclusion in Part 3 of the study. Consistent with past research, it was planned that one scene would include only drug paraphernalia and hence, it was planned that a Drug Video and Drug VR scene would include only drug paraphernalia. Drug Video 1 showed only drug paraphernalia and matched well the planned Drug VR scene that would include only drug paraphernalia (with no avatars). Similarly, Drug Video 2 included playground and outdoor scenes that matched the playground scene in the VR presentation.

Drug Videos 3, 5, 6, and 7 all included similar material (couch smoking). Drug Video 7 was chosen because, as shown in Table 2, it included all but one of the components that the other videos included as well as including a group scene that matched the party scene in the Drug VR presentation. Further, participant ratings indicated that Drug Video 7 had a higher average craving change score (*M*= 1.91) than Drug Video 5 (*M*= 1.82) and Drug Video 6 (*M*= 1.86). The expert panel rated Drug Video 7 as more useful for inducing craving (*M*= 5.4) than Drug Video 3 (*M*= 4.7), Drug Video 4 (*M*=5.1), Drug Video 5 (*M*= 4.6), and Drug Video 6 (*M*= 4.8). Drug Video 8 was not chosen since it had the lowest average craving rating (*M*= 0.91) out of all the videos as well as the lowest expert panel usefulness rating (*M*=3.5).

4.2. Study Part 2: Virtual Reality Paradigm Development

Part 2 included a sample of five youth who attended a focus group and were asked to rank eight possible scenes for scene relevance to adolescent cannabis use. They also individually saw the created VR drug presentation, after which they completed the SSQ and PQ and were asked for feedback.

4.2.1. Scene selection

Because participants were asked to rank all eight scenes, a Friedman test was used as a non-parametric, one-way repeated measures analysis of variance by ranks. Results of the Friedman test indicated that there was not a statistically significant difference among scene rankings, $\chi 2(7) = 9.733$, $p = 0.204$.

Although no significance was obtained with the omnibus Friedman Test, post-hoc analysis was done to explore how individual scenes compared to each other. Post-hoc analysis with Wilcoxon signed-rank tests was conducted on the five scenes with the highest mean rank. A Bonferroni correction was applied, resulting in a significance level set at $p < 0.01$ (.05/5). Friedman and Wilcoxon signed-rank tests provide interquartile range (IQR) scores: a measure of where the "middle fifty" is in a data set as well as upper and lower quartile scores. Median (IQR) rankings (with lower to upper quartile rankings in parentheses) for the Abandoned House, Other House, Playground, Apartment 1, Apartment 2, Restaurant, Motel, and Alley were 2 (1 to 4.5), 2 (1.5 to 4), 3 (1.5 to 7.5), 5 (3 to 6), 5 (4 to 7), 6 (3 to 7.5), 6 (3.5 to 6.5), and 7 (3.5 to 8), respectively. While the Abandoned House and Other House were the highest ranked scenes, given the similarities between the two houses and interest in both Apartment scenes, the latter two were compared. This was also done because the pilot group participants stated that while the two house scenes were well-designed, they found the layouts to be too similar. There were no significant differences between Apartment 1 and Apartment 2 ($Z = -1.414$, $p = 0.157$). Similarly, there were no significant differences between Apartment 2 and the Abandoned House ($Z = -2.414$, $p = 0.157$) suggesting that the Apartment 2 scene could be used in place of the Abandoned House. When the four participants were shown the Abandoned House scene and the Apartment scene and asked to choose one, two youth chose the Abandoned House and two chose the Apartment. Given this equal split and

that the statistical results favored the Abandoned House, the Abandoned House scene was used in the third part of the study.

These five youth also rated several stimuli for craving induction usefulness on a 0-6 Likert scale with '0' being not at all useful and '6' being maximally useful. Stimuli rated 0-2 were considered neutral and stimuli rated 3 or higher were considered drug-related. Based on these criteria, the youth rated 36 items as neutral and 33 items as drug-related. This determined which items were included in the Neutral VR task versus the Drug VR task in Part 3.

Stimuli were placed in Drug VR scenes to allow for one scene to include only paraphernalia with no avatars or talking (Abandoned House), one scene to include a party (Other House), and one scene to have paraphernalia and avatars who speak but with no party (Playground). This was done to be consistent with past research that considers the potentially differing effects of including avatars and social contact when creating craving. Consistent with the Drug VR presentation, the Neutral VR Abandoned House scene included neutral stimuli with no avatars (e.g., fruit, household products), while the Other House and Playground scenes included multiple avatars together.

While there was some overlap in what stimuli were included in each scene (playground, abandoned house, and other house), each scene included unique stimuli as well. For instance, all of the Drug VR scenes included a water pipe ("bong") but only the Playground scene included a marijuana roll, only the Abandoned House scene included

a glass pipe ("bowl"), and only the Other House scene included a marijuana plant. Efforts were made to include similarly-rated stimuli in each scene to ensure that one scene would not have all of the strongest craving-inducing stimuli.

Additionally, efforts were made to have approximately the same number of stimuli in each scene of both presentations. All three scenes in the Neutral VR presentation (Playground, Abandoned House, Other House) had the same number of stimuli in each scene. The scenes in the Drug VR presentation (Playground, Abandoned House, Other House) had approximately the same number of stimuli as each other included in each scene. Stimuli were positioned in the same scene locations across presentations. For instance, in the Drug VR Abandoned House scene, a bowl was placed on the floor of one of the rooms; in the Neutral VR Abandoned House scene, a wrench was placed in the same location. Of note, due to limitations in the number of drug stimuli available for use, stimuli that included alcohol and cigarettes were also included in the Drug VR presentation in addition to cannabis stimuli (based on participant craving induction usefulness ratings).

4.2.2. Presence

To assess the degree of presence of the VR environments, scores for participants on the four Presence Questionnaire (PQ) factor scales and overall total scores were calculated. The overall total mean score on the PQ was 136.88 (*SD*= 27.37). The mean score on the Involvement scale was 43.75 (*SD*= 13.40), on the Sensory Fidelity scale was

29.63 (*SD*= 5.02), on the Adaptation/Immersion scale was 33.50 (*SD*=9.12), and on the Interface Quality scale was 14.5 (*SD*= 2.96). A prior version of the PQ reported established norms among 152 college students for the Involvement/Control (*M*= 57.39, *SD*= 8.96) and Interface Quality (*M*= 14.65, *SD*= 3.40) factors scales (Witmer & Singer, 1998). While the Interface Quality average score for the virtual environment in the current study was similar to the established norms, the Involvement scale for the current study was more than one standard deviation lower than previously established norms. This may reflect the difference in the version of the scales or the age of the populations included. Conversely, this may reflect an expected difference. The Involvement scale inquires about perceived control of events in the virtual environment as well as how responsive the virtual environment was to user-initiated actions. Because, this study used a fixed VR path, which limited participant control, it is expected that this score would be lower than norms from studies that allow participants control over the virtual environment.

4.2.3. Additional questions

The created scale asking participants to respond to 12 statements regarding VR tolerability, likeability, and presence had good internal consistency reliability with a Cronbach's alpha of .86. Averages and frequencies were examined for participant responses to the 12 additional questions. The four participants who completed this portion of the study all rated "agree" or "strongly agree" with statements regarding VR

being interesting, enjoyable, an acceptable way to create craving, an effective way to create craving, and that they had a positive reaction to VR and would want to spend more time with it. They additionally all "agreed" or "strongly agreed" that they would be willing to use VR in treatment and that using VR would make them more interested in treatment and/or more likely to stay in treatment.

While all participants "agreed" or "strongly agreed" that the sounds were high quality, one participant "disagreed" with the statement that the visual layout was high quality. One person "disagreed" that VR would not have bad side effects; this participant may have been referring to craving as a side effect. Two people "disagreed" that VR was realistic. Due to this finding, efforts were made to incorporate participant feedback and suggestions to make the VR more realistic (e.g., where to place items) for Part 3 of the study.

4.3. Study Part 3: Paradigm Testing

Part 3 aimed to compare the VR and Video paradigms; it included a sample of 26 youth who viewed all four conditions (Neutral VR, Drug VR, Neutral Video, and Drug Video). Physiological (heart rate, skin conductance, skin temperature) and craving data were collected before and during the four conditions. For five-minutes before participants began the first condition, physiological data were collected while the participant was asked to sit still, which constituted a five-minute baseline that was used in analysis of the physiological data. Participants also completed a MCQ scale prior to

beginning the first presentation, which was used as a baseline score in analysis of the craving data. Single-item craving ratings were obtained before each scene as well as five times during each scene at approximately equivalent intervals based on when stimuli were presented. The single-item craving ratings obtained in the Video presentations were exactly matched for when they were obtained in the VR presentations. For the purposes of data reduction, an average craving rating was calculated for each of the four conditions representing the average of all single-item craving ratings made during that condition (e.g., Drug VR craving representing the craving ratings made during the three scenes of the Drug VR condition). These averaged craving ratings were used in all subsequent analyses.

Before beginning data analysis, all data entry was quality-checked by an individual unaffiliated with the study. This individual also quality-checked a third of the physiological data to ensure that filter and transformation procedures yielded similar data. The physiological data chosen for quality checking were determined based on use of an online random numbers generator.

4.3.1. Data analytic plan

Distributions of all possible covariates and outcome measures were examined for normality, skewness and kurtosis, as well as outliers. When needed, transformations and outlier accommodation methods were implemented. Two MANCOVAs were planned: one for self-report craving (MCQ and single-item craving) and another for

physiological data (heart rate, skin conductance, and skin temperature). First, in order to determine which variables to control for, correlations were conducted to determine which of the potential confounding variables were significantly related to any of the outcome variables. Those variables significantly related to any outcomes were to be controlled for in subsequent MANCOVA's. MANCOVA was selected to control for Type 1 error that would be accrued if models were tested separately and to allow for testing multiple covariates as well as multiple related dependent variables (e.g., single-item craving and MCQ scores).

4.3.2. Normality

Distributions of all continuous covariates and outcome measures were examined for normality via visual examination, Kolmogorov-Smirnov and Shapiro-Wilk analyses, as well as measures of skewness and kurtosis with cutoffs of +/- 1 for skew and +/- 3 for kurtosis considered acceptable. Potential outliers were assessed through extreme values analyses, visual examination of Boxplots, and calculation of z-scores (with a z-score greater than absolute value three indicative of an outlier).

Based on these techniques, the baseline MCQ Compulsivity score was found to have a non-normal distribution with elevated skew and kurtosis and one extreme outlier (participant 121). Visual examination of the participant's responses on the baseline MCQ evidenced that the participant's data were likely invalid. The participant had marked a '7' (the highest possible score) for every one of the 12 questions, which no other

participant did on any of the MCQ's completed throughout the study nor did this participant do on subsequent MCQ's, suggesting that this participant likely responded to this baseline MCQ without reading the questions and merely marking his response straight down the page. Hence, this participant's baseline MCQ data were excluded. The aforementioned normality examination was repeated excluding this participant's baseline data indicating that this variable was now within accepted normality limits.

Visual inspection of the univariate frequency distributions for two variables, average neutral VR craving and average neutral Video craving, indicated the presence of two outliers. While transformations of these variables could have been used to mitigate the unwanted influence of their values on the measures of central tendency for these variables, doing so would have changed the "metric" of these variables to something less familiar and easy to understand. As an alternative, a strategy of "outlier accommodation" known as "winsorizing", was employed to treat these outliers. More specifically, winsorizing involves recoding the outlying values to the closest, non-outlying observations in the frequency distribution and adding a small increment to the recoded score in order to retain the outliers' relative positions in the original frequency distribution (Clark, 1989). Following this procedure, skew and kurtosis were significantly reduced.

Among the covariate measures, three variables were found to have skew and kurtosis values outside of the accepted limits: Hollingshead SES, ASEBA Externalizing

Problems scale score, and recency of cannabis use; but this nonnormality was not attributable to the presence of extreme outliers for any of the three variables. Hence, transformations of the variables were conducted. Hollingshead SES and Externalizing Problems were negatively skewed; log transformations were imposed on both variables. Last cannabis use was positively skewed; a square root transformation was implemented. The transformed variables met normality assumptions and were used in subsequent analyses.

4.3.3. Preliminary analyses

Means, standard deviations, and bivariate correlations among the dependent variables collapsed across paradigms are presented in Table 4. A collapsed score was created by taking the mean across the two paradigms. For instance "Drug craving" was calculated by taking the mean of the craving change rating during the VR drug condition and the craving change rating during the Video drug condition. Change scores were used in all analyses and are included in the correlations presented. However, to allow for easier interpretation and consistency with past research, reported dependent variable means and standard deviations are included as raw scores. For all of the dependent variables but the single-item craving rating, scores for the neutral condition and drug condition were significantly correlated. While the craving and MCQ were correlated with each other, the physiological measures were not correlated with craving, MCQ, or each other.

Table 4: Intercorrelations among the dependent variables

	1	2	3	4	5	6	7	8	9	10
1. Drug Craving	1.00									
2. Neutral Craving	.12	1.00								
3. Drug MCQ	.43*	.41*	1.00							
4. Neutral MCQ	.35	.41*	.96**	1.00						
5. Drug HRT Peak	.08	-.02	.11	.22	1.00					
6. Neutral HRT Peak	.00	-.06	.18	.32	.57*	1.00				
7. Drug SC Peak	.01	-.04	.08	-.05	-.17	.00	1.00			
8. Neutral SC Peak	-.06	-.06	.04	-.07	-.12	.09	.96**	1.00		
9. Drug ST Peak	-.10	.23	-.08	-.11	.32	.11	.30	.37	1.00	
10. Neutral ST Peak	.04	.23	-.01	-.07	.27	.01	.18	.20	.92**	1.00
Mean	4.22	1.64	10.36	9.86	97.51	97.71	11.12	10.21	91.81	91.74
Standard Deviation	3.01	2.28	4.99	4.67	9.14	7.09	5.26	5.01	2.05	1.93

Note. N= 26. *p < .05. **p < .001. MCQ= Marijuana Craving Questionnaire, HRT= Heart Rate, SC= Skin Conductance, ST= Skin Temperature.

Means, standard deviations, and only the significant bivariate correlations among the dependent variables and potential covariates are presented in Tables 5-7. Of note, in Table 7, only the physiological variables that were significantly correlated with potential covariates are included in the table. Twelve potential covariates were considered including: age, gender, SES, race, externalizing problems, internalizing problems, game frequency, SAM baseline arousal, SAM baseline pleasure, BSSS total, age of first use, and recency of use. Considering potential covariates, only age, SES, and externalizing problems were significantly correlated with the craving dependent variables. Only age, gender, externalizing problems, baseline SAM pleasure, and game frequency were significantly correlated with the physiological dependent variables. Of note, when using a Bonferroni-corrected *p*-value of < .001 (.05/120 comparisons), none of the correlations remained significant.

Table 5: Significant correlations among craving variables and potential covariates

	1	2	3	4	5	6	7	8	9	10	11
1. VR D-Crav	1.00										
2. VR N-Crav	.06	1.00									
3. Video D-Crav	.54**	.19	1.00								
4. Video N-Crav	.08	.28	-.06	1.00							
5. VR D-MCQ	.17	.49*	.51**	.15	1.00						
6. VR N-MCQ	.09	.42*	.93 ***	.95 ***	.97 ***	1.00					
7. Video D-MCQ	.18	.46*	.58**	.15	.96 ***	.93 ***	1.00				
8. Video N-MCQ	.17	.39	.45*	.24	.93 ***	.95 ***	.94 ***	1.00			
9. Age	.11	.05	.12	-.43*	.08	.05	.03	.13	1.00		
10. SES	-.03	-.46*	-.01	-.15	.08	.05	.09	.04	-.15	1.00	
11. Externalizing Problems	.23	-.12	-.09	.51*	-.19	-.13	-.25	-.09	-.49*	-.01	1.00
Mean	3.78	1.73	4.66	1.54	10.26	9.70	10.46	10.01	16.92	50.13	2.93
Standard Deviation	2.96	2.42	3.35	2.26	4.89	4.66	5.15	4.79	1.26	9.99	1.31

Note. N= 26. *p < .05. **p < .001. ***p < .001. VR= Virtual Reality, D=Drug, N=Neutral, MCQ= Marijuana Craving Questionnaire, SES= Socioeconomic Status.

Table 6: Significant correlations among physiological variables and potential demographic covariates

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. VR D-HRT Peak	1													
2. VR N-HRT Peak	.24	1												
3. Video D-HRT Peak	.45*	.11	1											
4. Video N-HRT Peak	.45*	.31	.76**	1										
5. VR D-SC Peak	-.17	.14	-.26	-.10	1									
6. VR N-SC Peak	-.09	.18	-.18	-.12	.87***	1								
7. Video D-SC Peak	-.06	.03	-.17	.03	.90***	.74***	1							
8. Video N-SC Peak	.05	.17	-.16	.21	.74***	.51*	.91***	1						
9. VR D-ST Peak	.29	.09	.26	.16	.08	-.13	.20	.32	1					
10. VR N-ST Peak	.23	-.09	.34	.22	-.09	-.17	.03	.07	.89***	1				
11. Video D-ST Peak	.14	-.03	.19	.23	.24	.13	.45*	.59**	.76***	.63***	1			
12. Video N-ST Peak	.14	.04	.12	.06	.23	.07	.34	.46*	.87***	.70**	.83***	1		
13. Age	.10	-.21	.35	.22	.35	.29	.47*	.40	-.14	-.08	.05	-.03	1	
14. Gender	.28	-.03	.23	.45*	-.02	-.12	-.01	.11	-.11	-.13	.02	.05	.46*	1
Mean	97.34	98.92	97.40	96.72	11.08	10.51	11.05	10.23	91.73	91.60	91.80	91.93	16.92	
Standard Deviation	10.90	8.76	9.33	7.71	4.94	4.25	5.76	5.86	2.25	2.14	2.01	1.91	1.26	

Note. N= 26. *p < .05. **p < .001. ***p<.001. VR= Virtual Reality, D= Drug, N= Neutral, HRT= Heart Rate, SC= Skin Conductance, ST= Skin Temperature.

Table 7: Significant correlations among select physiological variables and additional potential covariates

	1	2	3	4	5	6	7
1. Virtual Reality Drug HRT Peak	1.00						
2. Virtual Reality Neutral HRT Peak	.24	1.00					
3. Video Neutral HRT Peak	.45*	.31	1.00				
4. Video Drug ST Peak	.14	-.03	.23	1.00			
5. Externalizing Problems	-.23	.47*	.11	-.08	1.00		
6. Game Frequency	-.23	-.30	.08	.41*	-.28	1.00	
7. SAM Baseline Pleasure	.58*	-.17	.43*	.39	-.06	.14	1.00
Mean	97.34	98.92	96.72	91.80	64.96		6.08
Standard Deviation	10.90	8.76	7.71	2.01	11.42		1.47

Note. N= 26. *p < .05. HRT= Heart Rate, ST= Skin Temperature, SAM= Self-Assessment Manikin.

Typical scoring of the MCQ yields four factor scores with no total score; due to having a small sample size and several variables of interest to include in analyses, for the purpose of data reduction, an exploratory factor analysis (EFA) using principal component analysis and promax rotation was conducted on the four baseline MCQ factor scores (Compulsivity, Emotionality, Expectancy, and Purposefulness) using SPSS version 21. The Kaiser-Meyer-Olkin measure of sampling adequacy was .74, indicating that the present data were suitable for principal components analysis. Similarly, Bartlett's test of sphericity was significant ($p < .001$), indicating sufficient correlation between the variables to proceed with the analysis.

Although scoring of the MCQ does not include a total score, the extraction yielded a one-factor solution. This factor accounted for 58.98% of the total variance. Communalities were high for each of the 4 scores, with a range of .53 to .82. This MCQ

composite score accounted for 65.53% of the variance and included all four of the MCQ scores. Corrected item-total correlations ranged from .55 to .80, and Cronbach's coefficient alpha was .82, indicating good subscale reliability. The composite MCQ factor score derived from the EFA was used in subsequent analyses as the representative baseline score. Similarly, a composite MCQ score was calculated for each of the four cue conditions (Neutral VR, Drug VR, Neutral Video, and Drug Video) by taking the mean of the MCQ factor scores (Compulsivity, Emotionality, Expectancy, and Purposefulness) for each condition. For instance, the neutral VR MCQ composite was created by taking the mean of the neutral VR Compulsivity, Emotionality, Expectancy, and Purposefulness scores.

Change-scores were calculated for all outcome variables to measure acute cue-induced craving and arousal, to eliminate baseline variability between participants, and to account for carry-over effects between cue conditions. MCQ change scores were calculated by subtracting the baseline MCQ composite score from the MCQ composite score collected after the respective cue condition. This yielded four MCQ change scores as outcome variables for each participant: Neutral VR, Drug VR, Neutral Video, and Drug Video. Change scores for single-item craving ratings in response to both neutral and cannabis cues were calculated by subtracting the rating made immediately before each of the four cue conditions from the averaged single-item ratings collected during each cue condition. Hence, consistent with the MCQ, there were four craving change

scores for each participant. Similarly, physiological change scores for heart rate, skin conductance, and skin temperature were calculated by subtracting the maximum value during the five-minute baseline from the maximum value collected during the respective cue condition. These change scores were used as the outcome variables in the two primary analyses. All analyses were conducted using SPSS version 21.

4.3.4. Presence

To assess the degree of presence of the VR environments, scores for participants on the four Presence Questionnaire (PQ) factor scales and overall total scores were calculated. The overall total mean score on the PQ was 121.42 (*SD*= 18.61). The mean score on the Involvement scale was 44.36 (*SD*= 10.52), on the Sensory Fidelity scale was 26.04 (*SD*= 5.40), on the Adaptation/Immersion scale was 34.92 (*SD*=6.40), and on the Interface Quality scale was 14.80 (*SD*= 3.76). A prior version of the PQ reported established norms among 152 college students for the Involvement/Control (*M*= 57.39, *SD*= 8.96) and Interface Quality (*M*= 14.65, *SD*= 3.40) factors scales (Witmer & Singer, 1998). Similar to results from Part 2, while the Interface Quality average score for the virtual environment in the current study was similar to the established norms, the Involvement scale for the current study was more than one standard deviation lower than previously established norms. As aforementioned, this finding could be related to the difference in PQ versions, population age differences, or could reflect the difference

in control afforded to participants in past studies compared to the current study that used a fixed VR path.

4.3.5. Differences in craving

Across both paradigms, all participants showed an increase in self-reported single-item craving in response to drug cues as compared to neutral cues. Across both paradigms, all participants showed an increase in MCQ scores in response to drug cues as compared to baseline MCQ scores. However, there existed greater variability when comparing MCQ scores between drug and neutral conditions. For VR, approximately 15% showed no change in their MCQ score between drug and neutral conditions; for Video, this was approximately 8%. For VR, approximately 69% showed an increase in MCQ scores in response to drug cues as compared to neutral cues. For Video, this rate was approximately 81%.

Because no potential confounding variables were found to be significantly correlated with the craving dependent variables using a Bonferroni-corrected cut-off *p*-value of .001, these possible covariates were excluded from further analyses and a MANOVA was conducted rather than a MANCOVA. Excluding covariates is consistent with past research that typically reported conducting analyses with no covariates (e.g., Bordnick et al., 2008; Gray et al., 2008; Niaura et al., 1998). Further, by utilizing a within-subjects design, each participant functioned as his/her own control.

To test the hypotheses that drug cues would produce more craving than neutral cues and to explore whether this varies by paradigm, a 2 (cue) x 2 (paradigm) within-subjects, mixed model MANOVA using the SPSS Mixed procedure was conducted with single-item craving and MCQ scores as the dependent variables (N= 26). There was no missing data.

The results of the mixed model multivariate analysis examining differences in craving yielded a significant multivariate effect of cue, $F(2, 20.40) = 24.28$, $p < .001$. The main effect of paradigm and the interaction of cue and paradigm were non-significant at the multivariate level ($p > .05$). Univariate analysis indicated that as predicted, drug cues produced higher single-item craving change scores (M=2.76, $p < .001$) and higher MCQ change scores (M=9.84, $p < .05$), than neutral cues. These effects and results of cue contrasts can be found in Table 8. [1]

Table 8: Summary of analyses following significant multivariate effect of cue in the prediction of MCQ change scores and single-item craving change scores

Outcome	*F-statistic* (*df*: 2, 22.88)	*Partial η2*	Cue (*M, SE*)	Contrast (*M difference, SE*)	*p-value* (contrast)
MCQ	4.31	.40	Drug (*M*= 9.84, *SE*= .92)	Drug – Neutral (*M*= .53, *SE*= .25)	.049
			Neutral (*M*= 9.31, *SE*= .86)		
Craving	52.03	.82	Drug (*M*= 2.76, *SE*= .34)	Drug – Neutral (*M*= 2.61, *SE*= .36)	.000
			Neutral (*M*= .15, *SE*= .16)		

[1] The same analysis was conducted as a MANCOVA including the three variables that were significantly correlated with dependent variables at a .05 *p*-value (age, SES, and externalizing problems); the results were the same.

4.3.5.1. Exploratory analysis of craving by scene

Each of the two paradigms included three scenes: 1) an outdoor, playground scene; 2) a paraphernalia-only scene; and 3) a party or group cannabis-using scene. To investigate if there were any differences in cue-induced craving by scene and to explore whether this varies by presentation, a 2 (paradigm) x 3 (scene) within-subjects, mixed model MANOVA using the SPSS Mixed procedure was conducted with single-item craving change scores as the dependent variable (N= 26).

The results of the mixed model multivariate analysis examining differences in craving yielded a significant multivariate effect of paradigm, $F(1, 25) = 5.29$, $p < .05$ as well as a significant multivariate effect of scene, $F(2, 25) = 32.93$, $p < .001$. The interaction of cue and presentation was significant at the multivariate level, $F(2, 25) = 4.94$, $p < .05$. Univariate analyses revealed a significant paradigm X scene effect for the paraphernalia-only scene (F (1, 25) = 4.98, $p < .05$) and party/group scene (F (1, 25) = 7.36, $p < .05$). Pairwise comparisons indicated that the mean change in craving differences between the two paradigms were significantly different for the aforementioned two scenes. Specifically, the Video paradigm elicited more craving (M= 4.54, SD= .69) than the VR paradigm (M=3.46, SD= .60) in the paraphernalia-only scene. Similarly, the Video paradigm elicited more craving (M= 5.47, SD= .67) than the VR paradigm (M=4.35, SD= .59) in the party/group scene. Results suggest that the scenes in the Video paradigm

elicited higher mean change in craving than did their counterparts in the Virtual Reality paradigm.

For consistency with past research (e.g., Bordnick, et al., 2009), the univariate scene effect was examined. Univariate analysis indicated that for VR, the outdoor/playground scene (*M*= 3.07, *SD*= .56) evoked lower mean changes in craving than the paraphernalia-only scene (*M*=3.46, *SD*= .60) or the party/group scene (*M*=4.35, *SD*= .59) and the paraphernalia-only scene elicited less mean change in craving than the party/group scene did. The same pattern of results was found for the Video paradigm whereby the outdoor/playground scene (*M*= 3.44, *SD*= .58) evoked lower mean changes in craving than the paraphernalia-only scene (*M*=4.45, *SD*= .69) or the party/group scene (*M*=5.47, *SD*= .67) and the paraphernalia-only scene elicited less mean change in craving than the party/group scene did.

4.3.6. Differences in arousal

Physiological data (heart rate, skin conductance, and skin temperature) were filtered before being analyzed to ensure the data did not contain artifacts. For each participant, interference in the heart rate function was eliminated by smoothing any spikes that lasted less than five seconds (by connecting endpoints and eliminating the spike). Interference lasting longer than five seconds would have to be omitted and a weighted average would have to be taken but no interference lasting longer than five seconds was found. As a data cleaning technique, post-acquisition transformations were

applied to skin conductance and skin temperature using finite impulse response (FIR) digital filtering using a low pass filter. This digital filtering was used to remove data due to noise. As was done with heart rate, interferences were eliminated for skin conductance and skin temperature by connecting endpoints and smoothing spikes; this was applied when a known artifact was present (e.g., participant sneezed or suddenly moved creating the spike, which was noted during data acquisition). Due to equipment error, one subject had no analyzable physiological data (participant 116). One additional participant did not have complete heart rate data. Two additional participants had incomplete skin conductance data and one additional participant had incomplete skin temperature data.

Because no potential confounding variables were found to be significantly correlated with the physiological dependent variables using a Bonferroni-correction, they were excluded from further analyses and a MANOVA was conducted rather than a MANCOVA.

To test the hypotheses that drug cues would induce greater arousal than neutral cues and to explore whether this varies by paradigm, a 2 (cue) x 2 (paradigm) within-subjects, mixed model MANOVA using the SPSS Mixed procedure was conducted with heart rate (beats per minute), skin conductance (GSR units of microSiemens; μS) and skin temperature as the dependent variables (N=25). Missing data was handled using pairwise deletion.

The results of the mixed model multivariate analysis examining differences in arousal yielded a significant multivariate effect of cue, $F(3, 23.43) = 6.41$, $p < .01$. The main effect of paradigm and the interaction of cue and paradigm were non-significant at the multivariate level ($p > .05$). Univariate analysis indicated that as predicted, drug cues produced higher change in skin conductance (M=3.99) than neutral cues (M=2.91), $p < .001$. These effects and results of cue contrasts can be found in Table 9. [2]

Table 9: Summary of analyses following significant multivariate effect of cue in the prediction of physiological change scores

Outcome	*F-statistic* (*df*: 3, 23.43)	*Partial η2*	Cue (*M, SE*)	Contrast (*M difference, SE*)	*p-value* (contrast)
Heart Rate	.00	.00	Drug (*M*= .71, *SE*= .1.36)	Drug – Neutral (*M*= -.02, *SE*= 1.21)	.988
			Neutral (*M*= .72, *SE*= 1.19)		
Skin Conductance	18.03	.66	Drug (*M*= 3.86, *SE*= .66)	Drug – Neutral (*M*= .95, *SE*= .22)	.000
			Neutral (*M*= 2.91, *SE*= .51)		
Skin Temperature	.37	.12	Drug (*M*= 1.66, *SE*= .33)	Drug – Neutral (*M*= .08, *SE*= .13)	.551
			Neutral (*M*= 1.58, *SE*= .28)		

4.3.7. Differences between the paradigms

Ten paired samples t-tests were conducted to compare VR and Video paradigms on the 10 questions regarding effectiveness, enjoyment, interest, acceptability, willingness to use in treatment, benefit of use in treatment, realism, likelihood of side effects, positive reaction, and desire to continue using (see Part 3 Additional Questions

[2] The same analysis was conducted as a MANCOVA including the five variables that were significantly correlated with dependent variables at a .05 *p*-value (age, gender, externalizing problems, baseline SAM pleasure, and game frequency); the results were the same.

in Appendix B). A Bonferroni-corrected *p* level of 0.005 (0.05/10) was used to protect against inflation of Type I errors produced by multiple comparisons. Participants were asked to select "Strongly Disagree" (0), "Disagree" (1), "Agree" (2), or "Strongly Agree" (3) in response to each question. The created scale including these 10 questions had good internal consistency reliability with a Cronbach's alpha of .83.

Across the 10 t-tests, only one was significant at the *p* level of .005. The Video (*M*= 2.27, *SD*=0.60) paradigm was rated as significantly more realistic than the VR (*M*= 1.50, *SD*= 0.81) paradigm, t(25)= 3.95, *p*< .005. As shown in Figure 2, the average response to each question was at least 1.5.

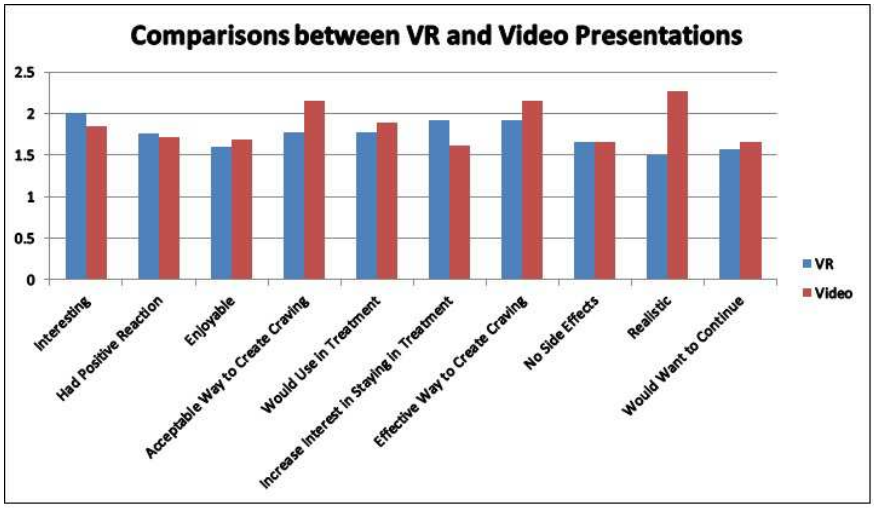

Figure 2: Average responses to questions regarding each paradigm.

As shown in Figures 3-12, most participants rated both paradigms as interesting, enjoyable, and as an acceptable and effective way to create cravings that would not have bad side effects. Most participants also endorsed having a positive reaction to both

paradigms, being willing to use the paradigms in treatment, and felt use of the paradigms would increase their interest in treatment and/or likelihood of staying in treatment. Participants who expressed disagreement with statements that the paradigms would not have bad side effects and explained their answer stated that they felt the paradigms could induce too much craving and a fear of using drugs. Similar responses were given by youth who expressed that they did not have a positive reaction to the paradigms due to experiencing craving. Of note, most participants reported that they felt VR was not realistic.

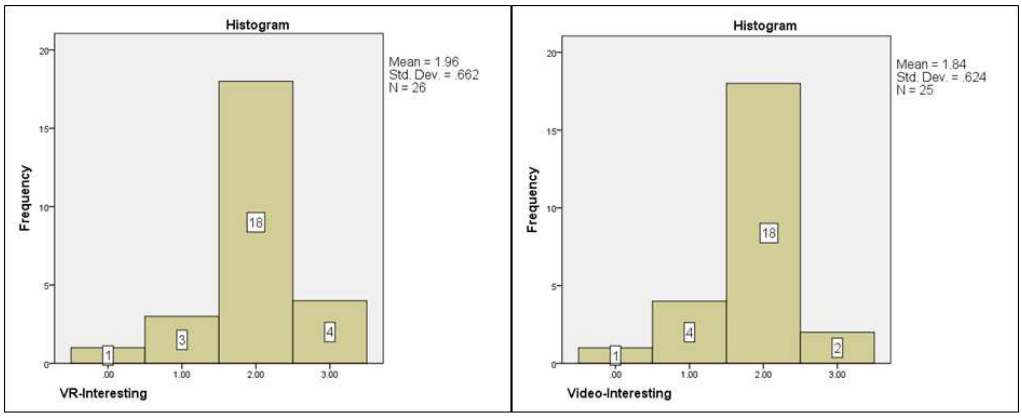

Figure 3: Participant responses to statements regarding how interesting the presentations were.

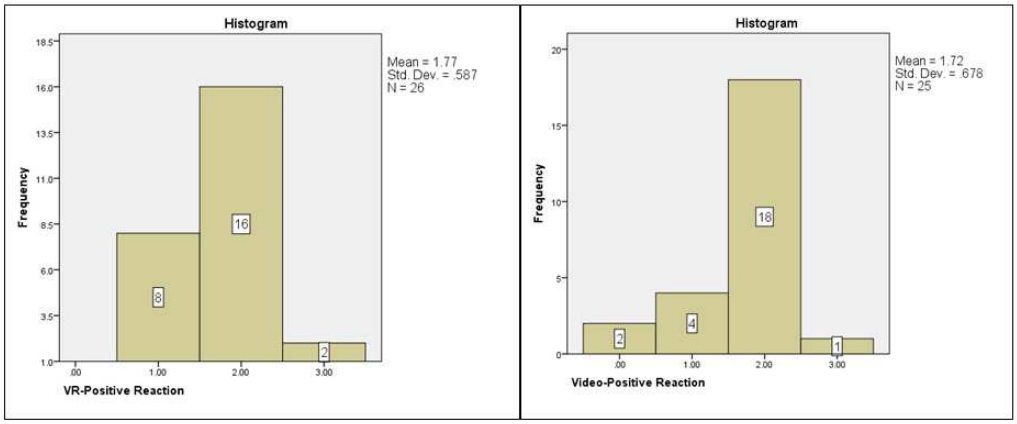

Figure 4: Participant responses to statements regarding whether they had a positive reaction to the presentations.

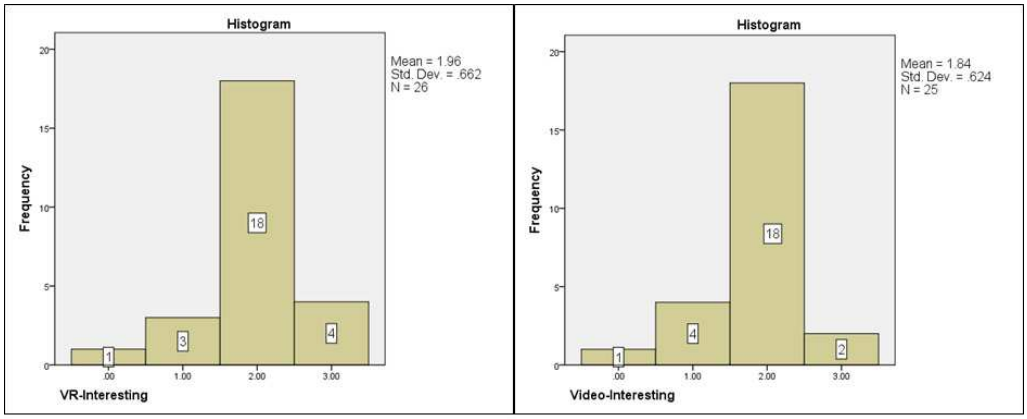

Figure 5: Participant responses to statements regarding how enjoyable the presentations were.

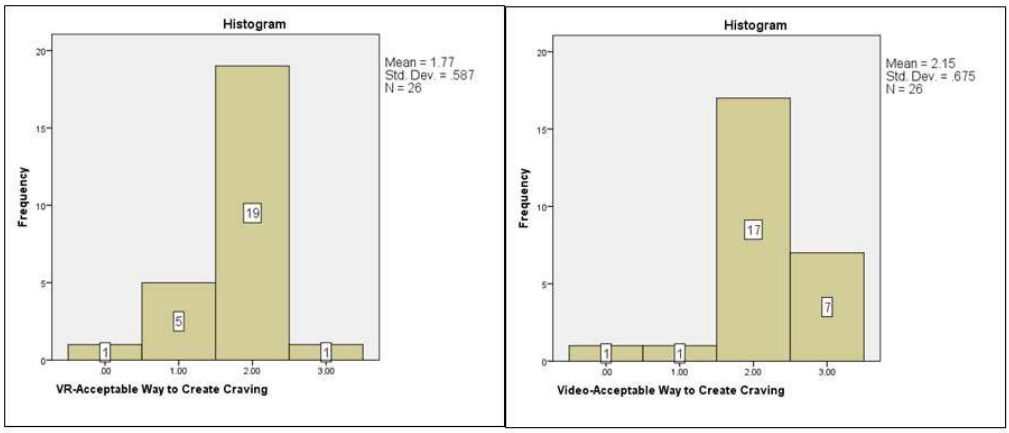

Figure 6: Participant responses to statements regarding acceptability of the presentations for creating craving/urges to use.

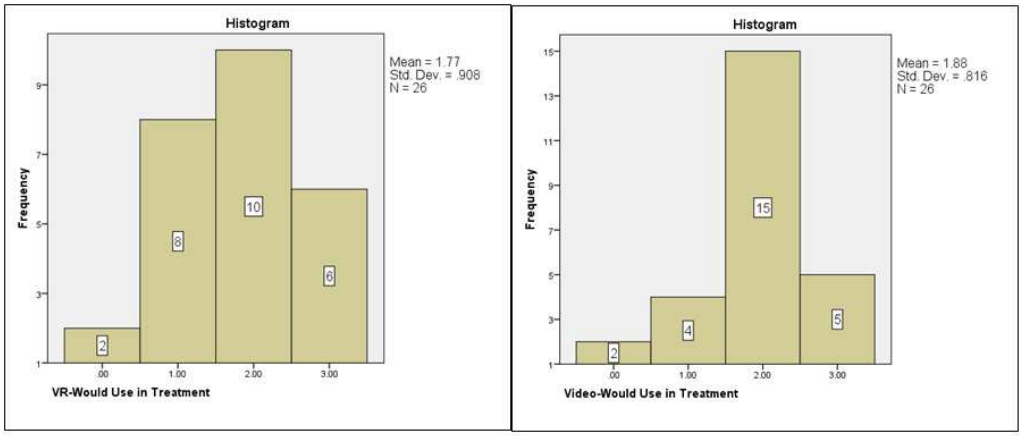

Figure 7: Participant responses to statements regarding willingness to use the presentations as part of treatment.

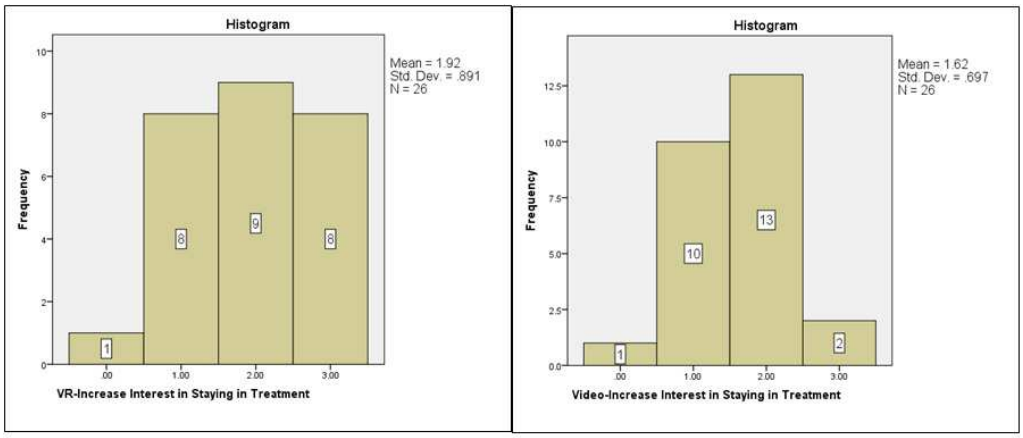

Figure 8: Participant responses to statements regarding whether the presentations would increase interest in treatment and/or likelihood of staying in treatment.

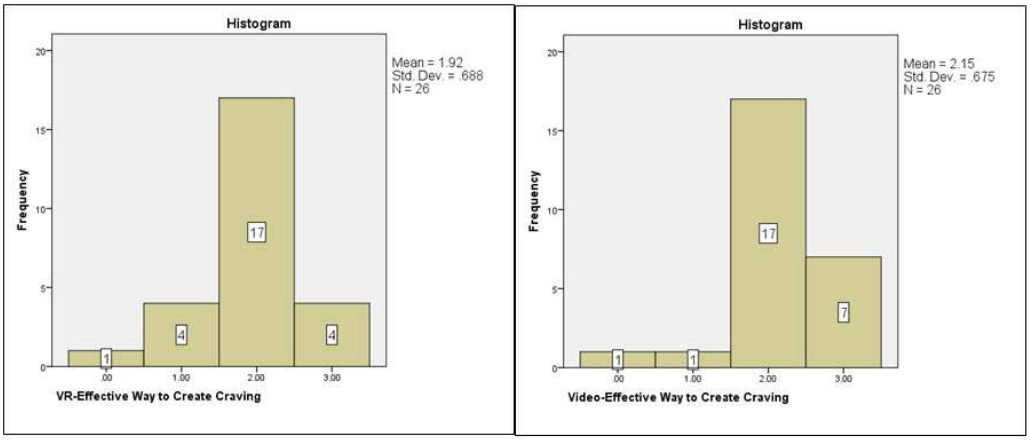

Figure 9: Participant responses to statements regarding the effectiveness of the presentations in creating cravings/urges to use.

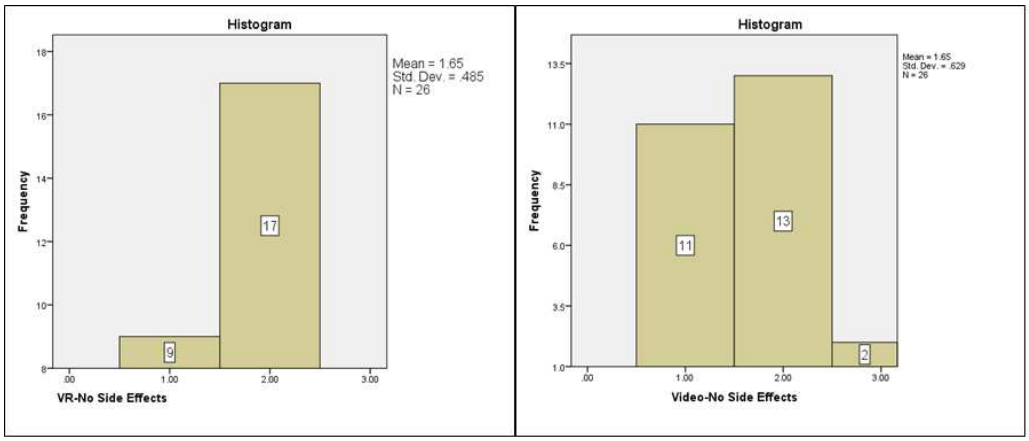

Figure 10: Participant responses to statements regarding the presentations *not* having bad side effects.

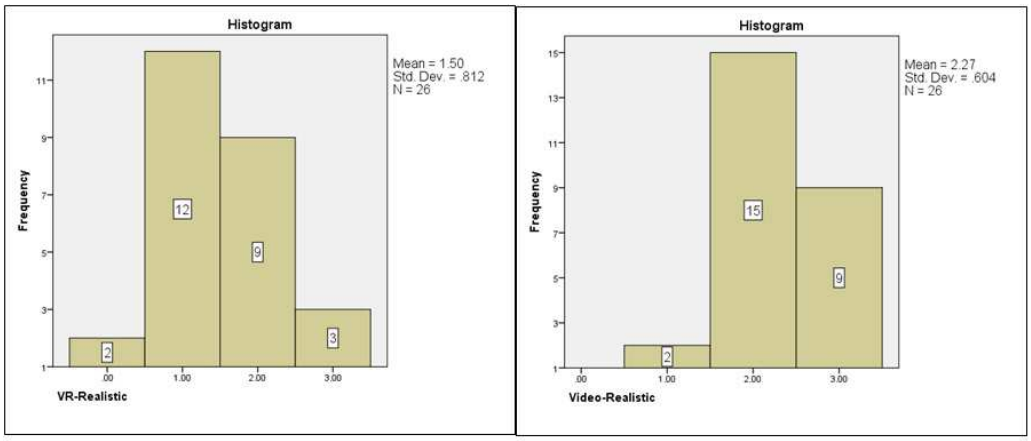

Figure 11: Participant responses to statements regarding how realistic the presentations were.

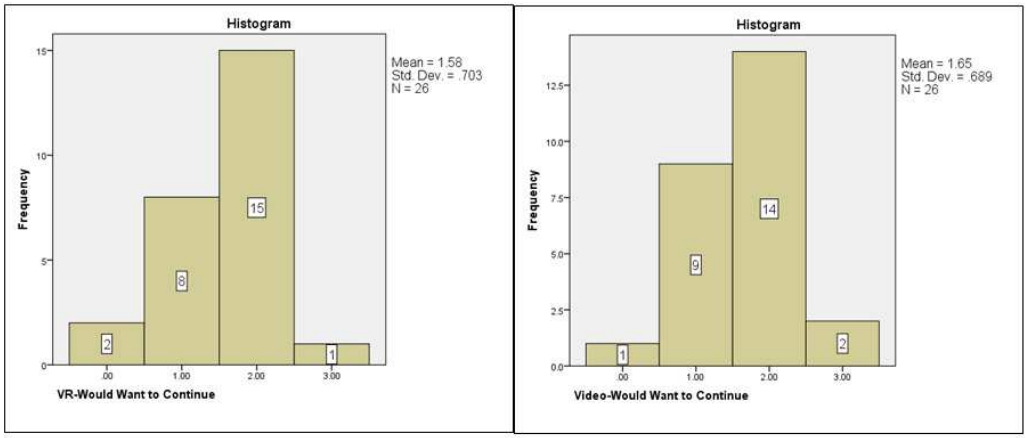

Figure 12: Participant responses to statements regarding whether the participant would want to continue interacting with the presentations.

As part of the additional questionnaire, youth were asked if they prefer the VR paradigm or Video paradigm and asked why. As shown in Table 10, all 26 youth responded, with 15 choosing Video and 11 choosing VR. Participants who preferred the Video paradigm related the choice to the Video being more realistic than the VR paradigm. Those who chose VR related the choice to VR being more interesting, that they enjoy video games, and that VR was more exciting, more engaging, provided more control over what occurs, and that they felt more a part of the presentation than with Video.

Table 10: Participant responses to which paradigm they prefer and why

Participant	Preferred Paradigm	Explanation
1	Video	"The videos were more authentic and gave me stronger urges".
2	Video	"More real life-based, actual people instead of a character. Craved more in video clips than VR".
3	VR	"Would engage me for longer periods of time".
4	VR	"It was very realistic and it made me feel I was right there".
5	Video	"The videos were real people smoking, the VR is still false even though it looks and sounds realistic".
6	Video	"Made more sense to me than the VR".
7	Video	"You can't get any more real than a video whereas with VR, it is not quite a person. Computers are not at the point of creating a real life human in that situation".
8	VR	"Because I felt more into it like I was the one walking around rather than watching someone else do it".
9	VR	"More exciting. Easier to relate to because you're choosing what happens".
10	VR	"Because it's like a video game. Awesome."
11	VR	"Video gets boring"!
12	Video	"Seemed more real, too often distracted by the randomness of VR".
13	VR	"When the characters in the VR talked back, I got the sense of what it was like to be there".
14	VR	"Because in the VR I can control what I can do. I would be able to make decisions on my own".
15	Video	"I prefer it just a tad".
16	Video	"Because video is more reality for me".
17	Video	"Because the VR got frustratingly slow since I was just observing it".

18	Video	“More realistic”.
19	VR	“I like video games”.
20	Video	“Video is more real world and relatable. I can see myself in the videos”.
21	Video	“It affected me more”.
22	VR	“It seemed more realistic”.
23	VR	“I like video games”.
24	Video	“Real”.
25	Video	“It wasn't as boring”.
26	Video	“Because it was more realistic”.

5. Discussion

The aim of this study was to develop and evaluate the feasibility of a virtual reality (VR) cue reactivity paradigm for adolescent cannabis abusers and compare it to a video cue reactivity paradigm. Video and VR paradigms were developed and reactivity to drug cues was compared to neutral cues on measures of subjective craving and physiological measures. The first hypothesis was supported indicating that across the paradigms, treatment-seeking adolescent cannabis abusers experienced higher levels of craving and physiological arousal (as measured by skin conductance) for drug cues compared to neutral cues. This is the first known study to show that virtual reality can be used to elicit craving among adolescent treatment-seeking, cannabis abusers.

The large effect sizes observed for cue-induced craving in the current study are consistent with past cue-reactivity research (e.g., Carter & Tiffany, 1999; Gray et al., 2008). Additionally, the finding that the party/group scene elicited significantly more change in cue-reactive craving is consistent with past research with adult alcoholics and VR suggesting that increased social pressure contributes to increased craving (Cho et el., 2008). Interestingly, in the current study the outdoor/playground scene also included groups of people smoking but also elicited the lowest mean change in craving when compared to the paraphernalia-only and party/group scenes. Being that the outdoor/playground scene was the first scene presented and each progressive scene yielded more craving than the prior one, this finding may imply a time or order effect

such that craving increased over time (carryover craving). For this initial feasibility study with a small sample, it was important to limit variation in what participants saw in order to allow for easier interpretation of findings and hence, all participants saw the same order of scenes; future research should counterbalance and randomize the order of scenes to investigate if there is an effect on craving. Another possible interpretation is that the group/party scene showed people or avatars having more fun and engagement with others (e.g., laughing, playing video games, conversing) than the outdoor/playground scene. Hence, it may be that when others are present, the level of craving elicited is dependent on the context of the scene.

Across paradigms, the skin conductance level was larger in response to drug-related cues as compared to neutral cues. This finding is consistent with previous research with cannabis and other substances such as alcohol, nicotine, cocaine, and heroin (Carter & Tiffany, 1999; Gray et al., 2008) suggesting that drug cues can elevate sympathetic activity with youth cannabis-abusers and further substantiates the potency of the drug cues. Results from the current study confirm that adolescents and young adults can experience subjective and physiological reactivity to drug cues.

While drug cues led to an increase in skin conductance as compared to neutral cues, no significant differences in arousal between drug and neutral cues were found for heart rate or skin temperature. These disparate results are consistent with that of several past research studies including Gray and colleagues (2008) who found that skin

conductance but not heart rate differed significantly between drug and neutral conditions for young adult cannabis abusers who were exposed to cannabis-related imagery, video, and in-vivo cues. Of note, in the current study, changes in heart rate, skin conductance, and skin temperature correlated poorly both with cue-induced craving and with each other; this is consistent with several other cue-reactivity studies (e.g., Tolliver et al., 2010). Across cocaine, alcohol, opioid, and nicotine dependence, according to Tolliver and colleagues (2010), "For each drug, effect sizes of cue exposure on subjective craving are significantly larger than the effect size of physiological responses, and multiple physiological measures are poorly predictive of craving in individuals" (p. 110). This finding that craving has a larger effect size than physiological reactivity has been evidenced in a meta-review across 41 studies investigating reactivity to cocaine, alcohol, heroin, and nicotine (Carter & Tiffany, 1999).

These results could reflect that participants are responding subjectively in the way they think the experimenter wants them to respond leading to higher reported craving than physiologic reactivity. However, due to participants being in treatment, they are more likely to suppress rather than exaggerate report of craving. Several participants in the current study expressed initial guilt at experiencing craving and required psychoeducation to explain that cravings are expected and are not due to a person's lack of commitment to treatment. Consistent with research examining reactivity to cocaine, methamphetamine, and cannabis (Gray et al., 2008; Robbins, Ehrman,

Childress, O'Brien, 1997; Tolliver et al., 2010), the current findings do not support the idea that subjective and physiological responses to cannabis-related cues represent a "unitary drug state" (Robbins et al., 1997).

The second and third hypotheses included that 1) the VR paradigm would elicit more arousal than the video paradigm and 2) participants would find the VR paradigm more enjoyable and acceptable than the video cue-exposure paradigm. Contrary to hypotheses, there were no significant differences between the paradigms either in elicitation of craving or arousal, or in ratings of tolerability or likeability. Of note, the only significant difference between the paradigms was that participants rated the video paradigm as significantly more realistic than the virtual reality paradigm; most participants reported that they felt VR was not realistic. This belief that VR was not realistic may have also been why VR did not evoke more craving or arousal than the video paradigm. Despite using advanced technology similar to what is used in popular video games, the VR software used in this study was created several years ago and may have felt outdated to the participants whereas the video condition utilized videos that were created this year and hence, may better reflect the youths' experiences.

Further, the VR environment had been originally developed for a sample of predominantly urban, low-income substance abusers and hence, included stimuli that the participants in this study may not have related to or felt was realistic for their own lives. Additionally, the VR software was not developed specifically for cannabis abusers,

and thus, included a wide range of drug-related stimuli rather than having an abundance of only cannabis cues whereas the video paradigm included exclusively cannabis cues, some of which could not be included in the VR paradigm. For instance, the video paradigm included clips of people preparing to use cannabis (e.g., weighing cannabis, grinding it, rolling papers, and pouring it into bowls or bongs) whereas the VR paradigm did not include any preparation of cannabis. Several adolescent participants and their substance abuse counselors mentioned that the preparation is a large craving trigger. The VR paradigm included static, unmoving drug paraphernalia and two scenes of people smoking cannabis whereas most of the video condition showed people actively engaging with the cannabis. While efforts were made to create paradigms that were similar, there remained some differences between the paradigms that may have masked the potential benefits of VR compared to video.

As already noted, a limitation to the current study is that the VR software did not include enough cannabis-specific stimuli and hence, other drug paraphernalia (e.g., cigarettes and alcohol) had to be included. Further, the VR software did not include any stimuli representing preparation for use (e.g., grinding the cannabis, packing a bowl). Future research should develop and test a VR paradigm with more cannabis-specific stimuli. Having more cannabis-related stimuli in a VR paradigm would also allow for tailoring the VR presentation to the specific participant viewing it, which would likely increase craving. Tailoring could include allowing the participant to select stimuli

reflecting various components of use (e.g., preparation, use, cleaning paraphernalia, aftermath of use), the mechanism of use (e.g., bowl versus bong versus joint), music heard, clothing, and what avatars say. Future research should evaluate whether the ability to tailor the VR presentation leads to greater craving elicitation than other cue-exposure paradigms such as video clips and whether variations in stimuli (e.g., number of cues, types of cues) influence craving ratings.

Similarly, in order to control for variability in what participants may see, the current study utilized a fixed path that did not allow participants the opportunity to guide themselves through the VR world; although this was done for reasons of improved experiemental control, this may have reduced immersion in the VR world and likely contributed to youth reporting that the VR was not realistic. Future research should consider whether using a fixed versus free path (when the participant chooses where to go and what to interact with in the virtual world) influences craving. It will also be important to examine if participant ability to control the path in the virtual world leads to greater craving elicitation than other cue-exposure paradigms.

A further limitation of the current study is the small sample size, which may have reduced the power needed to obtain significant results. Additionally, the sample included predominantly middle-upper class, White/Caucasian youth. It will be important for these hypotheses to be tested with a larger, more diverse sample.

Importantly, both paradigms (VR and Video) were able to utilize drug cues to elicit increased craving and arousal. Further, most participants rated both paradigms as interesting, acceptable, and effective at eliciting cravings. Notably, most participants marked that they would be willing to use the presentations in treatment and felt use of the presentations would increase their interest in treatment and/or likelihood of staying in treatment. This is significant when considering the potential long-term treatment use of cue-exposure paradigms. Both of the paradigms were shown to be effective at eliciting craving and the youth are willing to use them in treatment and believe the paradigms would help engage them in treatment. Retention in treatment is important but dropout among youth is high with attrition rates ranging from 20% to 50% among substance abusing adolescents (Winters, 1999). It has been shown that people who complete treatment for a substance use disorder are more likely to remain abstinent, have lower relapse rates and substance use, less unemployment, and lower rates of arrest than those who drop out of treatment (Hser et al., 2001; Stark, 1992). Hence, it is noteworthy that the youth in the current study believed that use of these cue-exposure paradigms would help engage them in treatment and/or stay in treatment.

Results from the current study evidencing that both virtual reality and video cue-exposure paradigms are effective at eliciting craving suggests that both can be used in clinical practice. It may be that one paradigm is more effective than the other for specific subpopulations of youth. Alternately, future research may find that the potential added

benefits of virtual reality (e.g., tailoring, participant-led interaction with what is shown through use of a free path) lead it to be a more powerful cue-exposure paradigm than others such as video clips.

Results evidencing that both paradigms elicited craving yields another important implication. Participants in the study were often surprised that they were able to experience high levels of craving that also quickly dissipated; all participants returned to baseline levels of craving prior to leaving the facility and no one had to be kept for extra time to enable craving to subside. Even just this one experience using cue-exposure was significant for the youth to learn that they could "ride out" the craving; using this approach repeatedly over time as an adjunct to treatment would enable the youth to practice and master this skill. Anecdotally in clinical work, clients have stated that they feel as though the cravings will never go away and that they cannot live (and remain abstinent) with cravings this strong. Repeated use of cue-exposure could lead to extinction whereby the cues no longer elicit craving, which would teach the youth that if they can remain abstinent, the intensity of cravings can be reduced.

Treatment outcomes for substance abusing youth are often dismal with findings indicating 70% of adolescents who have completed treatment programs relapse (Florsheim et al., 2008). The field needs innovative, interesting, and effective ways to improve treatment; cue-exposure paradigms and specifically virtual reality, offer a promising approach to improving outcomes for substance-abusing youth.

Appendix A

Part 2 Additional Questions

Please respond by marking an "X" in the appropriate box of the 5-point scale, in accordance with the question content and descriptive labels. Please consider the entire scale when making your responses and please be sure to answer all 12 questions and provide comments if applicable.

1. The virtual reality program was interesting:

Strongly Disagree	Disagree	Agree	Strongly Agree

2. Virtual reality is an acceptable way to create craving/urges to use:

Strongly Disagree	Disagree	Agree	Strongly Agree

3. If my substance abuse counselor were to say that virtual reality could be helpful, I would be willing to use this as part of my treatment:

Strongly Disagree	Disagree	Agree	Strongly Agree

4. Using virtual reality in my treatment would make me more interested in treatment and/or more likely to stay in treatment:

Strongly Disagree	Disagree	Agree	Strongly Agree

5. Overall, I had a positive reaction to the virtual reality program:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

6. Virtual reality is an effective way to create craving/urges to use:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

7. Virtual reality would not have bad side effects:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

8. The virtual reality program was realistic:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

9. The virtual reality program was enjoyable to use:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

10. If given the option of more time, I would want to continue interacting with the virtual reality program:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

11. The virtual reality program had high quality sounds:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

12. The virtual reality program had a high quality visual layout:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

If applicable, please provide any suggestions or comments on how to improve the virtual reality program:

Appendix B

Part 3 Additional Questions (after each condition, asked separately for Video and VR)

You just watched a video. Please answer these questions about it:

1. The VIDEO was interesting:

   |_____| |_____| |_____| |_____|

   Strongly Disagree Disagree Agree Strongly Agree

2. Overall, I had a positive reaction to the VIDEO:

   |_____| |_____| |_____| |_____|

   Strongly Disagree Disagree Agree Strongly Agree

3. The VIDEO was enjoyable:

   |_____| |_____| |_____| |_____|

   Strongly Disagree Disagree Agree Strongly Agree

Part 3 Additional Questions

Now that you have completed the experiment and saw two different presentations that both showed images of marijuana and/or people smoking (one being video and the other being virtual reality), please answer the following questions.

1.

a) VIDEO clips are an acceptable way to create craving/urges to use:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

b) VIRTUAL REALITY is an acceptable way to create craving/urges to use:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

2.

a) If my substance abuse counselor were to say that these VIDEO clips could be helpful, I would be willing to use this as part of my treatment:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

b) If my substance abuse counselor were to say that VIRTUAL REALITY could be helpful, I would be willing to use this as part of my treatment:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

3.

a) Using VIDEO clips in my treatment would make me more interested in treatment and/or more likely to stay in treatment:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

b) Using VIRTUAL REALITY in my treatment would make me more interested in treatment and/or more likely to stay in treatment:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

4.

a) VIDEO CLIPS are an effective way to create craving/urges to use:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

b) VIRTUAL REALITY is an effective way to create craving/urges to use:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

5.

a) VIDEO clips would *not* have bad side effects:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

b) VIRTUAL REALITY would *not* have bad side effects:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

If you marked Strongly Disagree or Disagree, please explain what side effects you think may happen:

_____

_____

6.

a) The VIDEO clips were realistic:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

b) The VIRTUAL REALITY program was realistic:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

7.

a) If given the option of more time, I would want to continue with the VIDEO clips:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

b) If given the option of more time, I would want to continue interacting with the VIRTUAL REALITY program:

|_____| |_____| |_____| |_____|

Strongly Disagree Disagree Agree Strongly Agree

- -

How often do you play arcade or video games? (OFTEN should be taken to mean every day or every two days, on average.)

|_____|_____|_____|_____|

NEVER SOMETIMES OFTEN

If you had to choose ***one*** paradigm that you most prefer, which one would you choose (select 1):

|_____| |_____|

Video Virtual reality

Why did you choose the one you did?

_____

_____

Do you have any other comments or feedback?

_____

_____

References

Achenbach, T. M., & Rescorla, L. A. (2001). Manual for the ASEBA school-age forms & profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Achenbach, T. M., & Rescorla, L. A. (2003). Manual for the ASEBA adult forms & profiles. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Bangay, S., & Preston, L. (1998). An investigation into factors influencing immersion in interactive virtual reality environments. B.K. Giuseppe & E.M. Wiederhold (Eds.). Amsterdam, Netherlands: Ios Press.

Baumann, S.B., & Sayette, M.A. (2006). Smoking cues in a virtual world provoke craving in cigarette smokers. Psychol Addict Behav, 20, 484–489.

Bordnick, P.S., Copp, H.L., Traylor, A., Graap, K.M., Carter, B.L., Walton, A., & Ferrer, M. (2009). Reactivity to cannabis cues in virtual reality environments. Journal of Psychoactive Drugs, 41(2), 105-12.

Bordnick, P.S., Graap, K.M., Copp, H., Brooks, J., Ferrer, M., & Logue, B. (2004). Utilizing virtual reality to standardize nicotine craving research: A pilot study. Addictive Behaviors, 29, 1889-1894.

Bordnick, P. S., Traylor, A., Copp, H. L., Graap, K. M., Carter, B., Ferrer, M., et al. (2008). Assessing reactivity to virtual reality alcohol based cues. *Addictive Behaviors*, *33*(6), 743-756.

Bradizza, C. M., Stasiewicz, P. R., & Maisto, S. A. (1994). A Conditioning Reinterpretation of Cognitive Events in Alcohol and Drug Cue Exposure. *Journal of Behavior Therapy and Experimental Psychiatry, 25*(1), 15-22.

Bradley, M.M., & Lang, P.J. (1994). Measuring Emotion: The self-assessment manikin and the semantic differential. *J Behav. Ther. & Exp. Psychiat., 25*, 49-59.

Brown, R.T., & Coupey, S.M. (1993). Illicit drugs of abuse. *Adolesc Med., 4*(2), 321-340.

Carter, B.L., & Tiffany, S.T. (1999). Meta-analysis of cue-reactivity in addiction research. *Addiction, 94*(3), 327-340.

Cho, S., Ku, J., Park, J., Han, K., Lee, H., Choi, Y. K., et al. (2008). Development and verification of an alcohol craving-induction tool using virtual reality: Craving characteristics in social pressure situation. *Cyberpsychology & Behavior, 11*(3), 302-309.

Chung, T., & Martin, C.S. (2011). Prevalence and clinical course of adolescent substance use and substance use disorders. In Y. Kaminer & K.C. Winters (Eds.), *Clinical manual of adolescent substance abuse treatment* (pp. 1-24). Washington, DC: American Psychiatric Publishing, Inc.

Clark, T. (1989). Managing Outliers: Qualitative Issues in the Handling of Extreme Observations in Marketing Research. *Marketing Research, 1*(2), 31-48.

Cobb, S.V.G., Nichols, S., Ramsey, A., & Wilson, J.R. (1999). Virtual reality-induced symptoms and effects (VRISE). *Presence, 8*(2), 169-186.

Copeland, W, Shanahan, L, Costello, E.J.,& Angold, A.(2011). A cumulative prevalence of psychiatric disorders by young adulthood: A prospective cohort analysis from the Great Smoky Mountains Study. *J. Am. Acad. Child Adolesc. Psychiatry*, *50*, 252-261.

Drummond, D. C. (2001). Theories of drug craving, ancient and modern. *Addiction, 96*(1), 33-46.

Drummond, D.C., & Glautier, S. (1994). A controlled trial of cue exposure treatment in alcohol dependence. *Journal of Consulting and Clinical Psychology, 62*(4), 809-817. doi: 10.1037/0022-006X.62.4.809

Eastwood, B., Bradley, B., Mogg, K., Tyler, E., & Field, M. (2010). Investigating the effects of a craving induction procedure on cognitive bias in cannabis users. *Addiction Research and Theory*, 18, 97-109. doi: 10.3109/16066350802699328

Faul, F., Erdfelder, E., Lang, A.G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Florsheim, P., Heavin, S., Tiffany, S., Colvin, P., & Hiraoka, R. (2008). An Experimental Test of a Craving Management Technique for Adolescents in Substance-abuse Treatment. *Journal of Youth and Adolescence, 37*(10), 1205-1215.

George, D. T., Gilman, J., Hersh, J., Thorsell, A., Herion, D., Geyer, C., et al. (2008). Neurokinin 1 receptor antagonism as a possible therapy for alcoholism. *Science, 319* (5869), 1536-1539.

Gilman, J. M., Ramchandani, V. A., Davis, M. B., Bjork, J. M., & Hommer, D. W. (2008). Why we like to drink: A functional magnetic resonance imaging study of the rewarding and anxiolytic effects of alcohol. *Journal Of Neuroscience, 28*(18), 4583-4591.

Gray, K. M., LaRowe, S. D., & Upadhyaya, H. P. (2008). Cue Reactivity in Young Marijuana Smokers: A Preliminary Investigation. *Psychology of Addictive Behaviors, 22*(4), 582-586.

Gregg, L., & Tarrier, N. (2007). Virtual reality in mental health. *Social Psychiatry & Psychiatric Epidemiology, 42*(5), 343-354. doi:10.1007/s00127-007-0173-4.

Hall, W. & Solowij, N. (1998). Adverse effects of cannabis. *The Lancet, 352*, 1611-1616.

Heishman, S.J., Singleton, E.G., & Liguori, A. (2001). Marijuana Craving Questionnaire: development and initial validation of a self-report instrument. *Addiction, 96*(7), 1023-34.

Hollingshead, A.B. (1975). A.B. *Four factor index of social status*. Unpublished manuscript, Yale University, New Haven, CT.

Hoyle, R.H., Stephenson, M.T., Palmgreen, P., Lorch, E.P., & Donohew, R.L. (2002). Reliability and validity of a brief measure of sensation seeking. *Personality and Individual Differences, 32*, 401-414.

Hser, Y. I., Grella, C. E., Hubbard, R. L., Hsieh, S. C., Fletcher, B. W., Brown, B. S., & Anglin, M.A. (2001). An evaluation of drug treatments for adolescents in 4 US cities. *Archives of General Psychiatry, 58(7)*, 689-95.

Immerse. (n.d.). In *Merriam-Webster online*. Retrieved from http://www.merriam-webster.com/dictionary/immersion

Jacobus, J., Bava, S., Cohen-Zion, M., Mahmood, O., & Tapert, S. F. (2009). Functional consequences of marijuana use in adolescents. *Pharmacology, biochemistry, and behavior, 92*(4), 559-65. Elsevier B.V. doi: 10.1016/j.pbb.2009.04.001.

Jellinek, E. M., Isbell, H., Lundquist, G., Tiebout, H. M., Duchene, H., Mardones, J., et al. (1955). The Craving for Alcohol. *Quarterly Journal of Studies on Alcohol, 16*(1), 34-66.

Johnston, L. D., O'Malley, P. M., Miech, R. A., Bachman, J. G., & Schulenberg, J. E. (2014). *Monitoring the Future national results on drug use: 1975-2013: Overview, Key Findings on Adolescent Drug Use*. Ann Arbor: Institute for Social Research, The University of Michigan.

Juliano, L.M., & Brandon, T.H. (1998). Reactivity to instructed smoking availability and environmental cues: Evidence with urge and reaction time. *Experimental & Clinical Psychopharmacology, 6*(1), 45-53.

Kennedy, R.S., Drexler, J.M., Compton, D.E., Stanney, K.M., Lanham, D.S., Harm, D.L. (2003). Configural scoring of simulator sickness, cybersickness and space adaptation syndrome: Similarities and differences. In L.J. Hettinger & M.W. Haas (Eds.), Virtual and adaptive environments: Applications, implications, and human performance issues (pp. 247-278). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kennedy, R.S., Lane, N.E., Berbaum, K.S., & Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology, 3*(3), 203-220.

Kuntze, M. F., Stoermer, R., Mager, R., Roessler, A., Mueller-Spahn, F., & Bullinger, A. H. (2001). Immersive virtual environments in cue exposure. *Cyberpsychology & Behavior, 4*(4), 497-501.

Lee, J., Lim, Y., Graham, S. J., Kim, G., Wiederhold, B. K., Wiederhold, M. D., et al. (2004). Nicotine craving and cue exposure therapy by using virtual environments. *Cyberpsychology & Behavior, 7*(6), 705-713.

Lee, J.H., Ku, J., Kim, K., Kim, B., Kim, I.Y., Yang, B.H., et al. (2003). Experimental application of virtual reality for nicotine craving through cue exposure. *Cyberpsychology & Behavior, 6*(3), 275-280.

Ludwig, A. M., & Wikler, A. (1974). Craving and Relapse to Drink. *Quarterly Journal of Studies on Alcohol-Part A, 35*(1), 108-130.

Marlatt, G. A. (1985). Cognitive factors in the relapse process. In G. A. Marlatt, and Gordon, J.R. (Ed.), *Relapse prevention* (pp. 128-200). New York: Guilford Press.

Mcauliffe, W.E, & Gordon, R. A. (1974). Test of Lindesmiths Theory of Addiction - Frequency of Euphoria among Long-Term Addicts. *American Journal of Sociology, 79*(4), 795-840.

Merikangas, K.R., He, J., Burstein, M., Swanson, S.A., Avenevoli, S., Cui, L., … Swendsen, J. (2010). Lifetime prevalence of mental disorders in US adolescents: Results from the National Comorbidity Study-Adolescent Supplement (NCS-A). *J Am Acad Child Adolesc Psychiatry, 49* (10), 980-989. doi: 10.1016/j.jaac.2010.05.017

Milin, R., Manion, I., Dare, G., & Walker, S. (2008). Prospective assessment of cannabis withdrawal in adolescents with cannabis dependence: A pilot study. *Journal of the American Academy of Child and Adolescent Psychiatry, 47*(2), 174-179.

Monti, P.M., Binkoff, J.A., Abrams, D.B., Zwick, W.R., Nirenberg, T.D., & Liepman, M.R. (1987). Reactivity of alcoholics and nonalcoholics to drinking cues. *Journal of Abnormal Psychology, 96*, 122–126.

Niaura, R., Shadel, W. G., Abrams, D. B., Monti, P. M., Rohsenow, D. J., & Sirota, A. (1998). Individual differences in cue reactivity among smokers trying to quit: Effects of gender and cue type. *Addictive Behaviors, 23*(2), 209-224.

Niaura, R. S., Rohsenow, D. J., Binkoff, J. A., Monti, P. M., Pedraza, M., & Abrams, D. B. (1988). Relevance of Cue Reactivity to Understanding Alcohol and Smoking Relapse. *Journal Of Abnormal Psychology, 97*(2), 133-152.

Paulus M.P. & Tapert, S.F. (2010). Neurocognition and social cognition in adolescent drug users: Vulnerability and consequences. In C.L. Cooper, J. Field, U. Goswami, R. Jenkins, & B.J. Sahakian (Eds.), *Mental capital and wellbeing* (pp. 119-128). Wiley-Blackwell.

Rankin, H., Hodgson, R., & Stockwell, T. (1983). Cue Exposure and Response Prevention with Alcoholics - a Controlled Trial. *Behaviour Research And Therapy, 21(4)*, 435-446.

Robertson, G., Czeminski, M., & van Dantzich , M. (1997). Immersion in desktop virtual reality. *Proceedings of UIST '97*, ACM Press, 11-19.

Robbins, S.J., Ehrman, R.N., Childress, A.R., O'Brien, C.P. (1997). Relationships among physiological and self-report responses produced by cocaine-related cues. *Addict Behav*, *22*, 157–167.

Rohsenow, D. J., & Monti, P. M. (1999). Does urge to drink predict relapse after treatment? *Alcohol Research & Health, 23*(3), 225-232.

Rohsenow, D. J., Monti, P. M., Rubonis, A. V., Sirota, A. D., Niaura, R. S., Colby, S. M., et al. (1994). Cue Reactivity As A Predictor Of Drinking Among Male Alcoholics. *Journal Of Consulting And Clinical Psychology, 62*(3), 620-626.

Sayette, M.A, Martin, C.S., Hull. J.G., Wertz, J.M., & Perrott, M.A. (2003). Effects of nicotine deprivation on craving response covariation in smokers. *Journal of Abnormal Psychology, 112*, 110-118.

Staiger, P. K., & White, J. M. (1991). Cue Reactivity in Alcohol Abusers - Stimulus Specificity and Extinction of the Responses. *Addictive Behaviors, 16*(5), 211-221.

Stark, M.J. (1992). Dropping out of substance abuse treatment: A clinically oriented review. *Clinical Psychology Review, 12*, 93-116.

Stewart, J., Dewit, H., & Eikelboom, R. (1984). Role of Unconditioned and Conditioned Drug Effects in the Self-Administration of Opiates and Stimulants. *Psychological Review, 91*(2), 251-268.

Substance Abuse and Mental Health Services Administration (2013). *Results from the 2012 National Survey on Drug Use and Health: Summary of National Findings*, NSDUH Series H-46, HHS Publication No. (SMA) 13-4795. Rockville, MD.

Tapert, S. F., Cheung, E. H., Brown, G. G., Frank, L. R., Paulus, M. P., Schweinsburg, A. D., et al. (2003). Neural response to alcohol stimuli in adolescents with alcohol use disorder. [Proceedings Paper]. *Archives of General Psychiatry, 60*(7), 727-735.

Tiffany, S. T., & Conklin, C. A. (2000). A cognitive processing model of alcohol craving and compulsive alcohol use. *Addiction, 95*(8), S145-S153.

Tolliver, B. K., McRae-Clark, A. L., Saladin, M., Price, K. L., Simpson, A. N., DeSantis, S. M., & ... Brady, K. T. (2010). Determinants of cue-elicited craving and physiologic reactivity in methamphetamine-dependent subjects in the laboratory. *The American Journal Of Drug And Alcohol Abuse, 36(2)*, 106-113. doi:10.3109/00952991003686402

Vandrey, R., Budney, A. J., Kamon, J. L., & Stanger, C. (2005). Cannabis withdrawal in adolescent treatment seekers. *Drug And Alcohol Dependence, 78*(2), 205-210.

Vora, J., Nair, S., Gramopadhye, A.K., Duchowski, A.T., Melloy, B.J., & Kanki, B. (2002). Using virtual reality technology for aircraft visual inspection training: presence and comparison studies. *Applied Ergonomics, 33*, 559-570.

Weinstein, A., Lingford-Hughes, A., Martinez-Raga, J., & Marshall, J. (1998). What makes alcohol-dependent individuals early in abstinence crave for alcohol: Exposure to the drink, images of drinking, or remembrance of drinks past? *Alcoholism-Clinical And Experimental Research, 22*(6), 1376-1381.

Winters , K.C. (1999). Treating adolescents with substance use disorders: An overview of practice issues and treatment outcome. *Substance Abuse, 20*(4), 203-225.

Wise, R. A. (1988). The Neurobiology of Craving - Implications for the Understanding and Treatment of Addiction. *Journal Of Abnormal Psychology, 97*(2), 118-132.

Wise, R. A., & Bozarth, M. A. (1987). A Psychomotor Stimulant Theory of Addiction. *Psychological Review, 94*(4), 469-492.

Witmer, B.G., Jerome, C.J., & Singer, M.J. (2005). The factor structure of the Presence Questionnaire. *Presence, 14*(3), 298-312.

Witmer, B.G., & Singer, M.J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence, 7*(3), 225-240.

Biography

Jacqueline R. Hersh was born in New York on May 19, 1983. She graduated Summa Cum Laude and received a bachelor of the arts degree in Sociology and with Highest Honors in Psychology in 2005. She also received a master of the arts degree in Clinical Psychology from Duke University in 2011 and will complete her clinical psychology internship at NYU/Bellevue Hospital Center in June 2014. Jacqueline has presented four posters, presented one symposium ("Psychosocial Treatments for Internalizing Adolescents with Substance Use Disorders" in 2010) and has been involved with four additional poster presentations. She has two first-author publications in peer-reviewed journals ("The Influence of Comorbid Depression and Conduct Disorder on MET/CBT Treatment Outcome for Adolescent Substance Use Disorders" and "Neurokinin 1 receptor antagonism as a possible therapy for alcoholism") as well as being an author on six additional peer-reviewed publications. Since obtaining the bachelor's degree, Jacqueline has received a Dissertation Research Award from the American Psychological Association (APA), a Webb Dissertation Research Award from Duke University, a Summer Research Fellowship from Duke University, a Graduate Fellowship from Duke University, the Claire Hamilton Conference Travel Award from Duke University, received Honorable Mention from the American Psychological Association Minority Fellowship, and been a T32 Training Grant Recipient for the National Institute of Mental Health through Duke University.