Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

An Optimal Control Framework for Adaptive Virtual Reality Interventions in Substance Use Rehabilitation

Charlotte Cheng

charlotte.q.cheng@gmail.com

George Mason University

Padmanabhan Seshaiyer

George Mason University

Research Article

Keywords: Mathematical Model, Optimal Control, Substance Use Disorder

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11708526/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

An Optimal Control Framework for Adaptive Virtual Reality Interventions in Substance Use Rehabilitation

Charlotte Cheng[1]
Mentor: Padmanabhan Seshaiyer[2]

[1]White Station High School, Memphis, TN, USA
[2]George Mason University, Fairfax, VA, USA

Abstract

Despite advances made over the years, substance use disorders are still a significant public health challenge, with relapse prevention continuing to be a major problem in rehabilitation. Traditional rehabilitation strategies often face barriers such as limited accessibility and difficulty recreating real-world environments that are associated with cravings. As a result, researchers have begun to study virtual reality (**VR**) as a tool for drug rehabilitation. I developed an Adaptive-Perturbed-Corrective-Relapse (**APCR**) compartmental model governed by differential equations that represent the transitions between stable recovery, cue-induced perturbation, corrective intervention, and relapse. The model incorporates a craving trigger parameterized using alcohol-related gaze data collected in a VR environment and a baseline recovery cue derived from physiological biomarkers. Transition rates were assigned based on literature-informed assumptions rather than clinical outcome data. To determine when and how strongly therapeutic interventions should be delivered, I formulated an optimal control problem in which intervention intensity is represented by a time-dependent control variable. Using Pontryagin's Minimum Principle, I derived an intervention strategy that balances reducing the perturbed state with the control effort over 60 minutes. Using the forward–backward sweep method, I obtained an optimal control strategy that reduced the simulated peak proportion of individuals in the Perturbed state by 33.26% compared with the passive baseline, highlighting the potential of mathematical optimization for adaptive VR rehabilitation. This work supports the United Nations Sustainable Development Goal 3 (Good Health and Wellbeing).

1 Introduction

Substance use disorder (**SUD**) is a chronic and relapsing condition during which triggering cues associated with specific substances can cause craving and increase vulnerability to relapse. Virtual reality (**VR**) provides an immersive, controlled, and repeatable setting to induce addiction related responses and, in response, deliver cue-based interventions, making the timing and intensity important variables for treatment optimization (Glavak-Tkalić et al., 2025; Mazza et al., 2021). Because an individual's movement between recovery-related states occurs over time and can be influenced by external therapeutic efforts, VR intervention can be framed as an optimal control problem for a dynamic system.

Differential equation models that originated from studying epidemiology have become widely used to study transition among states. Optimal control has been commonly applied to these systems. In smoking research, S. Zhang et al. (2024) developed a smoking epidemic model with different smoking degrees and analyzed its deterministic and stochastic dynamics in Mathematical Biosciences. In alcohol-use modeling, Z. Zhang et al. (2020) studied a delayed drinking model with optimal control, while Wireko et al. (2024) later formulated a joint drinking-and-smoking model with optimal control analysis. In drug-use modeling, Labzai et al. (2020) analyzed a discrete-time drug consumption model with optimal control, and Sowndarrajan et al. (2022) formulated and analyzed an optimal control heroin epidemic model. Similar compartmental addiction modeling has also been developed by Baca et al. (2024), who described the patient addiction journey through clinically meaningful transition states. In these situations, Pontryagin's Minimum Principle provides the necessary conditions for optimality through adjoint variables and a Hamiltonian. The resulting optimality system is then commonly solved using the forward–backward sweep method.

However, despite a substantial amount of research on optimal control models that are used to study smoking, drinking, drugs, etc., most focus on dividing populations into groups to study how general interventions such as education and treatment affect the groups (Baca et al., 2024). Even then, none of these studies consider responses induced by VR trigger exposure. As a result, existing optimal control methods have yet to be directly applied to VR-based interventions to determine when and how they should be delivered and designed to be most effective against relapse.

To address this gap, this paper develops a parameterized, optimal control framework for adaptive VR intervention through a compartmental model of recovery dynamics. Specifically, a four-compartment Adaptive–Perturbed–Corrective–Relapse (**APCR**) model is introduced to describe transitions among stable recovery, cue-induced perturbation, corrective intervention, and relapse states under cue-driven forcing. The trigger function is calibrated through the use of empirical gaze data assessing attentional bias toward alcohol-related cues in a VR environment. Since alcohol represents 59.6% of all people with a substance use disorder (Substance Abuse and Mental Health Services Administration, 2024), it provides a clinically representative benchmark of the time-dependent relapse dynamics. Afterwards, the existence, uniqueness, non-negativity, and boundedness of solutions within an invariant region for all admissible inputs are established. I then optimized the recovery cue to minimize perturbation and relapse subject to intervention cost; I further characterized the optimality system via Pontryagin's Minimum Principle and solved it numerically using the forward–backward sweep method.

The remainder of this paper is organized as follows. Section 2 introduces the APCR model and the equations governing the transitions. Section 3 formulates the optimal control problem, proves the existence of an optimal control, and derives the optimality system using Pontryagin's Minimum Princi-

ple. Section 4 describes the numerical method and presents the computational experiments. Section 5 concludes the paper.

2 Methods and Model

2.1 VR System Context

The VR framework acts as an immersive cue-exposure environment that is designed to assist with substance use rehabilitation. In this work, the VR system is represented through a computational framework that evaluates intervention dynamics. The data collected from the VR are used to parametrize the model that simulates state transitions. The rehabilitation process begins once the participant initiates the session. This transports them from the home-screen and places them into a controlled digital setting where the baseline state conditions for the system are established.

Once the participant enters, they experience a two-stage exposure process. In the first stage, the environment features a low-density of craving-inducing target cues. This allows the user to first acclimate to the environment. After this designated duration, the second stage begins where the environment then transitions into the high-density phase where there is a significant increase in the frequency and concentration of craving-inducing triggers.

Participant interaction with these virtual cues is evaluated through gaze tracking metrics, specifically gaze fixation and gaze frequency. Through measuring how the frequency and intensity of user focus on specific triggers, the system determines the strength of a target cue functioning as an active stimulus inducing cravings. These biometric gaze inputs serve as the foundation for the trigger function $T(t)$ (Section 2.3). This bridges real-time user behavior in the VR environment with the dynamic state transitions in the mathematical model.

2.2 Compartments and State Dynamics

I modeled recovery using four mutually exclusive compartments: Adaptive, Perturbed, Corrective, and Relapse. At the population level, $A(t)$, $P(t)$, $C(t)$, and $R(t)$ represent the proportion of individuals in a population that occupy a specific recovery state at time t. Since the individuals are not able to exit the system, the cohort proportions must sum to one and the state variables satisfy:

$$A(t)+P(t)+C(t)+R(t)=1$$

for all t.

- **Adaptive** $A(t)$ state - stable recovery
- **Perturbed** $P(t)$ state - cue induced cravings
- **Corrective** $C(t)$ state - undergoing recovery
- **Relapse** $R(t)$ state - relapse

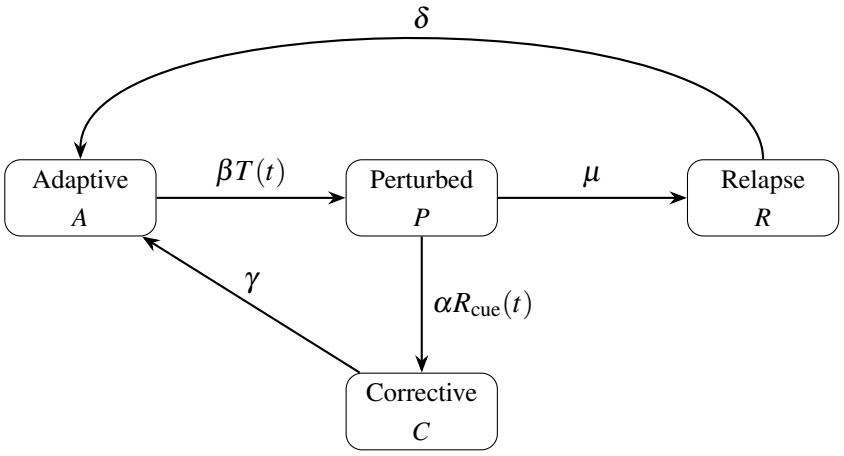
The following system (1) is the corresponding baseline Kolmogorov forward equation for this process. In this system, each equation balances the proportion entering a state against the proportion leaving

it, at the transition rates listed in Table 1. A, P, C, and R can be interpreted as the cohort proportions occupying each state.

The dynamics between the states are governed by four coupled ordinary differential equations:

$$\begin{aligned}
\frac{dA}{dt} &= -\beta\, T(t) A + \gamma C + \delta R \\
\frac{dP}{dt} &= \beta\, T(t) A - \alpha\, R_{\text{cue}}(t) P - \mu\, P, \\
\frac{dC}{dt} &= \alpha\, R_{\text{cue}}(t) P - \gamma C, \\
\frac{dR}{dt} &= \mu\, P - \delta R.
\end{aligned} \tag{1}$$

with nonnegative initial data of $(A(0), P(0), C(0), R(0))$ summing to one. In this system, $T(t)$ represents the craving **trigger** and $R_{\text{cue}}(t)$ is the therapeutic **recovery cue** (Section 2.3); the transition parameters are found and defined in Table 1. A trigger induces the transition from Adaptive to Perturbed at a rate $\beta\, T(t)$. From the Perturbed state, the recovery cue drives the transition from Perturbed to Corrective at rate $\alpha R_{\text{cue}}(t)$, whereas progression towards Relapse occurs at rate μ. The Corrective to Adaptive and Relapse to Adaptive transitions then occur at rates of γ and δ respectively. The compartments and transition rates are visualized in the diagram below with the appropriate values.



A visual representation of the APCR compartments and transitions. All individuals begin in the adaptive state before transitioning to perturbed due to the trigger $T(t)$. From the perturbed state individuals either go to the corrective state due to a recovery cue $R_{\text{cue}}(t)$ that is replaced in section 3.1 for $u_1(t)$. From the latter two states, individuals route back to the adaptive state ensuring no individuals exit the system.

Table 1: Model parameters (per minute) and objective weights.

Symbol	Meaning	Value
β	trigger susceptibility	**0.20**
γ	successful reintegration	**0.10**
α	responsiveness to intervention	**0.15**
μ	progression toward relapse	**0.03**
δ	recovery from relapse	**0.005**
A_1	penalty weight on the Perturbed state	10
B_1	control-effort weight	0.10
$[u_{\min}, u_{\max}]$	control bounds	$[0,1]$
t_f	time horizon	60
(A_0, P_0, C_0, R_0)	initial condition	$(1.00, 0, 0, 0)$

Table 2: Literature-informed justification for the APCR transition parameters. The parameters are model-specific latent transition rates and are not directly estimated from the single-session gaze data. The cited studies support the relative assumptions of each transition.

Symbol	Justification
β	Drug cues can induce trigger cravings quickly within a single cue exposure session. Carter and Tiffany (1999) found increases in self-reported craving following exposure to drug-related cues across 41 cue-reactivity studies. Consistent with this evidence, β was assigned a relatively high rate to represent the potential for a rapid cue-induced perturbation.
γ	The transition from the perturbed to the corrective state represents a positive response to the intervention. However, going back from corrective to adaptive is a more gradual recovery process. Kober et al. (2010) show that craving can be reduced during cognitive regulation. In addition, Conklin and Tiffany (2002) frame cue-exposure treatment as extinction-based learning. This motivates modeling the return to the Adaptive state as a recovery process rather than an instantaneous reset. As a result the model assumes that the transition back to adaptive is slower than the initial transition to corrective.
α	The value of α was chosen to be slightly smaller than β because the intervention acts in response to cue-induced perturbation rather than initiating the perturbation itself. This ordering is consistent conceptually with evidence that cognitive regulation can reduce craving within an experimental session (Kober et al., 2010). However, α cannot be much smaller than β, since an intervention that acts on a much slower time scale would be unable to effectively counteract the trigger.

Symbol	Justification
μ	Cue-induced perturbations can develop on a relatively short timescale. Epstein et al. (2009) observed increases in cue exposure and mood changes during the five hours preceding cocaine use and heroin craving. Accordingly, in this model, $\mu < \beta$ was adopted to reflect the intended distinction between rapid cue-induced perturbation and slower progression toward relapse . Similarly, $\mu < \alpha$ was imposed as a modeling assumption so that corrective dynamics operate on a faster timescale than progression toward relapse.
δ	Recovery from relapse to adaptive was modeled as the slowest transition. Recovery was modeled as a gradual process requiring sustained support, so δ was chosen much smaller than γ and μ.

2.3 External Input Signals

Trigger $T(t)$: The trigger was activated across a single exposure window $[t_{\text{on}}, t_{\text{off}}]$ and remains at zero otherwise. Its behavior throughout this window was determined through a weighted combination of two behavioral benchmarks obtained from a VR session

$$T(t) = \begin{cases} w_1\, g_{\text{fixation}} + w_2\, g_{\text{freq}}, & t_{\text{on}} \leq t \leq t_{\text{off}} \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where $w_1 + w_2 = 1$ for preservation of original scale.

40 individuals' reactions to alcohol related trigger objects in a VR setting (Simon et al., n.d.) were used to validate and parameterize the variables, providing a data-driven empirical value for modeling responses to substance-related cues. The gaze fixation (g_{fixation}) was determined through analyzing the average fixation of an individual spent on a triggering object. The gaze frequency (g_{freq}) was determined by analyzing the frequency of the fixations an individual has on the trigger objects. Both the gaze fixation and the gaze frequency were normalized via min-max normalization as shown in equation 3

$$x_{\text{normalized}} = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \tag{3}$$

$$x_{\text{normalized}} = \frac{x_{\text{max}} - x}{x_{\text{max}} - x_{\text{min}}} \tag{4}$$

Recovery Cue $R_{\text{cue}}(t)$: In the passive baseline, the therapeutic recovery cue was activated across a single exposure window from t_{start} until the end of the period. Its behavior throughout the window was determined through a weighted combination of heart rate, heart rate variability, and stress score (Yfantidou et al., 2022). These physiological measures were obtained from a publicly available Fitbit data set rather than collected directly from the VR intervention. The heart rate variability was normalized via min-max normalization as shown in equation 3. The inverse of the heart rate and stress score were used as higher values dictate worse results. These two indicators were normalized through inverse min-

max normalization as shown in equation 4. They are to be used as a representative of physiological signals for the formulation of the mathematical model illustratively rather than clinical validation.

$$R_{\text{cue}}(t) = \begin{cases} n_1\,\text{hrv} + n_2\,\text{hr} + n_3\,\text{stress}, & t \geq t_{\text{start}} \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

where $n_1 + n_2 + n_3 = 1$ for preservation of original scale and units.

Rather than using visual biometrics, $R_{\text{cue}}(t)$ is derived directly from physiological response metrics: heart rate variability (hrv), heart rate (hr), and stress scores (stress) gathered from wearable Fitbit logs. Within the active time frame, these normalized physiological signals are combined via a weighted linear sum.

The values used for the external input signal can be seen in Table 3

Table 3: External Input Signal Values.

Symbol	Meaning	Value
g_{fixation}	normalized gaze fixation	**0.548**
g_{freq}	normalized gaze frequency	**0.572**
hrv	normalized heart rate variability	**0.508**
hr	normalized inverse heart rate	**0.506**
stress	normalized inverse stress score	**0.316**
(w_1, w_2)	chosen $T(t)$ weighted coefficients	**(0.50,0.50)**
(n_1, n_2, n_3)	chosen $R_{\text{cue}}(t)$ weighted coefficients	**(0.40,0.30,0.30)**

Justification R_{cue} coefficients: HRV was weighted slightly more because it is used as an indicator of autonomic regulation (Cannon, 2025), which is directly relevant to craving regulation. The inverse heart rate and stress inverse score were included as complementary but less specific physiological indicators that could be affected by other measures. Therefore, $(0.40, 0.30, 0.30)$ was used as a near-equal nominal weighting with slight HRV emphasis.

3 Optimal Control Formulation and Mathematical Analysis

3.1 Optimal Control Formulation

The baseline recovery cue $R_{\text{cue}}(t)$ is a prescribed and passive signal. Because a VR system can deliver a therapeutic cue adaptively upon detecting elevated craving, rather than relying on a fixed schedule, I replaced $R_{\text{cue}}(t)$ in the Perturbed to Corrective transition by a bounded, time-dependent **control** $u_1(t)$,

giving the controlled dynamics:

$$\begin{aligned}
\frac{dA}{dt} &= -\beta T(t) A + \gamma C + \delta R, \\
\frac{dP}{dt} &= \beta T(t) A - \alpha u_1(t) P - \mu P, \\
\frac{dC}{dt} &= \alpha u_1(t) P - \gamma C, \\
\frac{dR}{dt} &= \mu P - \delta R,
\end{aligned} \tag{6}$$

where the control input $u_1(t)$ is constrained to the admissible set:

$$\mathscr{U} = \{u_1(t) | u_{\min} \leq u_1(t) \leq u_{\max} \text{ for all } 0 \leq t \leq t_f\}$$

Cost Function: I sought a control that minimizes the time spent in the Perturbed which indirectly reduces the relapse state, while balancing the cost of intervention, through the cost function:

$$J(u_1) = \int_0^{t_f} \left(A_1 P(t) + \tfrac{1}{2} B_1 u_1(t)^2 \right) dt. \tag{7}$$

The control-effort cost $\frac{1}{2} B_1 u_1^2$ penalizes large control actions in order to prevent excessive intensity, high costs, or aggressive intervention. The coefficient B_1 represents the cost weight. The state cost $A_1 P(t)$ penalizes the system for spending time in the perturbed state. The coefficient A_1 represents the penalty weight.

Hamiltonian: Combining the J cost function with the state equations weighted by adjoint (costate) variable $\lambda = \lambda_A, \lambda_P, \lambda_C, \lambda_R$ results in the formulation of the Hamiltonian:

$$\begin{aligned}
H = {} & A_1 P + \tfrac{1}{2} B_1 u_1^2 \\
& + \lambda_A \big(-\beta T A + \gamma C + \delta R \big) + \lambda_P \big(\beta T A - \alpha u_1 P - \mu P \big) \\
& + \lambda_C \big(\alpha u_1 P - \gamma C \big) + \lambda_R \big(\mu P - \delta R \big).
\end{aligned} \tag{8}$$

Adjoint system: By Pontryagin's Minimum Principle, the co-states are governed by $\dot{\lambda}_i = -\partial H / \partial x_i$. The transversality conditions $\lambda_i(t_f) = 0$ ensure at the end of the time period (t_f) no longer affect the final cost or performance of the system:

$$\begin{aligned}
\frac{d\lambda_A}{dt} &= \beta T(t) (\lambda_A - \lambda_P), \\
\frac{d\lambda_P}{dt} &= -A_1 + (\alpha u_1 + \mu) \lambda_P - \alpha u_1 \lambda_C - \mu \lambda_R, \\
\frac{d\lambda_C}{dt} &= \gamma (\lambda_C - \lambda_A), \\
\frac{d\lambda_R}{dt} &= \delta (\lambda_R - \lambda_A)
\end{aligned} \tag{9}$$

Optimal control: Pointwise minimization H through the first-order condition $\partial H / \partial u_1 = 0$ yields the equation $B_1 u_1 - \alpha P (\lambda_P - \lambda_C) = 0$. Accounting for the upper and lower control limits, the optimal

control $u^*(t)$ yields the characterization:

$$u_1^*(t) = \min\left\{u_{\max}, \max\left\{u_{\min}, \frac{\alpha P(t)(\lambda_P(t) - \lambda_C(t))}{B_1}\right\}\right\}. \tag{10}$$

The coupled problem defined by the state system (6), the adjoint system (9), and the control law (10) is solved numerically by the forward–backward sweep method in Section 4.

3.2 Mathematical Analysis

To analyze the state dynamics, I define the feasible region of values as the closed, positively invariant simplex:

$$\Delta = \left\{(A,P,C,R) \in \mathbb{R}^4_{\geq 0} \;\middle|\; A+P+C+R = 1\right\}. \tag{11}$$

Proposition 1. *Given an initial condition $x(0) \in \Delta$ and every admissible control $u_1 \in \mathscr{U}$, System 6 contains a unique solution on $[0,t_f]$.*

Proof. Consider System 6 as

$$\dot{x} = f(t,x,u_1(t)).$$

For each admissible control u_1, the function f is measurable with respect to t because $T(t)$ and $u_1(t)$ are measurable. For each fixed t, the function f is continuous in the state variables $x = (A,P,C,R)$. Furthermore, each component of f is linear in the state variables A,P,C,R, and is therefore locally Lipschitz in x. Since $T(t)$ is bounded and $u_1(t) \in [u_{\min}, u_{\max}]$, the coefficients in the state system are bounded on $[0,t_f]$. Therefore, by the Carathéodory existence and uniqueness theorem (Filippov, 1988), System 6 has a unique solution on $[0,t_f]$ for every $x(0) \in \Delta$. □

Proposition 2. *Define the state vector as*

$$x(t) = \begin{pmatrix} A(t) \\ P(t) \\ C(t) \\ R(t) \end{pmatrix}$$

Given a state $x(0) \in \Delta$, the solution trajectory $x(t)$ in System 6 lies in $x(t) \in \Delta$ for all $t \in [0,t_f]$.

Proof. This means that there is both a conservation of total mass and the non-negativity of each state component variable.

Conservation of Mass: Adding up the four differential equations of System 6 results in:

$$\frac{d}{dt}\big(A(t)+P(t)+C(t)+R(t)\big) = \dot{A}+\dot{P}+\dot{C}+\dot{R} = 0$$

Since each term appears exactly twice among the four differential equations (once negative and once positive), the terms cancel out. This ensures that the total sum is constant:

$$A(t)+P(t)+C(t)+R(t) = A(0)+P(0)+C(0)+R(0) = 1$$

Non-Negativity: In order to demonstrate that the variables stay non-negative over time, Nagumo's condition (Song, 2022) requires verifying the direction of the system when any one variable drops to zero. If the rate of change of a variable stays non-negative ($\dot{x} \geq 0$), then the trajectory cannot be negative:

$$\begin{aligned} A = 0 &\implies \dot{A} = \gamma C + \delta R \geq 0, \\ P = 0 &\implies \dot{P} = \beta T(t) A \geq 0, \\ C = 0 &\implies \dot{C} = \alpha u_1(t) P \geq 0, \\ R = 0 &\implies \dot{R} = \mu P \geq 0. \end{aligned}$$

Since the derivatives of the state are non-negative on the boundaries of $\mathbb{R}^4_{\geq 0}$, trajectories do not cross into negative values.

Together, conservation and non-negativity show that $x(t) \in \Delta$ for all $t \in [0, t_f]$. □

Proposition 3. *Let $\mathscr{U} = \{u_1 : [0,t_f] \to [u_{\min}, u_{\max}]$ measurable$\}$ be the set of admissible controls. Then, there exists an optimal control $u_1^* \in \mathscr{U}$ such that:*

$$J(u_1^*) = \min_{u_1 \in \mathscr{U}} J(u_1).$$

Proof. The existence of an optimal control is derived from the Filippov–Cesari Existence Theorem (Martcheva, 2015). All following conditions are verified:

Non-Empty Set of Admissible Controls and Boundedness: The set of admissible controls $\mathscr{U}$ is non-empty, as the control remaining constant as $u_1(t) = u_{\min}$ satisfying $u_1(t) \in [u_{\min}, u_{\max}]$ for all $t \in [0, t_f]$ can serve as an admissible control.

Compactness and Convexity of Control Values: It is clear that the control set $[u_{\min}, u_{\max}]$ is nonempty, compact and convex. The admissible controls are measurable functions whose values lie in this compact set. The state system is affine in the control, and its solutions are bounded in the invariant simplex. Moreover, the integrand is convex in u_1, and is bounded from below.

Affine Control Dynamics: The function is affine in the control u_1 because it can be written as a term independent of u_1 plus a term that is proportional to u_1.

Convexity and Lower Bound of the Running Cost: The running cost is $A_1 P + \frac{1}{2} B_1 u_1^2$. It has a lower bound of 0 since all terms are non-negative. The cost is also strictly convex with respect to the control u_1, since:

$$\frac{\partial^2}{\partial u_1^2} \left(A_1 P + \frac{1}{2} B_1 u_1^2 \right) = B_1 > 0$$

Since all conditions are satisfied, an optimal control $u_1^* \in \mathscr{U}$ exists. □

4 Numerical Methods and Results

A forward-backwards sweeping algorithm is used to find the numerical solution to the optimization control problem. For this method, the state equations are defined by initial conditions and the adjoint equations are defined by terminal conditions. Using an initial guess for the control function, the state equations are first solved "forward" from $t_i = 0$ to $t_f = 60$ minutes under the initial conditions that $x(0) = (1.00, 0, 0, 0)$ because the initial condition places all cohorts mass in the adaptive state. Using the state and adjoint solutions, an updated control profile from the optimality condition is calculated. This updated control is then combined with the previous control through a weighted average to improve stability. This process is repeated iteratively, numerically integrating the differential equations at each step until the variations of the state function and control profile are within a specific limit. The system is integrated numerically over a 60-minute period and continued iterations until the change in the control profile fell below the tolerance of 10^{-4}.

4.1 Baseline Dynamics

Under the passive recovery cue, a fixed constant cue determined through a weighted linear equation involving heart rate, heart rate variability, and stress score is used. The trigger which is a weighted combination of the normalized gaze fixation and gaze frequency is initiated at $t = 10$ as shown in Figure 1. As a result in Figure 2, individuals begin to move from the adaptive to perturbed state at $t = 10$ minutes. This eventually raises the proportion of individuals in the perturbed state to a peak of $P = 0.439$ before it decreases as the recovery cue causes a proportion of individuals to move towards the corrective state. This can be seen with the perturbed state line beginning to flatten before eventually decreasing at $t = 15$. At the end the proportion of individuals in all cohorts plateau to a fixed value.

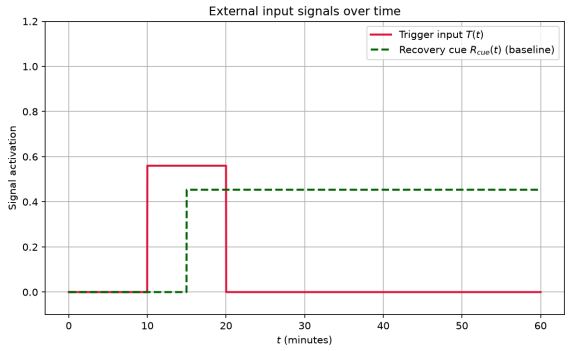


Figure 1: External input signals involving the craving trigger $T(t)$ (active over $10 \leq t \leq 20$ min) which equated to 0.560 and the baseline recovery cue $R_{\text{cue}}(t)$ (active starting at $t = 15$ min) which equated to 0.450.

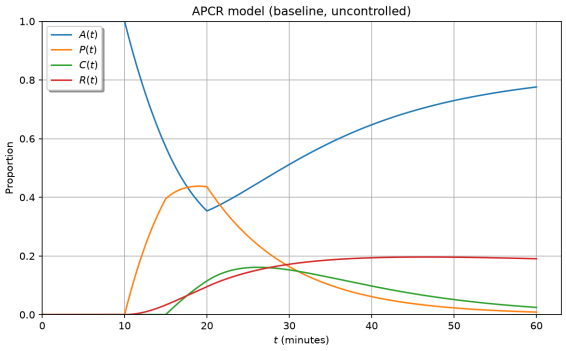


Figure 2: Compartment trajectories of the baseline dynamics under the passive cue.

4.2 Optimal Control and Convergence

The optimal control $u_1(t)$ responds to the trigger. Before trigger firing, the value is extremely small, but at $t = 10$ minutes, it experiences a sharp increase once craving is provoked. From there, it reaches a stationary value where the control intensity is 1. However, the control intensity does not stay at 1 for the remainder of the session. Following the trigger window, the control gradually decreases as the need for intervention decreases. The model does not use maximum intervention for the entire duration of the session, indicating that the optimal intervention is determined through the analysis of craving and trigger periods rather than being applied continuously. In addition, as the quantity of iterations increase, the cost function decreases indicating that the forward-backward sweep continuously improves the control strategy. Each iteration updates the control based on current solutions allowing the algorithm to identify effective controls. As the cost function continues decreasing, the optimization approaches a stable solution, suggesting that the control is converging towards an optimal intervention strategy.

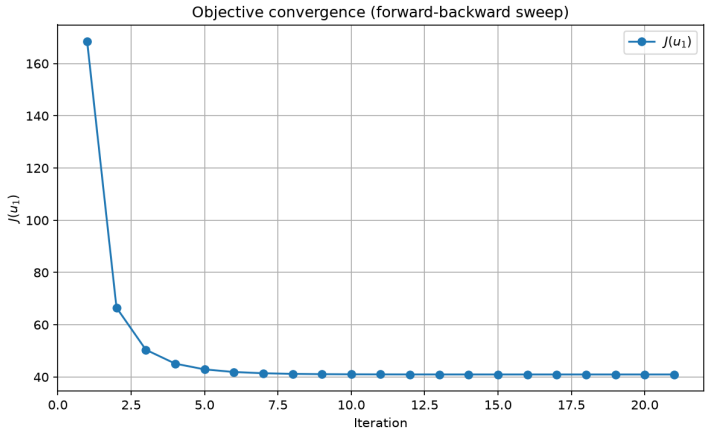


Figure 3: Convergence of the cost function $J(u_1)$ across forward-backward sweep iterations.

4.3 Effect of Control on Perturbed State

Compared to the passive baseline, the optimal control lowers the modeled peak proportion of the craving. The peak Perturbed proportion falls from 0.439 to 0.293, representing a decrease of 33.26%. The decrease is due to the recovery cue now becoming a dynamic variable rather than a static value. This decrease in the Perturbed state (shown closely in Figure 5) also results in changes to the other states. The modeled proportion present in the Corrective state with the control is also significantly higher. (Figure 2 and Figure 4).

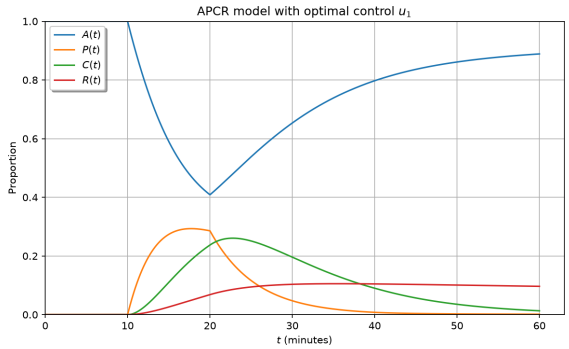


Figure 4: Compartment trajectories of the dynamics under the optimal control u_1.

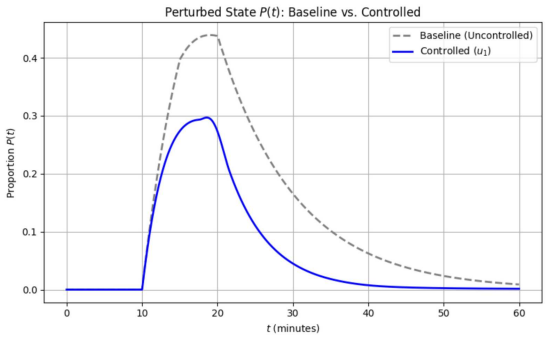


Figure 5: Comparison of the Perturbed proportion for baseline and optimal control.

4.4 Recovery Cue Comparison

The recovery cue was calculated through a weighted linear equation of physiological that was measured to trigger from $t = 15$ onwards. However, the recovery cue was parameterize with physical data not originating from a VR experiment. In addition, the exact time the trigger is released was chosen rather than justified based on clinical studies. In turn, I examined how the perturbed state would be affected if the recovery cue was set to the max value of $R_{cue} = 1$ or if R_{cue} was released starting at $t = 0$ on-wards rather than $t = 15$ in comparison to both the baseline and the optimal control. As shown in Figure 6, both the passive continuous recovery cue and the passive maximum recovery cue outperform the baseline. However, none of the 3 passive cues outperform the optimal control.

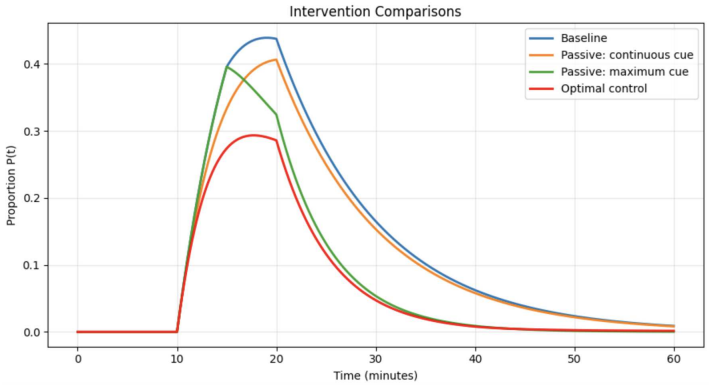


Figure 6: Perturbed state trajectory comparisons between the passive baseline, passive continuous cue, passive maximum cue, and optimal control. This shows that even under these circumstances the stable cue still doesn't outperform the optimal control.

4.5 Balance of Perturbed State with Weights

The performance of the optimal control is sensitive to the relative weighting factors assigned within the cost function. Figure 7 illustrates the dynamic response of the perturbed state variable $P(t)$ under varying weight allocations for state suppression (A_1) and control implementation cost (B_1). Adjusting these parameters reveals the fundamental trade-off between reducing the proportion of individuals in the perturbed state while minimizing the treatment cost and intensity that dictate a quadratic optimal control. Increasing the state penalty weight A_1 suppresses the peak proportion of individuals in P(t) and helps the

system recover much faster. However, raising A_1 to a certain point diminishes returns offering very little benefit despite requiring more therapeutic effort. On the other hand, as shown in Figure 8 increasing the control cost weight B_1 puts a heavier penalty on treatment expenses which causes the model to limit therapy costs. As a result, higher quantities of B_1 result in the Perturbed state approaching closer to the passive baseline. These tradeoffs assist in the identification of an approach that balances keeping the proportion of individuals in the Perturbed state low without overexerting on treatment.

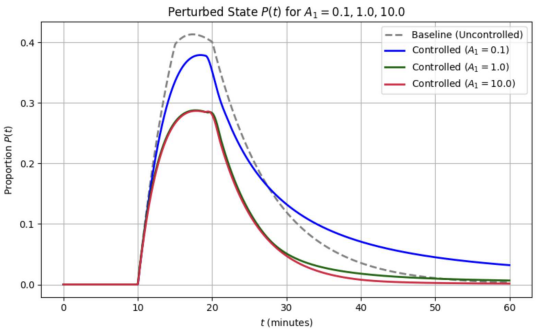


Figure 7: The effect of different state penalty weight (A_1) quantities of the quadratic optimal control on the perturbed state. Decreasing the value of A_1 increases the peak perturbed proportion as less of a weight is applied on suppressing the perturbed state.

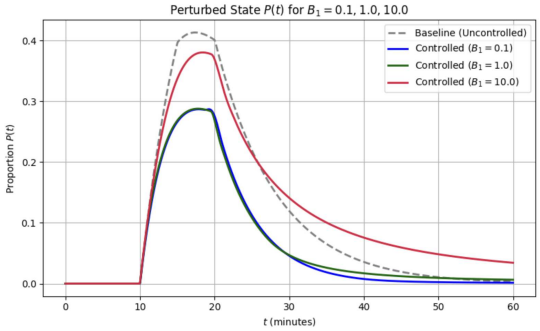


Figure 8: The effect of different control cost weight (B_1) quantities of the quadratic optimal control on the perturbed state. Increasing the value of B_1 increases the peak perturbed proportion as a heavier of a weight is applied on suppressing the cost. However, the effects diminish and become negligible as B decreases to an extent. $B_1 = 1$ and $B_1 = 0.1$ overlap.

4.6 Cost Analysis

Figure 9 and Figure 10 represent the expenditure of the treatment over the course of the 60 minute session. As shown in Figure 9 and 10, the instantaneous cost remains at 0 until the trigger starts the perturbation at $t = 10$ minutes. Once $t = 10$ both figures increase sharply. However, Figure 9 reaches a peak for a brief moment during the session after which the penalty cost slopes largely downwards before plateauing. This maximum treatment effort directly coincides with the highest proportion of the Perturbed state, where heavy intervention is necessary to suppress the peak reaction. Figure 10 reaches its peak of 0.05 at $t = 10$ after which the cost stays at that for almost thirty minutes before decreasing exponentially and plateauing at a cost of 0. As the Perturbed state subsides, the required VR intensity quickly declines towards zero. This demonstrates that high-intensity VR treatment is only needed during a narrow window rather than the full session allowing the total costs to be manageable.

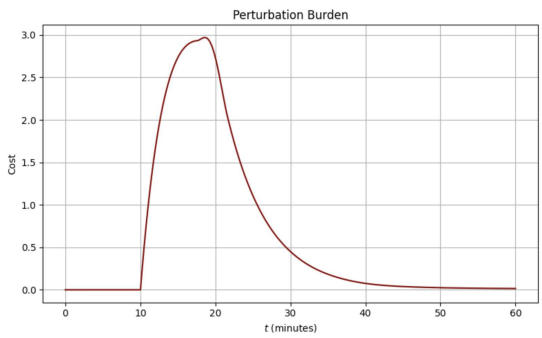


Figure 9: The model's penalty for remaining in the Perturbed state across the 60 minute session.

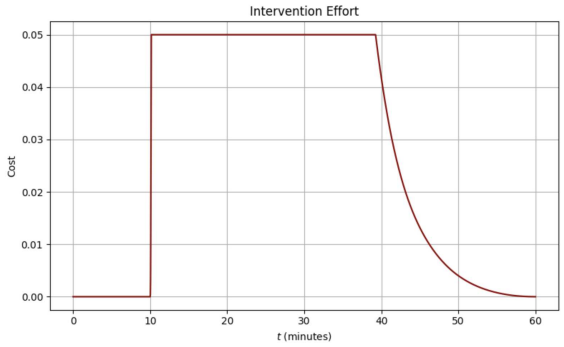


Figure 10: The mathematical cost applied to the intervention effort across the 60 minute session.

4.7 Sensitivity Analysis of Parameters

Because the parameter values reported in Table 1 are informed by literature and modeling assumptions rather than directly estimated by empirical data, it is important to assess how variations in these values affect the model's outcomes. A sensitivity analysis was thus performed by varying each model parameter independently to 50%, 100% and 150% its nominal value while holding the remaining parameters constant. For each parameter change, the resulting peak Perturbed proportion under the optimal control was recorded. The sensitivity analysis revealed that β and α had the greatest influence on the peak Perturbed proportion. Increasing β from 50% to 150% of its nominal value increased the peak Perturbed proportion from 0.190 to 0.364 an increase of 91.58%. While increasing α from 50% to 150% of its nominal value decrease the peaked proportion from 0.392 to 0.236 a decrease of 39.8% This correlation between parameter value and peak perturbed proportion for alpha is expected as alpha represents the transition out of perturbed and into the corrective state while β represents the transition from adaptive to perturbed. This indicates that larger transition rates into the Perturbed state substantially increases the model's peak. By comparison, γ and δ produced only minor changes in the intensity of the peak across the tested range, while the parameter μ showed a moderate inverse relationship.

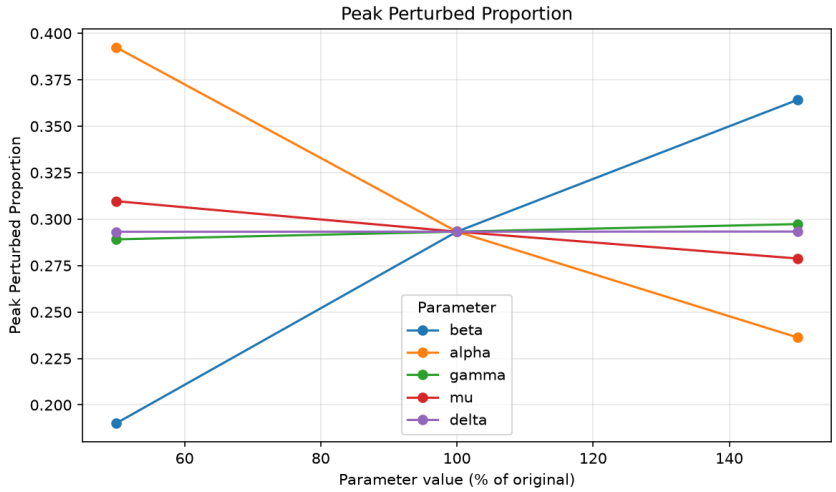


Figure 11: Peaked perturbed proportion under different parameter values.

Table 4: Sensitivity analysis of model parameters and their effect on the peak Perturbed proportion. Table created by the student researcher through Python, 2026.

Parameter	Parameter Value	Peak Perturbed
β	50% (0.10)	0.190170
β	100% (0.20)	0.293239
β	150% (0.30)	0.364138
α	50% (0.075)	0.392354
α	100% (0.15)	0.293239
α	150% (0.225)	0.236319
γ	50% (0.05)	0.289107
γ	100% (0.10)	0.293239
γ	150% (0.15)	0.297312
μ	50% (0.015)	0.309642
μ	100% (0.03)	0.293239
μ	150% (0.045)	0.278796
δ	50% (0.0025)	0.293181
δ	100% (0.005)	0.293239
δ	150% (0.0075)	0.293296

4.8 Validation of Population Conservation

To check that the model and solver is proper, the total mass $A+P+C+R$ is tracked over each run. The sum of the proportions remains equivalent to 1 for both the baseline and the controlled trajectories. This shows that throughout the 60 minute session no individuals entered or exited the system. They only moved between compartments. This numerically confirms the invariance of the simplex established in section 3.

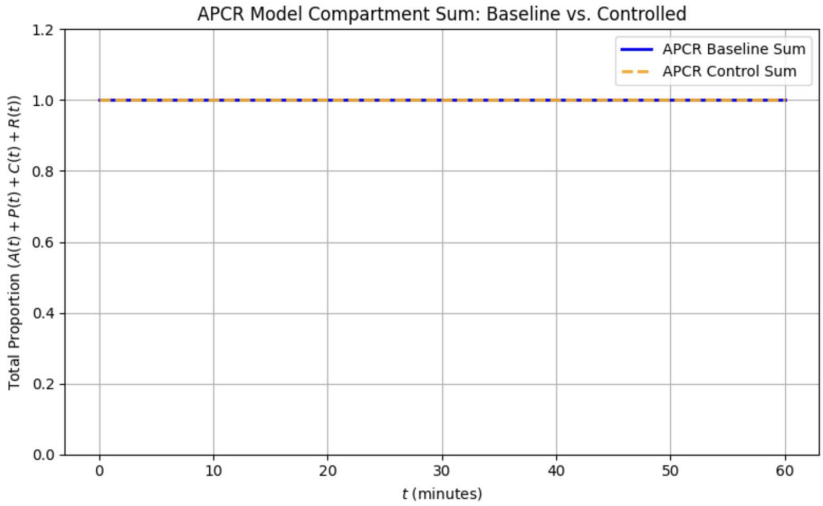


Figure 12: Compartment sums for the baseline and controlled trajectories.

5 Discussion and Conclusion

I developed an optimal control framework for adaptive VR intervention for substance use rehabilitation. The four-compartment APCR model represents transitions among cohorts between stable recovery, cue-induced perturbation, corrective intervention, and relapse. By replacing a stable recovery cue with a time-dependent control, the framework lets intervention intensity respond dynamically to need rather than follow a fixed schedule.

The numerical results demonstrate the potential benefit of this adaptive strategy. Relative to the passive-cue baseline, the optimal control reduced the peak proportion in the Perturbed state of the simulation from 0.439 to 0.293 representing a 33.26% decrease. The simulated results shown throughout the model suggest that an adaptive intervention is more effective than a static, prescribed recovery cue.

VR therapy is not applied at maximum intensity throughout the simulation. The intensity and marginal cost increase sharply following activation of the craving trigger and decrease as the need for intervention diminished. This illustrates the advantage of the state dependent formulation. Rather than being applied the whole time, the therapeutic effort is concentrated at periods where the proportion perturbed is the highest, when the risk is highest. The quadratic formulation of the cost function shows the balance between reducing perturbation and while limiting intervention intensity. Increasing the state weight (A_1) suppresses the Perturbed state but can cause negative effects on the cost both fiscally and physically. On the other hand, increasing the control weight (B_1) reduces the intensity of treatment which can reduce effectiveness. The optimal strategy depends on reducing perturbation while balancing the relative cost.

Several limitations should be considered when interpreting these findings. The APCR model mathematically represents the dynamics of the VR simulation but does not fully capture the complexities of real-world human behavior, including physiological and behavioral factors. In addition, the trigger and recovery cue signals that are defined by weighted combinations of behavioral and psychological measurements require additional empirical work to determine exact weights and calibration for individual users, as I used estimations. Due to limitations in human gaze data available in public databases, the trigger signal was parameterized only through alcohol data that may not reveal the full complexities of substance use disorder for other substances such as opioids and cannabis. The sensitivity analysis showed that the peak perturbed proportion was most responsive to β and α. Increasing β from 50% to 150% of its nominal value increased the peak from 0.190 to 0.364, while increasing α decreased it from 0.392 to 0.236. As a result deviations in β and α can cause significant changes to the peak perturbed proportion simulated. In comparison, γ and δ produced relatively small changes, indicating that the model's peak Perturbed proportion is more sensitive to the transition rates associated with entering and leaving the Perturbed state. Finally, the model is made under assumptions of a constant rate of transfer between state compartments whose values are approximated.

Despite these limitations, the framework provides a quantitative foundation for designing adaptive VR interventions that can respond to modeled changes in the recovery state. This model establishes a pathway towards personalized intervention strategies with adaptive therapeutic intensity based on individual responses through integration of behavioral and physiological signals into a state dependent controlled framework. Future experimental studies aim to use data to calibrate the exact model parameters to enable validation of the predicted optimal control strategy in VR settings. This framework represents a step towards transforming VR from a fixed therapeutic stimulus to a dynamically responsive

one that allocates treatment when and where it is needed the most.

References

Baca, A., Gonzalez, D., Ogueda, A. G., Matto, H. C., & Seshaiyer, P. (2024). Modeling, analysis and simulation of patient addiction journey. *CODEE Journal*, *18*(1), Article 4. Retrieved from https://scholarship.claremont.edu/codee/vol18/iss1/4/

Cannon, C. P. (2025, November). *How relevant is heart rate variability?* Harvard Health Publishing, Harvard Medical School. Retrieved from https://www.health.harvard.edu/heart-health/how-relevant-is-heart-rate-variability

Carter, B. L., & Tiffany, S. T. (1999). Meta-analysis of cue-reactivity in addiction research. *Addiction*, *94*(3), 327–340. Retrieved from https://pubmed.ncbi.nlm.nih.gov/10605857/

Conklin, C. A., & Tiffany, S. T. (2002). Applying extinction research and theory to cue-exposure addiction treatments. *Addiction*, *97*(2), 155–167. Retrieved from https://doi.org/10.1046/j.1360-0443.2002.00014.x doi: 10.1046/j.1360-0443.2002.00014.x

Epstein, D. H., Willner-Reid, J., Vahabzadeh, M., Mezghanni, M., Lin, J.-L., & Preston, K. L. (2009). Real-time electronic diary reports of cue exposure and mood in the hours before cocaine and heroin craving and use. *Archives of General Psychiatry*, *66*(1), 88–94. Retrieved from https://doi.org/10.1001/archgenpsychiatry.2008.509 doi: 10.1001/archgenpsychiatry.2008.509

Filippov, A. F. (1988). *Differential equations with discontinuous righthand sides*. Springer Netherlands. Retrieved from https://doi.org/10.1007/978-94-015-7793-9 doi: 10.1007/978-94-015-7793-9

Glavak-Tkalić, R., Šimunović, M., Perić Pavišić, K., Razum, J., & Colombo, D. (2025). Virtual reality in prevention and treatment of substance-related disorders: A systematic review of randomized controlled trials. *Clinical Psychology & Psychotherapy*, *32*(4), e70144. Retrieved from https://doi.org/10.1002/cpp.70144 doi: 10.1002/cpp.70144

Kober, H., Mende-Siedlecki, P., Kross, E. F., Weber, J., Mischel, W., Hart, C. L., & Ochsner, K. N. (2010). Prefrontal-striatal pathway underlies cognitive regulation of craving. *Proceedings of the National Academy of Sciences*, *107*(33), 14811–14816. Retrieved from https://doi.org/10.1073/pnas.1007779107 doi: 10.1073/pnas.1007779107

Labzai, A., Kouidere, A., Khajji, B., Balatif, O., & Rachik, M. (2020). Mathematical modeling and optimal control strategy for a discrete time drug consumption model. *Discrete Dynamics in Nature and Society*, *2020*, 5671493. Retrieved from https://doi.org/10.1155/2020/5671493 doi: 10.1155/2020/5671493

Martcheva, M. (2015). *An introduction to mathematical epidemiology* (Vol. 61). Springer. Retrieved from https://ndl.ethernet.edu.et/server/api/core/bitstreams/c1c06d6b-e3f3-4480-88f7-33dc26158e37/content

Mazza, M., Kammler-Sücker, K., Leménager, T., Kiefer, F., & Lenz, B. (2021). Virtual reality: A powerful technology to provide novel insight into treatment mechanisms of addiction. *Translational Psychiatry*, *11*(1), 617. Retrieved from https://doi.org/10.1038/s41398-021-01739-3 doi: 10.1038/s41398-021-01739-3

Simon, J., Quertemont, E., & Kelmendi, A. (n.d.). *Assessing attentional bias for alcohol-related cues using eye tracking in a virtual reality environment.* OSF project. Retrieved from https://osf.io/438mn/overview?view_only=c10c7c8e67aa48f5adf1327c58d894a4

Song, Y. (2022). *Positive invariance condition for continuous dynamical systems based on nagumo theorem.* arXiv preprint arXiv:2207.05429 [math.DS]. Retrieved from https://doi.org/10.48550/arXiv.2207.05429 doi: 10.48550/arXiv.2207.05429

Sowndarrajan, P. T., Shangerganesh, L., Debbouche, A., & Torres, D. F. M. (2022). Optimal control of a heroin epidemic mathematical model. *Optimization*, *71*(11), 3107–3131. Retrieved from https://doi.org/10.1080/02331934.2021.2009823 doi: 10.1080/02331934.2021.2009823

Substance Abuse and Mental Health Services Administration. (2024, July). *Key substance use and mental health indicators in the United States: Results from the 2023 National Survey on Drug Use and Health* (HHS Publication Nos. PEP24-07-021, NSDUH Series H-59). Rockville, MD: Center for Behavioral Health Statistics and Quality, Substance Abuse and Mental Health Services Administration. Retrieved from https://www.samhsa.gov/data/sites/default/files/reports/rpt47095/National%20Report/National%20Report/2023-nsduh-annual-national.pdf

Wireko, F. A., Ndogum, S., Nasiru, B. A., Ansah-Asamoah, D., Adu, I. K., & Asamoah, J. K. K. (2024). A dynamical model of drinking and smoking with optimal control analysis. *PLOS ONE*, *19*(12), e0311835. Retrieved from https://doi.org/10.1371/journal.pone.0311835 doi: 10.1371/journal.pone.0311835

Yfantidou, S., Karagianni, C., Efstathiou, S., Vakali, A., Palotti, J., Giakatos, D. P., ... Girdzijauskas, Š. (2022). LifeSnaps, a 4-month multi-modal dataset capturing unobtrusive snapshots of our lives in the wild. *Scientific Data*, *9*, 663. Retrieved from https://doi.org/10.1038/s41597-022-01764-x doi: 10.1038/s41597-022-01764-x

Zhang, S., Meng, Y., Chakraborty, A. K., & Wang, H. (2024). Controlling smoking: A smoking epidemic model with different smoking degrees in deterministic and stochastic environments. *Mathematical Biosciences*, *368*, 109132. Retrieved from https://doi.org/10.1016/j.mbs.2023.109132 doi: 10.1016/j.mbs.2023.109132

Zhang, Z., Zou, J., & Kundu, S. (2020). Bifurcation and optimal control analysis of a delayed drinking model. *Advances in Difference Equations*, *2020*, 522. Retrieved from https://doi.org/10.1186/s13662-020-02987-1 doi: 10.1186/s13662-020-02987-1