

# A Model-Based Feedback-Control Approach to Behavior Modification Through Reward-Induced Attitude Change

J. Ni, D. Kulić, and D.E. Davison

**Abstract**—We consider the problem of trying to change the behavior of a person by offering him a sequence of rewards, with each reward granted only if he performs the desired behavior to a certain degree. We develop a suitable cognitive model based on well-established psychological principles, and then use this model to devise a control algorithm that controls the person’s attitude. By exploiting psychological dynamics to drive his attitude to a large enough value in finite time, the person’s behavior can be brought to any desired level. This approach has the significant advantage that, at the end of the time interval, no reward is needed to maintain the desired behavior. We consider both open-loop and feedback implementations of the algorithm.

## I. INTRODUCTION

One of the strengths of control theory is its applicability in a wide variety of fields. A fascinating application area lies at the intersection of *social psychology* [1], which is the branch of psychology dealing with the dynamics of how people influence one another, and *behavior modification* [2], which refers to the set of techniques used to change a person’s behavior. Both of these fields deal with important health issues such as obesity and smoking.

The use of a control engineering mindset in psychology and behavior problems has a rich history (e.g., see [3]–[5] for a tiny sampling). In our previous work (see [6]–[8]) we focus on the system in Figure 1, in which a person (say a mother) offers a sequence of rewards (denoted  $R[k]$  at time sample  $k$ ) to induce another person (say her son) to perform a certain behavior towards which he initially has a negative attitude. In [6]–[8] we assume that the mother can monitor the behavior intent of the child, and we propose a double-integrator feedback control law for the mother to determine  $R[k]$ . The resulting feedback control law exhibits robustness, but it leads to undesirable transients and, significantly, the approach requires  $R[k] > 0$  for all  $k \geq 0$  [7]–[8].

This paper continues our research in [6]–[8]. The main objective is unchanged, but the approach is fundamentally different. To help keep the discussion concrete, suppose that the mother is trying, ultimately, to encourage her son to jog daily for  $B_d^*$  (where  $B_d^* > 0$ ) minutes. Assume  $k$  is measured in days, and let  $B[k]$  denote the actual number of minutes the child jogs on day  $k$ . The mother offers the child  $R[k]$  reward units (say, dollars) at the end of day  $k$ , contingent on the child jogging  $B_d[k+1]$  minutes (where  $B_d[k+1] \geq 0$ ) on

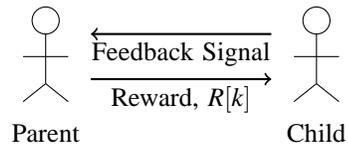


Fig. 1. The basic setup in which a parent offers a reward  $R[k]$  (with  $R[k] \geq 0$ ), to a child to drive the child’s behavior to a desired level.

the following day; that is, the child receives the reward  $R[k]$  if and only if  $B[k+1] \geq B_d[k+1]$ . The value of  $B_d[k+1]$ , which is determined by the mother at the end of day  $k$ , can vary with  $k$ , and it may be less than, equal to, or greater than  $B_d^*$ . Note that the child knows the value of  $B_d[k+1]$  but not the value of  $B_d^*$ . In contrast to our earlier work, in which  $B_d[k+1]$  is effectively fixed at zero, we consider here both  $R[k]$  and  $B_d[k+1]$  to be control signals. We make use of this extra degree of freedom to devise a control algorithm that guarantees (assuming no model uncertainty), after a finite number of days  $\bar{k}$ , two properties:  $B[k] \geq B_d^*$  for  $k \geq \bar{k}$  (i.e., the child’s behavior is driven to, or beyond, the desired level), and  $R[k] = 0$  for  $k \geq \bar{k}$  (i.e., no reward is required after day  $\bar{k}$ ). The first property is shared by the control algorithms in [6]–[8]. The second property, which is not shared by those algorithms, has great practical appeal. The keys to weaning the child off the reward are the use of  $B_d[k+1]$  as an additional control signal, and a focus on controlling the child’s internal attitude (denoted  $A_{out}[k]$ ) to indirectly affect his behavior. Accordingly, in this paper we assume the parent can measure  $A_{out}[0]$  and, if closed-loop implementation is desired,  $A_{out}[k]$  for other  $k$ . We omit discussion on how the attitude might be measured in practice other than noting that psychometrics (the field that includes attitude measurement) is well developed (e.g., see [9]).

A cognitive model of the child is given in Section II, building on the model development in [6]–[8]. The model is constructed from three foundational psychological theories: the *theory of planned behavior*, *cognitive dissonance theory*, and *overjustification theory*. The main difference from our earlier work is that the model now allows for  $B_d[k] > 0$ . The new control algorithm is developed in Section III. The algorithm drives the child’s cognitive state through two main stages: In Stage 1 we exploit the child’s cognitive dissonance dynamics to force his attitude positive, while in Stage 2 we exploit the same dynamics, while carefully avoiding overjustification effects, to make the child’s attitude sufficiently large that his behavior reaches  $B_d^*$ . Simulations and conclusions are given in Sections IV and V.

This research is supported by Natural Sciences and Engineering Research Council of Canada (NSERC).

The authors are with the Department of Electrical and Computer Engineering, University of Waterloo, Ontario, N2L 3G1, Canada. The corresponding author is D.E. Davison, ddavison@uwaterloo.ca.

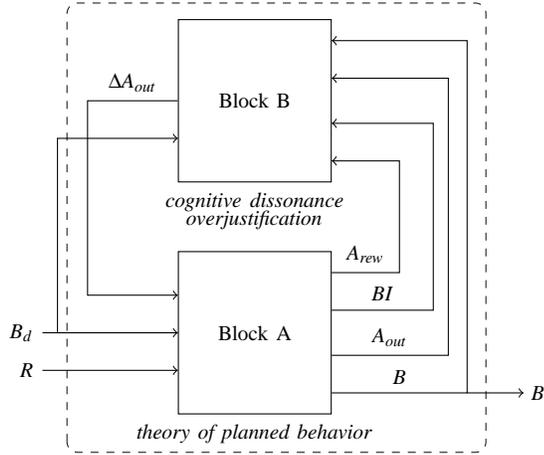


Fig. 2. The overall plant model (dashed box), decomposed into Blocks A and B. The system inputs are  $R[\cdot]$  and  $B_d[\cdot]$ , and the system output is  $B[\cdot]$ . This figure is revised from [6, Figure 1].

## II. SYSTEM MODEL

Building upon [6]–[8], a block diagram of the system model is provided in Figure 2. The model combines three psychological components:

### 1. Theory of Planned Behavior (Block A)

The *theory of planned behavior* [1][10][11] models how various beliefs and values (which we group together as *attitudes*) influence a person’s intent to behave, which in turn influences his actual behavior. We model these relationships as follows:

$$A_{out}[k] = A_{out}[k-1] + \Delta A_{out}[k-1], \quad (1)$$

$$A_{rew}[k] = r_1 A_{rew}[k-1] + \mu_1 (1 - r_1) R[k-1], \quad (2)$$

$$BI[k] = A_{out}[k] + A_{rew}[k], \quad (3)$$

$$B[k] = \begin{cases} B_d[k] & \text{if } BI[k] \geq B_d[k] \text{ and } A_{out}[k] \leq B_d[k] \\ A_{out}[k] & \text{if } (BI[k] < B_d[k] \text{ and } A_{out}[k] \geq 0) \\ & \text{or } A_{out}[k] > B_d[k] \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

In (1)  $A_{out}[k]$  denotes the child’s intrinsic attitude towards the behavior or, more precisely, towards an *outcome* of the behavior that the mother is trying to induce. (Outcomes of jogging might include improved health, tiredness, sweat, dirty clothes, improved social ties with other joggers, etc.) If there are multiple outcomes, each outcome is associated with a corresponding attitude; for simplicity in the following we consider a single outcome (e.g., improved health). Equation (1) describes how  $A_{out}[k]$  evolves, where  $\Delta A_{out}[k-1]$  arises from cognitive dissonance and overjustification pressures (Block B in Figure 2). A second attitude term,  $A_{rew}[k]$ , arises in (2) from the reward that is offered by the mother. Since we assume  $R[k]$  is measured in dollars, parameter  $\mu_1$  (with  $\mu_1 > 0$ ) in (2) captures how much the child values one dollar. As in [6]–[8] we assume first-order mental processing dynamics in (2), where  $r_1 \in [0, 1)$  determines the time constant. According to the theory of planned behavior,

attitudes sum to determine the child’s behavior intention ( $BI[k]$ ), as reflected in (3).

Signal  $BI[k]$  determines, in part, the child’s actual behavior ( $B[k]$ ) through (4). Three separate cases are included in (4). To help explain them, suppose  $B_d[k]$  is 30 minutes:

- If the child’s intrinsic attitude towards the jogging task is less than 30 minutes ( $A_{out}[k] \leq 30$ ), but the reward is large enough that his behavior intent is at least 30 minutes ( $BI[k] \geq 30$ ), then the child will accept the reward and do no more than what is required to accept the reward (i.e.,  $B[k] = 30$ ). This is the first case in (4).
- If the child’s intrinsic attitude towards the jogging task is more than 30 minutes ( $A_{out}[k] > 30$ ), then he will accept the reward but jog longer than what is required to earn the reward, and we assign  $B[k] = A_{out}[k]$ . Likewise, we set  $B[k] = A_{out}[k]$  in the situation where the child has a non-negative attitude towards the jogging task ( $A_{out}[k] \geq 0$ ) but his overall behavior intent is not enough to earn the reward ( $BI[k] < 30$ ). This is the second case in (4).
- The only remaining case is where the child has a negative attitude towards the jogging task ( $A_{out}[k] < 0$ ) and his intent to behave is not enough to earn the reward ( $BI[k] < 30$ ). In this case the child rejects the reward and, since he has no intrinsic desire to jog, he does not jog, i.e.,  $B[k] = 0$ . This is the third case in (4).

Note that  $A_{out}[k]$ ,  $\Delta A_{out}[k]$ ,  $A_{rew}[k]$ ,  $BI[k]$ ,  $B[k]$ ,  $B_d[k]$ , and  $B_d^*$  all have the same units (e.g., minutes of jogging for the scenario in this paper).

As a final comment, we recognize that the theory of planned behavior includes an aspect called *perceived behavioral control* (PBC) [11] that is ignored in our model. PBC models the idea that a person, despite having a high  $BI[k]$ , may actually not attempt the behavior because he thinks that he is not capable of doing the task, or that there is some obstacle preventing him from doing the task. For simplicity we assume here that the child is capable, and thinks he is capable, of successfully jogging for whatever duration is specified by the mother. Obviously this assumption breaks down in some scenarios (e.g., the child sprains his ankle).

### 2. Cognitive Dissonance Theory (Block B)

Cognitive dissonance theory [1][12][13] is one of the most celebrated theories in psychology. There are two aspects to the theory: first, if a person’s behavior is inconsistent with one of his attitudes, then an uncomfortable tension called *dissonance pressure* arises in the person, and, second, a person experiencing such tension will consequently try to reduce it either by changing his attitudes or his behavior.

For the setup considered here, dissonance pressure arises in the child if he declines the reward because he is turning down something that he values positively; on the other hand, if the child accepts the reward to jog more than he intrinsically wants to, then he again experiences dissonance pressure, but now because he is doing something that he does not enjoy doing. Following [13] and our previous work, we quantify dissonance pressure as the fraction of the child’s cognitive pairs that are inconsistent. To this end, we introduce

two indicator variables. First, define

$$B_{sgn}[k] = \begin{cases} +1 & \text{if } B[k] \geq B_d[k] \text{ or } A_{out}[k] \geq 0 \\ -1 & \text{otherwise.} \end{cases} \quad (5)$$

In words,  $B_{sgn}[k]$  is set to +1 if the child does any amount of jogging on day  $k$ , and to  $-1$  otherwise, with the conditions in (5) carefully stated to properly interpret the case  $B[k] = 0$  (which could arise, technically, in any of the three cases in (4)). Second, define

$$B_{rel}[k] = \begin{cases} +1 & \text{if } B[k] \geq B_d[k] \\ -1 & \text{otherwise.} \end{cases} \quad (6)$$

That is,  $B_{rel}[k]$  is set to +1 if, on day  $k$ , the child jogs at least the number of minutes required to earn the reward specified by the mother on the previous day, and to  $-1$  otherwise. Using these two indicator variables we can concisely compute the raw dissonance pressure of the child,  $P_{raw}^{CD}[k]$ , as follows:

$$M_{incon}^1[k] = \begin{cases} |A_{rew}[k]| & \text{if } \text{sgn}(A_{rew}[k]) \neq B_{rel}[k] \\ 0 & \text{otherwise,} \end{cases} \quad (7)$$

$$M_{incon}^2[k] = \begin{cases} |A_{out}[k]| & \text{if } \text{sgn}(A_{out}[k]) \neq B_{sgn}[k] \\ 0 & \text{otherwise,} \end{cases} \quad (8)$$

$$M_{con}^1[k] = \begin{cases} |A_{rew}[k]| & \text{if } \text{sgn}(A_{rew}[k]) = B_{rel}[k] \\ 0 & \text{otherwise,} \end{cases} \quad (9)$$

$$M_{con}^2[k] = \begin{cases} |A_{out}[k]| & \text{if } \text{sgn}(A_{out}[k]) = B_{sgn}[k] \\ 0 & \text{otherwise,} \end{cases} \quad (10)$$

$$M_{incon}[k] = \sum_{i=1}^2 M_{incon}^i[k], \quad M_{con}[k] = \sum_{i=1}^2 M_{con}^i[k], \quad (11)$$

$$P_{raw}^{CD}[k] = \begin{cases} B_{sgn}[k] \frac{M_{incon}[k]}{M_{incon}[k] + M_{con}[k]} & \text{if } M_{incon}[k] + M_{con}[k] > 0 \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

The fractional dissonance calculation appears in (12). The term  $B_{sgn}[k]$  in (12) is included to help formulate how the child reduces dissonance pressure.

As in our previous work, we keep the model reasonably simple by assuming that the child reduces dissonance pressure by changing only  $A_{out}$ , his intrinsic attitude towards jogging. (Other dissonance-reduction techniques, such as the introduction of new cognitions that are consistent with behavior, could be incorporated, but then an additional level of cognitive modeling would be required to deal with how the child decides on which technique is used in a particular situation.) In [6]–[8], where only binary behaviors are considered, it was simple to determine which direction  $A_{out}$  should be changed to reduce the raw dissonance pressure. The situation is more complicated here because it is possible for the child to desire to jog for a shorter duration than that which is required to earn the reward. Based on our previous work and the equations above, the typical situation would be for the child to reduce his dissonance pressure magnitude by increasing  $A_{out}$  if  $B_{sgn}[k] = +1$  and by decreasing  $A_{out}$  if  $B_{sgn}[k] = -1$ . However, it is possible for the situation to be reversed. For example, imagine the case where the child has a small positive intrinsic attitude towards jogging, and the mother offers a small reward for the child to jog for a very long duration; in such a case the child would decline

the reward. To reduce the resulting dissonance pressure, he could increase  $A_{out}$ , but he might instead “give up” the whole idea of jogging and decrease  $A_{out}$  substantially, making it highly negative, and that too will have the effect of reducing dissonance pressure. To capture this *attitude reversal* possibility, we introduce a new indicator variable  $r[k]$  which is +1 when reversal takes place and  $-1$  otherwise:

$$r[k] = \begin{cases} +1 & \text{if } B_d[k] - BI[k] > \alpha_{rev} A_{out}[k], A_{out}[k] \geq 0, \\ & K_1 P_{raw}^{CD}[k] > 2A_{out}[k], \text{ and } A_{rew}[k] > 0, \\ -1 & \text{otherwise.} \end{cases} \quad (13)$$

In (13), the condition  $B_d[k] - BI[k] > \alpha_{rev} A_{out}[k]$  models the idea that reversal takes place only if  $A_{out}[k]$  is small enough (for fixed  $B_d[k]$  and  $BI[k]$ ) or, conversely, only if  $B_d[k]$  is large enough (for fixed  $BI[k]$  and  $A_{out}[k]$ ). Parameter  $\alpha_{rev}$  represents the sensitivity of the child to exhibiting attitude reversal; a larger value is associated with a child who is less likely to “give up” on jogging. For the above argument to make sense,  $\alpha_{rev}$  should be at least 1; we use the range  $1 < \alpha_{rev} < 3$  in simulations. The condition  $K_1 P_{raw}^{CD}[k] > 2A_{out}[k]$  in (13) is required to ensure that the magnitude of attitude reversal, computed in (15) below, actually does result in a decrease in dissonance pressure.

As before, and as in our earlier work, we assume first-order mental processing dynamics (with time constant  $r_2 \in [0, 1)$ ) to map raw dissonance pressure  $P_{raw}^{CD}$  to the actual experienced dissonance pressure, denoted  $P^{CD}$ :

$$P^{CD}[k] = \begin{cases} (1 - r_2) P_{raw}^{CD}[k] & \text{if } r[k-1] = 1 \\ r_2 P^{CD}[k-1] + (1 - r_2) P_{raw}^{CD}[k] & \text{otherwise.} \end{cases} \quad (14)$$

Note that in the special case where attitude reversal occurs, we reset the dissonance pressure in (14) because it is unlikely that any residual  $P^{CD}$  will matter after such a dramatic change in psychological state.

Finally, assume the change in  $A_{out}[k]$  arising from cognitive dissonance reasons is proportional to dissonance pressure, with proportionality constant  $K_1 > 0$ :

$$\Delta A_{out}^{CD}[k] = \begin{cases} -K_1 P^{CD}[k] & \text{if } r[k] = 1 \\ +K_1 P^{CD}[k] & \text{otherwise.} \end{cases} \quad (15)$$

The term  $\Delta A_{out}^{CD}[k]$  affects  $\Delta A_{out}[k]$  through (21), given below.

### 3. Overjustification Theory (Block B)

Overjustification theory [1][14][15] applies when a reward is given to a person to do something that the person already enjoys doing. The theory states that such rewards are counterproductive in that they *reduce* the intrinsic desire of the person towards that behavior. Although the overjustification effect has been validated in various experiments, it is not nearly as well studied as is cognitive dissonance theory, and psychology researchers do not yet have a general model of the degree to which intrinsic desire is reduced. We denote by  $B_t[k]$  the minimal attitude level to which the overjustification effect can drive  $A_{out}[k]$ , and for simplicity we assume  $B_t[k]$  is a constant fraction of  $B_d[k]$ , i.e.,

$$B_t[k] = \alpha_{B_d} \cdot B_d[k], \quad (16)$$

for some constant  $0 < \alpha_{B_d} < 1$ . So, for example, if  $\alpha_{B_d} = 0.2$ ,  $A_{out}[k] = 20$  minutes, and  $B_d[k] = 30$  minutes (with  $R[k] > 0$ ), then (16) implies that overjustification pressure can reduce the child's intrinsic attitude from 20 minutes to as low as 6 minutes. If it happens that  $B_t[k] > A_{out}[k]$ , then our model implies that overjustification pressure does not decrease  $A_{out}$ .

To describe the above phenomenon mathematically, we introduce a relative attitude with respect to  $B_t[k]$ :

$$A_{out}^{rel}[k] = \max\{0, A_{out}[k] - B_t[k]\}. \quad (17)$$

Then the raw and filtered overjustification pressures, and the resulting change in intrinsic attitude, are computed just as in our previous work, but using  $A_{out}^{rel}$  instead of  $A_{out}$ , as follows:

$$P_{raw}^{OJ}[k] = \begin{cases} A_{out}^{rel}[k]A_{rew}[k] & \text{if } A_{out}^{rel}[k] > 0 \text{ and } A_{rew}[k] > 0 \\ & \text{and } B[k] \geq B_d[k] \\ 0 & \text{otherwise,} \end{cases} \quad (18)$$

$$P^{OJ}[k] = r_3 P^{OJ}[k-1] + (1-r_3) P_{raw}^{OJ}[k], \quad (19)$$

$$\Delta A_{out}^{OJ}[k] = \begin{cases} -K_2 P^{OJ}[k] & \text{if } K_2 P^{OJ}[k] \leq A_{out}^{rel}[k] \\ -A_{out}^{rel}[k] & \text{otherwise.} \end{cases} \quad (20)$$

In (18) the raw pressure is computed, then filtered in (19) using first-order mental processing dynamics, with time constant  $r_3 \in [0, 1)$ . The resulting change in  $A_{out}$ , computed in (20), is assumed to be proportional to  $P^{OJ}[k]$  with proportionality constant  $K_2 > 0$ . The effect is saturated in (20) so that attitude is not driven below  $B_t[k]$ . The final attitude change, denoted  $\Delta A_{out}^{OJ}[k]$  in (20), is added to the final attitude change that resulted from cognitive dissonance pressure, in (15), to complete the plant model:

$$\Delta A_{out}[k] = \Delta A_{out}^{CD}[k] + \Delta A_{out}^{OJ}[k]. \quad (21)$$

### III. PROPOSED CONTROL APPROACH

The overall model of the plant in Figure 2 is given in (1)–(21). We next describe an approach for the mother to compute, on day  $k$ ,  $R[k]$  and  $B_d[k+1]$  such that, after a finite interval  $\bar{k}$ ,  $B[k] \geq B_d^*$  for  $k \geq \bar{k}$  and  $R[k] = 0$  for  $k \geq \bar{k}$ . We assume the mother knows the various plant parameters ( $\mu_1$ ,  $r_1$ ,  $r_2$ ,  $r_3$ ,  $\alpha_{rev}$ ,  $\alpha_{B_d}$ ,  $K_1$ , and  $K_2$ ). The plant's initial conditions are

$$P^{CD}[0] = P^{OJ}[0] = A_{rew}[0] = 0, \quad A_{out}[0] = A_0^*, \quad (22)$$

where we assume the mother also knows  $A_0^*$ .

#### A. Controller Development

The proposed control approach is to drive the child's behavior,  $B[k]$ , to  $B_d^*$  in two stages, both of which emphasize control of the child's intrinsic attitude,  $A_{out}[k]$ . In Stage 1, a sequence of impulsive rewards is applied that, through dissonance pressure, results in  $A_{out}[k]$  increasing at a high rate until it becomes positive. In Stage 2, we again take advantage of dissonance pressure dynamics, but this time to increase  $A_{out}[k]$  to  $B_d^*$  which results in  $B[k]$  also being increased to  $B_d^*$ .

For simplicity, we assume  $r_1 = 0$  from now on. We allow for  $r_2 > 0$  by introducing an integer  $N \geq 2$  that can be

thought of as the settling time of the system. Specifically, if an impulsive reward is applied at time 0, a transient of the form  $(1 - r_2^{k-1})$  appears, and our approach is to wait for the transient to settle before applying the next impulsive reward is applied. In the special case when  $r_2 = 0$ , we use  $N = 2$ . Otherwise, if  $r_2 > 0$ , we use the approximation that after the transient has decayed by 98%, it can be ignored. Hence, the number of days to wait,  $N$ , should be chosen as

$$N = \begin{cases} 2 & \text{if } r_2 = 0 \\ \lceil \max\{2, 1 + \log_{r_2} 0.02\} \rceil & \text{if } r_2 > 0. \end{cases} \quad (23)$$

The algorithm is summarized in Figure 3. We next discuss Stages 1 and 2 in detail. If  $A_0^* < 0$ , the system starts in Stage 1; otherwise it starts in Stage 2.

*Stage 1:*  $A_{out}[k] < 0$  for  $k = 0, N, 2N, 3N, \dots$

During Stage 1, the mother applies a reward that (nearly) maximizes dissonance pressure, resulting in the child's attitude increasing at a high rate. Stage 1 is repeated until the attitude becomes positive. Lines 1-13 in the algorithm (Figure 3) implement Stage 1.

We start the analysis on day  $k$  assuming "initial conditions"  $P^{CD}[k] = P^{OJ}[k] = A_{rew}[k] = 0$ ,  $A_{out}[k] = A_0$ . The mother offers a reward of  $R[k]$  for the child to jog for  $B_d[k+1]$  minutes; the goal here is to determine suitable values for  $R[k]$  and  $B_d[k+1]$ . For cognitive dissonance to make  $A_{out}[k+1]$  larger than  $A_{out}[k]$ ,  $P_{raw}^{CD}[k+1]$  needs to be positive, which requires (see (12)) that  $B_{sgn}[k+1] = +1$ , i.e., the reward  $R[k]$  has to be large enough to force  $B[k+1] > 0$ . From (4),  $B[k+1]$  can be made positive only by ensuring  $BI[k+1] \geq B_d[k+1]$ . However, from (3) and (22) (and using the fact that  $A_{out}[k+1] = A_{out}[k]$  for the above initial conditions), the child's behavior intention on day  $k+1$  is

$$BI[k+1] = A_{out}[k+1] + A_{rew}[k+1] = A_0 + \mu_1 R[k]. \quad (24)$$

Consequently,  $BI[k+1] \geq B_d[k+1]$  holds if and only if

$$R[k] \geq \frac{B_d[k+1] + |A_0|}{\mu_1}. \quad (25)$$

The associated dissonance pressure, from (12), is

$$P_{raw}^{CD}[k+1] = \frac{B_{sgn}[k+1] \cdot |A_{out}[k+1]|}{|A_{out}[k+1]| + A_{rew}[k+1]} = \frac{|A_0|}{|A_0| + \mu_1 R[k]}. \quad (26)$$

Maximizing (26) subject to (25) results in  $B_d[k+1] = 0$  and  $R[k] = |A_0|/\mu_1$ . For improved robustness, it is better to set the desired behavior level slightly positive, so we take

$$B_d[k+1] = \epsilon \quad (27)$$

for small but positive  $\epsilon$ . We also set the reward to

$$R[k] = \frac{2B_d[k+1] + |A_{out}[k]|}{\mu_1} = \frac{2\epsilon + |A_{out}[k]|}{\mu_1}, \quad (28)$$

where the extra factor of 2 is included as an additional safety margin to allow for increased robustness against perturbations or noise that might result in  $R[k]$  being insufficient to drive  $B_{sgn}[k+1]$  positive, which will lead to a large *negative* attitude change instead of a large positive attitude change.

The above values of  $R[k]$  and  $B_d[k+1]$  appear in Lines 10 and 13 in Figure 3.

By iterating the plant dynamics, we find that the reward (28), assuming no reward is offered on days  $k+1$ ,  $k+2$ , etc., affects the child's attitude on day  $k+n$  (with  $n \geq 0$ ) as follows:

$$A_{out}[k+n] = \begin{cases} A_{out}[k] & \text{if } n = 0, 1 \\ A_{out}[k] + K_1 P_{raw}^{CD}[k+1](1 - r_2^{n-1}) & \text{if } n \geq 2. \end{cases} \quad (29)$$

Hence, if  $r_2 = 0$ , the child's attitude increases to  $A_0 + K_1 P_{raw}^{CD}[k+1]$  at time  $k+2$ , and then it remains fixed at that value. If  $r_2 > 0$ , the child's attitude converges to  $A_0 + K_1 P_{raw}^{CD}[k+1]$  exponentially, so we use the approximation that  $A_{out}[k+N] \approx A_0 + K_1 P_{raw}^{CD}[k+1]$ . Also note that  $P^{CD}[k+N] = 0$  if  $r_2 = 0$  (and  $P^{CD}[k+N] \approx 0$  if  $r_2 > 0$ ) and  $P^{OJ}[k+N] = A_{rew}[k+N] = 0$ , so the plant's state is suitable for Stage 1 to be executed again, on day  $k+N$ , if the child's attitude needs to be boosted beyond  $A_0 + K_1 P_{raw}^{CD}[k+1]$ .

For implementation purposes, it is useful to be able to compute, for  $R[k]$  as in (28) and  $B_d[k+1]$  as in (27), the value of  $A_{out}[k+N]$  on day  $k$ . The calculation is straightforward because the plant parameters are assumed to be known: substitute (28) into (26), and then (26) into (29) with  $n = N$  to obtain

$$A_{out}[k+N] = A_{out}[k] + K_1 \frac{|A_{out}[k]|}{2|A_{out}[k]| + 2\epsilon} (1 - r_2^{N-1}). \quad (30)$$

This calculation, which is used to keep track of how the child's attitude increases, appears in Lines 2 and 11 of Figure 3.

Two special situations need to be mentioned. First, if  $A_{out}[k]$  happens to be near zero, the reward calculation in (28) breaks down because the  $2\epsilon$  term starts to dominate the calculation. To avoid this difficulty, we add a condition to the algorithm that, if the expected value of  $A_{out}[k+N]$  satisfies  $|A_{out}[k+N]| < A_{th}$  (where  $A_{th} > 2\epsilon$  is a small threshold parameter), then, instead of using the reward (28), we use the reward value that forces  $A_{out}[k+N] = A_d$  where  $A_d = -A_{th}$  is the "desired" attitude. The value of  $R[k]$  that forces  $A_{out}[k+N] = A_d$  is determined using (26) and (29) (with  $n = N$ ):

$$R[k] = \frac{|A_{out}[k]|}{\mu_1} \left( \frac{K_1(1 - r_2^{N-1})}{|A_{out}[k]| + A_d} - 1 \right). \quad (31)$$

This special case appears in Lines 6-8 in Figure 3.

The second special situation, which arises only if  $K_1$  is very large or  $B_d^*$  is very small, concerns the possibility that the attitude  $A_{out}[k+N]$  exceeds the desired value,  $B_d^*$ . Although there might be no practical worry if the child's attitude is driven higher than desired, it is simple enough to determine the reward value that forces  $A_{out}[k+N] = A_d$  where  $A_d$ , the "desired" attitude, is now set to  $B_d^*$ . Indeed, the calculation is exactly the same as in the preceding paragraph, so the appropriate reward is (31) with  $A_d = B_d^*$ . This second special case appears in Lines 3-5 in Figure 3. If this special case arises, then there is no need to go on to Stage 2;

otherwise, once an impulse drives  $A_{out}[k]$  positive, Stage 2 is entered.

*Stage 2:*  $0 \leq A_{out}[k] < B_d^*$  for  $k = 0, N, 2N, 3N, \dots$

The goal of Stage 2 is to use the child's cognitive dissonance dynamics to increase his attitude further, until it reaches  $B_d^*$ . Again we use a sequence of reward impulses, with an impulse applied every  $N$  days. On day  $k$  we seek a reward  $R[k]$  that is small enough that it will be rejected by the child since this leads to cognitive dissonance pressure that will raise the child's attitude, and it also avoids exciting overjustification dynamics that would have the undesirable effect of lowering the child's attitude. In addition, we want to avoid the possibility of attitude reversal. Stage 2 appears in Lines 14-20 in Figure 3.

We start the analysis on day  $k$  assuming "initial conditions"  $P^{CD}[k] = P^{OJ}[k] = A_{rew}[k] = 0$ ,  $A_{out}[k] = A_0$ . To ensure the child rejects the reward  $R[k]$ , we force the second case in (4), i.e., force  $BI[k+1] < B_d[k+1]$ , which, from (2)-(3) and the assumption that  $r_1 = 0$ , is equivalent to

$$R[k] < \frac{B_d[k+1] - A_{out}[k]}{\mu_1}. \quad (32)$$

In addition, we deduce from (13) that attitude reversal is avoided on day  $k+1$  if  $R[k]$  is chosen such that  $B_d[k+1] - BI[k+1] \leq \alpha_{rev} A_{out}[k+1]$ , which, again from (2)-(3), is equivalent to

$$R[k] \geq \frac{B_d[k+1] - (\alpha_{rev} + 1) \cdot A_{out}[k]}{\mu_1}. \quad (33)$$

As long as the mother chooses  $B_d[k+1]$  such that

$$B_d[k+1] > A_{out}[k], \quad (34)$$

there exists a positive  $R[k]$  that satisfies both (32) and (33). By iterating through the plant dynamics, we can determine, for any such  $R[k]$ , the child's attitude on day  $k+N$ :

$$\begin{aligned} A_{out}[k+N] &= A_{out}[k] + K_1 P_{raw}^{CD}[k+1](1 - r_2^{N-1}) \\ &= A_{out}[k] + K_1 \frac{\mu_1 R[k]}{A_{out}[k] + \mu_1 R[k]} (1 - r_2^{N-1}). \end{aligned} \quad (35)$$

We see from (35) that  $A_{out}[k+N]$  is greater than  $A_{out}[k]$  for any positive choice of  $R[k]$ , and larger rewards lead to larger increases in attitude. In the limit  $R[k] \rightarrow \infty$  (subject to  $B_d[k+1]$  chosen large enough to satisfy (32)), the attitude increases to

$$A_{out}[k+N] = A_{out}[k] + K_1(1 - r_2^{N-1}). \quad (36)$$

The practical concern of keeping  $R[k]$  at a reasonable level cannot be ignored. By introducing a controller tuning parameter  $\beta \in (0, 1)$ , the aggressiveness of attitude increase can be adjusted:

$$A_d = \beta A_{out}[k] + (1 - \beta)(A_{out}[k] + K_1(1 - r_2^{N-1})). \quad (37)$$

In (37),  $A_d$  denotes the "desired" value of  $A_{out}[k+N]$ . If  $\beta \approx 1$ , the increase in attitude is small and  $R[k]$  is small; if  $\beta \approx 0$ , the increase in attitude approaches the maximum bound (36) and  $R[k]$  is very large. To get an expression for

the associated  $R[k]$ , set the left-hand side of (35) to  $A_d$  and solve for  $R[k]$ :

$$R[k] = \frac{A_{out}[k]}{\mu_1} \left( \frac{K_1(1-r_2^{N-1})}{A_{out}[k] + K_1(1-r_2^{N-1}) - A_d} - 1 \right). \quad (38)$$

One further detail is needed: to avoid driving the attitude higher than needed (i.e., beyond  $B_d^*$ ), we add a saturator to (37), as follows:

$$A_d = \min\{B_d^*, \beta A_{out}[k] + (1-\beta)(A_{out}[k] + K_1(1-r_2^{N-1}))\}. \quad (39)$$

The reward calculation based on (38) and (39) is included in Lines 15-16 of Figure 3.

Lastly, the value of  $B_d[k+1]$  needs to be determined. Substitution of (38) into (32) yields the inequality  $B_d[k+1] > B_{dmin}[k]$  where

$$B_{dmin}[k] = A_{out}[k] \left( \frac{K_1(1-r_2)^{N-1}}{A_{out}[k] + K_1(1-r_2)^{N-1} - A_d} \right). \quad (40)$$

Manipulation of (39) in (40) reveals that any  $B_d[k+1]$  satisfying  $B_d[k+1] > B_{dmin}[k]$  also satisfies (34), so (34) does not have to be explicitly included in the bound. In addition, substitution of (38) into (33) yields the inequality  $B_d[k+1] \leq B_{dmax}[k]$  where

$$B_{dmax}[k] = A_{out}[k] \left( \frac{K_1(1-r_2)^{N-1}}{A_{out}[k] + K_1(1-r_2)^{N-1} - A_d} + \alpha_{rev} \right). \quad (41)$$

Hence, any choice of  $B_d[k+1]$  satisfying

$$B_{dmin}[k] < B_d[k+1] \leq B_{dmax}[k]. \quad (42)$$

is acceptable. We introduce a second tuning parameter,  $\gamma \in (0,1)$ , to choose a particular  $B_d[k+1]$  within this range:

$$B_d[k+1] = \gamma B_{dmin}[k] + (1-\gamma)B_{dmax}[k]. \quad (43)$$

For robustness reasons, it is safest to use  $\gamma \approx 0.5$  to reduce the risk of violating either (32) or (33). The calculation of (40), (41), and (43) appears in Lines 17-19 of Figure 3. This completes Stage 2.

### B. Theoretical Guarantees

The algorithm in Figure 3, by construction, is meant to drive the child's attitude, and therefore behavior, to  $B_d^*$ . The following precise statement can be made:

**Theorem 1:** Assume the plant parameters  $\mu_1$ ,  $r_1$ ,  $r_2$ ,  $r_3$ ,  $\alpha_{rev}$ ,  $\alpha_{B_d}$ ,  $K_1$ , and  $K_2$  are known and that there is no plant uncertainty. For simplicity take  $r_1 = 0$ . Let the plant initial conditions be (22) and define  $N$  as in (23). Denote the ultimate desired behavior by  $B_d^*$ . For any controller parameters,  $\varepsilon > 0$  and  $A_{th} > 2\varepsilon$  satisfying

$$\frac{K_1(1-r_2^{N-1})}{4} > A_{th} + \varepsilon, \quad (44)$$

and any  $\gamma \in (0,1)$  and  $\beta \in (0,1)$ , there exists a  $\bar{k}$  such that  $B[k] \geq B_d^*$  and  $R[k] = 0$  for  $k \geq \bar{k}$ . There are three cases:

- If  $A_0^* \geq B_d^*$ , then  $\bar{k} = 0$  and  $B[k] = A_0^* \geq B_d^*$  and  $R[k] = 0$  for all  $k \geq 0$ .

- If  $0 \leq A_0^* < B_d^*$ , then

$$\bar{k} = N \cdot \left\lceil \frac{B_d^* - A_0^*}{(1-\beta)K_1(1-r_2^{N-1})} \right\rceil, \quad (45)$$

and  $B[k] = B_d^*$  and  $R[k] = 0$  for all  $k \geq \bar{k}$ .

- If  $A_0^* < 0$ , then if  $\varepsilon = 0$  necessarily  $\bar{k} \leq \bar{k}_1 + \bar{k}_2$  where

$$\bar{k}_1 = N \cdot \left\lceil \frac{|A_0^*|}{0.5K_1(1-r_2^{N-1})} + 1 \right\rceil, \quad (46)$$

$$\bar{k}_2 = N \cdot \left\lceil \frac{B_d^* - A_{th}}{(1-\beta)K_1(1-r_2^{N-1})} \right\rceil, \quad (47)$$

and  $B[k] = B_d^*$  and  $R[k] = 0$  for  $k \geq \bar{k}$ . For  $\varepsilon > 0$ ,  $\bar{k}$  exists but it may be larger than  $\bar{k}_1 + \bar{k}_2$ .

*Proof:* Condition (44) is needed to ensure that Stage 1 doesn't stall with  $A_{out}[k]$  "stuck" between  $-A_{th}$  and  $A_{th}$ . In the worst case,  $A_{out}[k]$  is driven to  $-A_{th}$ ; from (30), the next impulse results in an attitude increase of

$$K_1 \frac{|A_{th}|}{2|A_{th}| + 2\varepsilon} (1-r_2^{N-1}). \quad (48)$$

Forcing (48) to be greater than  $2A_{th}$  (which, after simple algebra, is equivalent to (44)) guarantees that  $A_{out}[k+N]$  will be driven to some value larger than  $A_{th}$ .

Now consider the three cases: First, if  $A_0^* \geq B_d^*$ , the algorithm jumps to Line 21 at  $k = 0$ , so  $\bar{k} = 0$  and  $B[k] = A_0^* \geq B_d^*$  for all  $k \geq 0$ . Second, If  $0 \leq A_0^* < B_d^*$ , the algorithm skips Stage 1. Each impulse in Stage 2, by (37), results in an increase in attitude equal to  $(1-\beta)K_1(1-r_2^{N-1})$ . Hence, the number of impulses required is  $(B_d^* - A_0^*) / [(1-\beta)K_1(1-r_2^{N-1})]$ , rounded up. Each impulse and subsequent waiting period uses  $N$  days. This leads to (45). Third, if  $A_0^* < 0$ , the algorithm executes in both Stage 1 and Stage 2. In Stage 1, if  $\varepsilon = 0$ , it follows from (30) that each impulse results in an increase in attitude equal to  $0.5K_1(1-r_2^{N-1})$ . Hence, the number of impulses required is  $|A_0^*| / [0.5K_1(1-r_2^{N-1})]$ , rounded up. Each impulse and subsequent waiting period requires  $N$  days. One additional impulse may be required because of the threshold  $A_{th}$  patch. Hence, the number of days needed for Stage 1 is no more than  $\bar{k}_1$  in (46). For Stage 2, the number of days required is the same as in (45) except the initial attitude is no longer  $A_0^*$  but now some value larger than  $A_{th}$ ; hence  $A_{th}$  is included in (47) as the worst-case value. For  $\varepsilon > 0$ , the number of days required to exit Stage 1 may increase, but detailed analysis of (30), omitted due to space limits, shows that Stage 1 will still be exited after a finite number of iterations (e.g., as a very conservative bound, Stage 1 will be exited after  $N \cdot \lceil |A_0^*| / 2A_{th} \rceil$  days for any  $A_{th} > 2\varepsilon$  and  $\varepsilon > 0$  satisfying (44)).  $\square$

Notice that (45), (46), and (47) are all decreasing functions of  $K_1$ , an intuitive relationship since a child who is more sensitive to cognitive dissonance pressure (larger  $K_1$ ) experiences greater attitude change after each impulse, so fewer impulses (smaller  $\bar{k}$ ) are required. Also note that (45) and (47) increase with  $B_d^*$ , reflecting the idea that additional impulses are needed if the ultimate desired behavior is larger.

Finally, (45) and (47) increase as  $\beta$  tends from 0 to 1, consistent with the intuition about the role played by  $\beta$ ; see the discussion around (37).

### C. Open-Loop Implementation

A straightforward open-loop implementation of the controller described by (27), (28), (31), (38), (39), (40), (41), and (43) is provided in Figure 3. The algorithm is executed for each  $k$  from  $k = 0$  until  $B_d$  reaches  $B_d^*$  at time  $k = \bar{k}$ . Reward impulses are applied on days 0,  $N$ ,  $2N$ , etc. The expected value of  $A_{out}[k]$ , for  $k = 0, N, 2N, \dots$ , is stored in variable  $\hat{A}_{out}^N[k]$ .<sup>1</sup> At  $k = 0$  the algorithm needs to be initialized with  $\hat{A}_{out}^N[0] = A_0^*$ . Between impulses, at times  $k = 1, \dots, N-1, N+1, \dots, 2N-1, 2N+1, \dots$ , the algorithm executes in Lines 22-24. During those intervals,  $R[k]$  is set to zero and  $B_d[k+1]$  is held at the value imposed at the previous impulse. Finally, the algorithm also executes Lines 22-24 once  $B[k]$  reaches  $B_d^*$ .

### D. Feedback Implementation

Intuitively, any open-loop implementation requires that the plant parameters be known accurately. Potentially, robustness can be improved through the use of feedback. Assuming the mother can measure  $A_{out}[k]$  for each day  $k$ , the algorithm can be converted into a feedback implementation by adding a “Line 0” to Figure 3 in which the value of  $A_{out}[k]$  measured by the mother on day  $k$  is stored in variable  $\hat{A}_{out}^N[k]$ . If there is no uncertainty or noise, the feedback implementation will behave the same as the open-loop implementation; if there is uncertainty or noise, we expect the feedback implementation to demonstrate better performance.

An additional benefit of the feedback implementation is flexibility: on days that  $A_{out}[k]$  is not measured, “Line 0” can be skipped and the algorithm will run as an open-loop implementation. There is also flexibility in the requirement that impulse rewards be applied on days  $k = 0, N, 2N, \dots$  since, over intervals where the mother cannot measure the child’s attitude or is not available to offer a reward with her son, the value of  $N$  can be chosen to be larger than in (23).

## IV. SIMULATION RESULTS

We simulated the plant, described by (1)–(21), and controller, described by (27), (28), (31), (38), (39), (40), (41), and (43) and in Figure 3, for various initial conditions and parameter values. Representative results are included in Figures 4–6. The plant and control parameters are identical in the three figures except the tuning parameter  $\beta$  varies from 0.25 to 0.5 to 0.75, showing how the Stage 2 aggressiveness varies with  $\beta$ . In each figure, the top plot shows the ultimate desired behavior ( $B_d^* = 50$  min), the child’s actual behavior ( $B[k]$ ), and the behavior level set by the mother ( $B_d[k]$ ). In all cases the child’s behavior is successfully driven to

<sup>1</sup>The hat and the superscript  $N$  are included in  $\hat{A}_{out}^N[k]$  to indicate that, for  $k = 0, N, 2N, \dots$ , the variable stores an estimated value of  $A_{out}[k+N]$ ; it is an estimate since, due to noise and uncertainty,  $\hat{A}_{out}^N[k]$  may in fact differ from  $A_{out}[k+N]$ . For graphing purposes in simulations, we keep  $\hat{A}_{out}^N[k]$  fixed during each inter-impulse interval (see Line 24), i.e., for any  $p \geq 0$ ,  $\hat{A}_{out}^N[k] = \hat{A}_{out}^N[pN]$  for  $k = pN, pN+1, \dots, (p+1)N-1$ .

```

1: if  $k \bmod N = 0$  and  $\hat{A}_{out}^N[k] < 0$  then ▷ Stage 1
2:    $\Delta \hat{A}_{out}^N[k] \leftarrow K_1 \left( \frac{|\hat{A}_{out}^N[k]|}{2|\hat{A}_{out}^N[k]|+2\epsilon} \right) (1 - r_2^{N-1})$ 
3:   if  $\hat{A}_{out}^N[k] + \Delta \hat{A}_{out}^N[k] \geq B_d^*$  then
4:      $R[k] \leftarrow \frac{|\hat{A}_{out}^N[k]|}{\mu_1} \left( \frac{K_1(1-r_2^{N-1})}{|\hat{A}_{out}^N[k]|+B_d^*} - 1 \right)$ 
5:      $\hat{A}_{out}^N[k+1] \leftarrow B_d^*$ 
6:   else if  $-A_{th} < \hat{A}_{out}^N[k] + \Delta \hat{A}_{out}^N[k] < A_{th}$  then
7:      $R[k] \leftarrow \frac{|\hat{A}_{out}^N[k]|}{\mu_1} \left( \frac{K_1(1-r_2^{N-1})}{|\hat{A}_{out}^N[k]|-A_{th}} - 1 \right)$ 
8:      $\hat{A}_{out}^N[k+1] \leftarrow -A_{th}$ 
9:   else
10:     $R[k] \leftarrow \frac{2\epsilon+|\hat{A}_{out}^N[k]|}{\mu_1}$ 
11:     $\hat{A}_{out}^N[k+1] \leftarrow \hat{A}_{out}^N[k] + \Delta \hat{A}_{out}^N[k]$ 
12:  end if
13:   $B_d[k+1] \leftarrow \epsilon$ 
14: else if  $k \bmod N = 0$  and  $\hat{A}_{out}^N[k] < B_d^*$  then ▷ Stage 2
15:    $A_d \leftarrow \min\{B_d^*, \beta \hat{A}_{out}^N[k] +$ 
16:      $(1 - \beta)(\hat{A}_{out}^N[k] + K_1(1 - r_2^{N-1}))\}$ 
17:    $R[k] \leftarrow \frac{\hat{A}_{out}^N[k]}{\mu_1} \left( \frac{K_1(1-r_2^{N-1})}{\hat{A}_{out}^N[k]+K_1(1-r_2^{N-1})-A_d} - 1 \right)$ 
18:    $B_{dmin} \leftarrow \hat{A}_{out}^N[k] \left( \frac{K_1(1-r_2^{N-1})}{\hat{A}_{out}^N[k]+K_1(1-r_2^{N-1})-A_d} \right)$ 
19:    $B_{dmax} \leftarrow \hat{A}_{out}^N[k] \left( \frac{K_1(1-r_2^{N-1})}{\hat{A}_{out}^N[k]+K_1(1-r_2^{N-1})-A_d} + \alpha_{rev} \right)$ 
20:    $B_d[k+1] \leftarrow \gamma B_{dmin} + (1 - \gamma) B_{dmax}$ 
21:    $\hat{A}_{out}^N[k+1] \leftarrow A_d$ 
22: else ▷ Between impulses or when done
23:    $R[k] \leftarrow 0$ 
24:    $B_d[k+1] \leftarrow B_d[k]$ 
25:    $\hat{A}_{out}^N[k+1] \leftarrow \hat{A}_{out}^N[k]$ 
26: end if

```

Fig. 3. Open-loop implementation of the control algorithm. The algorithm is executed for each  $k$  from  $k = 0$  until  $B_d$  reaches  $B_d^*$  at time  $k = \bar{k}$ . At  $k = 0$  the algorithm is initialized with  $\hat{A}_{out}^N[0] = A_0^*$ . The code has three parts: Stage 1 (Lines 1-13), Stage 2 (Lines 14-20), and a “no reward” section, activated between impulses or once  $B[k]$  reaches  $B_d^*$  (Lines 21-24).

$B_d^*$ . The middle plots show the reward  $R[k]$  offered by the mother at times  $k = 0, N, 2N, \dots$  (where  $N = 3$ ). The word “YES” is indicated above any reward that is accepted by the child, and the word “NO” otherwise. As expected, the rewards are accepted during Stage 1 and rejected during Stage 2. The total value of the accepted rewards for each of these three simulations is \$132.92. The third plot in each figure shows the child’s intrinsic attitude,  $A_{out}[k]$ , and the change in attitude resulting from cognitive dissonance,  $\Delta A_{out}^{CD}[k]$ . By design, there is no attitude change arising from overjustification pressure. Note that the value of  $\bar{k}$ , in each figure, is consistent with Theorem 1.

As a preliminary robustness study, we allowed for uncertainty in  $K_1$ . Denote the true plant parameter by  $K_1$ , and the associated value in the control algorithm by  $K_{1c}$ . All parameters are set as in Figure 5. Simulations reveal that  $B[k]$  is successfully driven to  $B_d^*$  when  $29.0 \leq K_{1c} \leq 30.4$  (for the open-loop implementation) and for  $0 < K_{1c} < \infty$  (for the feedback implementation). Analogous simulations

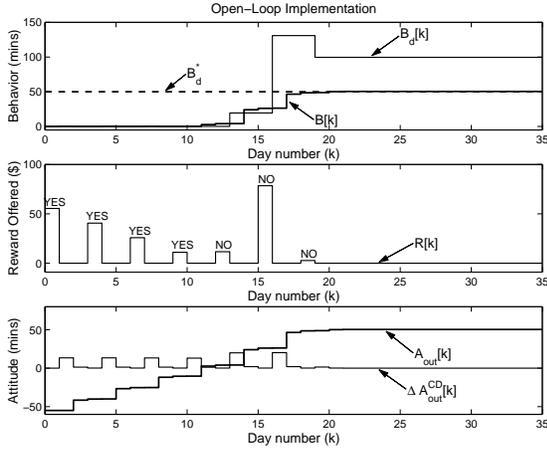


Fig. 4. System simulation with plant parameters  $\mu_1 = 1 \text{ min}/\$, r_1 = 0, r_2 = 0.1, r_3 = 0.5, \alpha_{B_d} = 0.8, \alpha_{rev} = 2, K_1 = 30 \text{ min}, K_2 = 0.1 \text{ min}^{-1}$ , control parameters  $\epsilon = 0.2, A_{th} = 2 \text{ min}, \gamma = 0.5, \beta = 0.25, N = 3, B_d^* = 50 \text{ min}$ , and initial conditions (22) with  $A_{out}^* = -55 \text{ min}$ .

for uncertainty in  $\mu_1$  reveal good performance when  $0.908 \leq \mu_{1c} \leq 1.003$  (for the open-loop implementation) and for  $0.501 \leq \mu_{1c} \leq 1.003$  (for the feedback implementation). As expected, the feedback implementation has better robustness.

## V. CONCLUSIONS

We have developed a novel model-based behavior-modification algorithm where both  $R[k]$  and  $B_d[k+1]$  are used as control signals to change a person's attitude, and, hence, behavior. Practical strengths of this approach are that no rewards are required in the long term, the transient behavior is good (e.g., no overshoot), each controller parameter is easily interpreted in terms of the plant parameters, and the timing of the control scheme is flexible. On the other hand, the approach requires good knowledge of some plant parameters (especially  $\mu_1$ ), as well as regular measurement of  $A_{out}[k]$  if a closed-loop implementation is used. Our current work focuses on experimental validation of the plant model and assumptions, as well as a systematic study of performance in the face of noise, uncertainty, and disturbances. We are also exploring how the model and controller can be modified to handle situations where the goal, instead of increasing desirable behaviors (e.g., exercising), is to reduce undesirable behaviors (e.g., smoking).

## REFERENCES

- [1] D. G. Myers and S. J. Spencer, *Social Psychology*. McGraw Hill Ryerson, 3rd Canadian ed., 2006.
- [2] G. Martin and J. Pear, *Behavior Modification: What it is and how to do it*. Saddle River, NJ: Pearson Prentice Hall, 2007.
- [3] W. T. Powers, *Behavior: The control of perception*. Aldine de Gruyter, 1973.
- [4] C. Carver and M. F. Scheier, *On the Self-Regulation of Behavior*. Cambridge University Press, 1998.
- [5] D. Rivera, M. Pew, and L. Collins, "Using engineering control principles to inform the design of adaptive interventions: A conceptual introduction," *Drug and Alcohol Dependence*, vol. 88(supp 2), pp. S31–S40, 2007.
- [6] R.-A. Vanderwater and D. E. Davison, "A dynamic control approach to studying the effectiveness of rewards in inducing behavior and attitude change," in *Proceedings of the IEEE International Conference on Control and Automation*, (Christchurch, New Zealand), 2009.

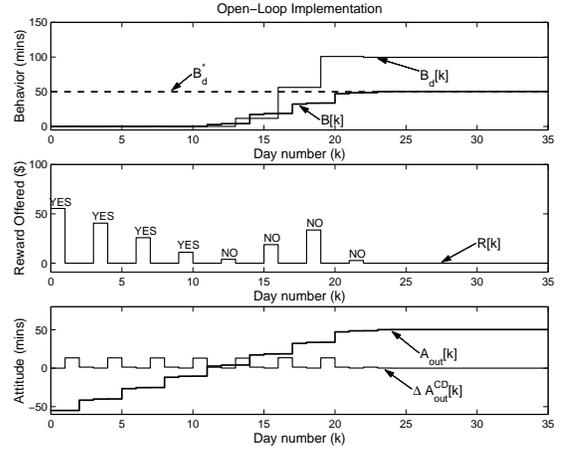


Fig. 5. System simulation using the same parameters as in Figure 4 except for  $\beta = 0.5$ . The Stage 2 behavior is less aggressive than in Figure 4.

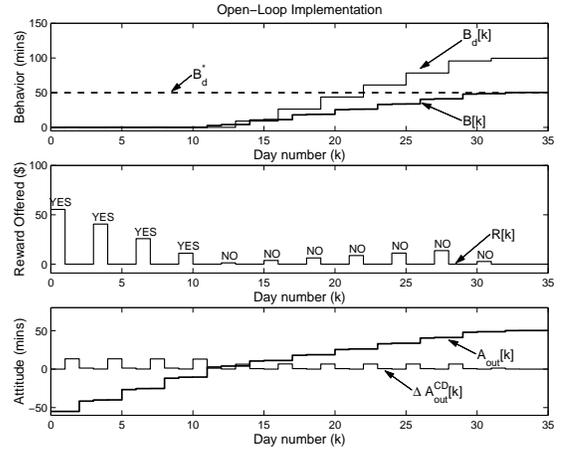


Fig. 6. System simulation using the same parameters as in Figure 4 except for  $\beta = 0.75$ . The Stage 2 behavior is even less aggressive than in Figure 5.

- [7] R. Vanderwater and D. Davison, "Using rewards to change a person's behavior: A double-integrator output-feedback dynamic control approach," in *American Control Conference*, (San Francisco), 2011.
- [8] D. Davison, R. Vanderwater, and Z. Zhou, "A control-theory reward-based approach to behavior modification in the presence of social-norm pressure and conformity pressure," in *American Control Conference*, (Montreal), 2012.
- [9] P. Erwin, *Attitudes and Persuasion*. Psychology Press, 2001.
- [10] I. Ajzen, *Attitudes, Personality, and Behavior*. Open University Press, 2nd ed., 2005.
- [11] I. Ajzen, "Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior," *Journal of Applied Social Psychology*, vol. 32, no. 4, pp. 665–683, 2002.
- [12] L. Festinger, *A Theory of Cognitive Dissonance*. Evanston, 1957.
- [13] J.-L. Beauvois and R. Joule, *A Radical Dissonance Theory*. Taylor and Francis, 1996.
- [14] E. L. Deci, "Effects of externally mediated rewards on intrinsic motivation," *Journal of Personality and Social Psychology*, vol. 18, no. 1, pp. 105–115, 1971.
- [15] E. L. Deci, R. Koestner, and R. M. Ryan, "A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation," *Psychological Bulletin*, vol. 125, no. 6, pp. 627–668, 1999.