2011 American Control Conference
on O'Farrell Street, San Francisco, CA, USA
June 29 - July 01, 2011

Using Rewards to Change a Person's Behavior: A Double-Integrator Output-Feedback Dynamic Control Approach

R. Vanderwater and D.E. Davison

***Abstract*— Using a nonlinear discrete-time model that captures how rewards affect a person's attitudes and behavior, we consider in this paper how to find a reward sequence that drives behavior to (or beyond) some desired level. We develop an output-feedback controller by re-arranging the plant equations so that the plant and controller form a linear feedback loop with a "disturbance" that captures all nonlinear effects. We then argue that the "disturbance" signal is rate limited, and propose a controller that contains two discrete-time integrators, a zero, and a saturator. Some simulation and analysis results are included to show that the control scheme is effective.**

I. Introduction

In this paper we deal with a novel application of control theory in the field of psychology. The basic setup, indicated in Figure 1, involves one person offering a reward sequence to a second person, with acceptance of the reward by the second person contingent on him or her carrying out some specific behavior. The use of rewards to influence behavior in such a manner is very common in the workplace, in politics, and in daily life.

To make the discussion concrete, assume that the person offering the reward is a mother and the person to whom the reward is being offered is her carnivorous child. Let's also say that the parent is trying to induce the child to eat his vegetables at supper, and that the reward being offered is measured in dollars. In terms of timing, we imagine that in the evening of every day (with the day number denoted by k, $k \geq 0$), the parent offers a reward ($R[k]$, $R[k] \geq 0$) to the child which he will receive if, during supper on the next day, he does indeed eat his vegetables. Over night, the child ponders the reward offer and, on day $k+1$, he wakes up with a certain attitude towards eating his vegetables ($A_{out}[k+1]$) and a certain attitude towards the offered reward ($A_{rew}[k+1]$). At supper on day $k+1$, the child either (a) eats his vegetables and accepts the associated reward ($B[k+1] \geq 0$), or (b) he does not eat his vegetables, and he is not allowed to take the reward ($B[k+1] < 0$). As explained below, the behavior $B[k+1]$ can lead to a change in attitude, $\Delta A_{out}[k+1]$. At some point after supper, the mother then offers a new reward ($R[k+1]$) to her child that deals with eating vegetables on day $k+2$, and the cycle repeats. Putting this all together, in control jargon, the parent is the controller, the child is the plant, $R[\cdot]$ is the control signal, and $B[\cdot]$ is the plant output. The goal, from the parent's perspective, is to control $B[\cdot]$.

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.

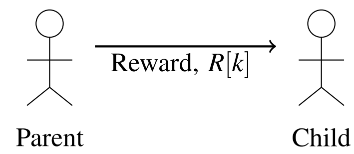

Fig. 1. The basic setup in which a parent offers a reward sequence, $R[k]$ (with $R[k] \geq 0$), to a child to induce a certain behavior.

In this work we emphasize the *dynamical* nature of the child's decision making. The importance of dynamical modeling is perhaps taken for granted by control engineers, but the vast majority of psychology researchers think in terms of "cause and effect" or "independent and dependent variable" paradigms, both which exclude important dynamical phenomena such as instability, limit cycles, and convergence. Important exceptions include work in *system dynamics theory* [1][2], where dynamic modeling of social and industrial situations is the focus, and *perceptual control theory* [3], where a dynamic control model of cognitive dynamics is the focus. Our work, which deals both with modeling and control, fits nicely within these two fields.

This paper continues our earlier work [4], where a discrete-time nonlinear dynamic model of decision making is developed. The underlying psychology and the (slightly simplified) model from [4] are concisely reviewed below; see [5] for a description of the model that is even more detailed than that in [4]. Two feedback control schemes were also proposed in [4], but they assume that the child's full cognitive state is measurable by the parent—hardly a realistic assumption. The main contribution of this paper is to develop an output-feedback controller, which requires measurement of only $B[k]$, to drive $B[\cdot]$ to (or above) a desired value, $B_d > 0$. Our controller is developed based on a linear approximation (not classic Jacobian linearization) of the plant, and we argue that good steady-state performance requires the use of two integrators. Some analysis, simulation results, and practical considerations are included.

II. System Model

A. Relevant Psychology and Model Development

Many introductory psychology textbooks (e.g., [6]) offer a thorough and very readable overview of the academic branches, modeling approach, and culture of the psychology research community. The model used here depends on the following three well-established psychological theories that relate behavior and attitude:

1) The Theory of Planned Behavior: The theory of planned behavior [6][7][8] is a popular model of how people make reasoned decisions about carrying out, or not carrying out, some behavior. The essence of the theory is that behavior is a function of various beliefs and values, grouped together here (for simplicity) under the name of *attitudes*. In our setup, we are concerned with the child's behavior, $B[k]$, and the only two attitudes in question are $A_{out}[k]$ (the child's present attitude towards the behavior, or more precisely an expected *outcome* of the behavior, that the mother is trying to induce) and $A_{rew}[k]$ (the child's present attitude towards the reward that is being offered). All of $B[k]$, $A_{out}[k]$, and $A_{rew}[k]$ can be positive or negative, and the magnitude indicates the strength of the behavior or attitude. Equation (1) below shows how $\Delta A_{out}[\cdot]$ (discussed below) drives $A_{out}[k]$, while (2) models how $A_{rew}[k]$ is formed from $R[k-1]$ (where $r_1 \in [0,1)$ determines the mental processing time constant and $\mu_1 > 0$ reflects the value that the child places on one dollar). Equation (3) shows how $B[k]$ is generated by summing the two attitudes. Note that $B[k]$, $A_{out}[k]$, and $A_{rew}[k]$ all have the same units, called "attitude units" here.

2) Cognitive Dissonance Theory: Cognitive dissonance theory [6][9][10] is possibly the most studied, and certainly one of the best known, theories in psychology. The theory has two main components: first, if a person's behavior is inconsistent with one of his attitudes, then a tension called *dissonance pressure* arises in the person, and, second, due to the discomfort of dissonance pressure, a person will take measures (typically changing his behavior and/or his attitude) to reduce the pressure. As an example, if on day k the child's attitude towards eating vegetables is negative (i.e., $A_{out}[k] < 0$) but the child accepts his mother's reward to eat his vegetables (i.e., $B[k] > 0$), he will experience dissonance pressure because he has just chosen to do something that he does not like doing; on the other hand, if the child declines the reward and does not eat his vegetables, he will experience dissonance pressure, not from any inconsistency between $B[k]$ (which is negative) and $A_{out}[k]$ (which is also negative), but because he is declining a reward that he values positively. To keep the model reasonably simple, we make the following assumption about dissonance pressure reduction:

Assumption 1: The child reduces dissonance pressure through only one way, namely, changing his attitude towards the behavior, $A_{out}[k]$.

Consistent with standard practice [10], the raw dissonance pressure, denoted $P^{CD}_{raw}[k]$, is computed as $M_{incon}[k]/(M_{incon}[k]+M_{con}[k])$, where $M_{incon}[k]$ is a measure of the total inconsistency between behavior and attitudes, and where $M_{con}[k]$ is a measure of the total consistency between behavior and attitudes; see (4)–(9). The first "case" in (9) captures this relationship, while the second "case" in (9), to avoid division by zero, handles the special situation where $B[k] = 0$. The $\mathrm{sgn}(B[k])$ term is included in (9) so that $P^{CD}_{raw}[k]$ drives the attitude change $\Delta A^{CD}_{out}[k]$ (discussed below) in a direction that always *reduces* dissonance pressure. We model the mental processing of dissonance pressure in (10), where $P^{CD}[k]$ is the *experienced* dissonance pressure and $r_2 \in [0,1)$ indicates the processing time constant. Finally, the change in attitude resulting from dissonance pressure reduction, $\Delta A^{CD}_{out}[k]$, is computed as in (11), where $K_1 > 0$ is a parameter reflecting how much attitude change is induced for each unit of experienced dissonance pressure. The term $\Delta A^{CD}_{out}[k]$ is one of two terms that affect $\Delta A_{out}[k]$, as in (15).

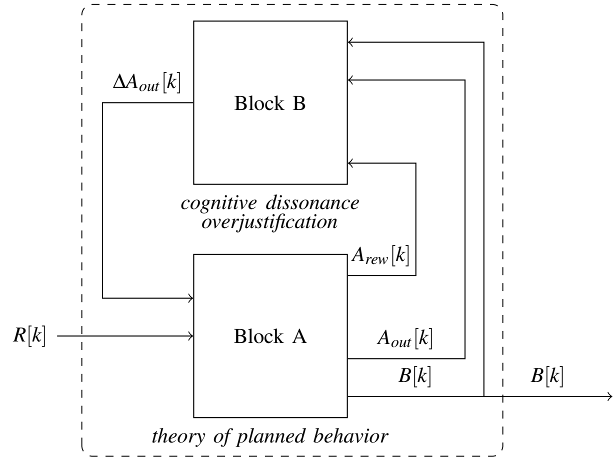

Fig. 2. The overall psychological system (dotted box), decomposed into Blocks A and B. The system input is $R[k]$ and the system output is $B[k]$. This figure is adapted from [4, Figure 1].

3) Overjustification Theory: Overjustification theory [11] deals with the special case where a reward is offered to a person to do something that the person already enjoys doing. In our case, this means overjustification theory applies only when $A_{out}[k] > 0$. The theory states that such rewards are counter-productive in that they *reduce* the intrinsic desire of the person towards that behavior. That is, if the mother offers a reward to the child when $A_{out}[k]$ is already positive, then the child will feel pressure (perhaps unconsciously) to reduce $A_{out}[\cdot]$. Our model captures this counter-intuitive, but thoroughly experimentally-verified, phenomenon in (12)–(15). In these equations, $P^{OJ}_{raw}[k]$ denotes the raw overjustification pressure, $P^{OJ}[k]$ denotes the experienced overjustification pressure (with mental processing time constant $r_3 \in [0,1)$), $\Delta A^{OJ}_{out}[k]$ denotes the attitude change that results from the experienced overjustification pressure, and $K_2 > 0$ is a parameter reflecting how much attitude change occurs per unit of pressure. Note that, to be consistent with the underlying psychology, (14) is arranged so that overjustification pressure can never force $A_{out}[\cdot]$ to be negative. Finally, the term $\Delta A^{OJ}_{out}[k]$ contributes to $\Delta A_{out}[k]$ through (15).

B. Mathematical Model and Parameter Values

The equations of the overall model are collected below, while the block diagram in Figure 2 shows how the various signals are connected.

Theory of Planned Behavior (Block A):

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

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

$$B[k] = A_{out}[k] + A_{rew}[k]. \tag{3}$$

Cognitive Dissonance Effects (Block B):

$$M^1_{incon}[k] = \begin{cases} |A_{rew}[k]B[k]| & \text{if } A_{rew}[k] > 0, B[k] < 0 \\ 0 & \text{otherwise,} \end{cases} \tag{4}$$

$$M^2_{incon}[k] = \begin{cases} |A_{out}[k]B[k]| & \text{if } A_{out}[k] < 0, B[k] > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

$$M^1_{con}[k] = \begin{cases} |A_{rew}[k]B[k]| & \text{if } A_{rew}[k] > 0, B[k] > 0 \\ 0 & \text{otherwise,} \end{cases} \tag{6}$$

$$M^2_{con}[k] = \begin{cases} |A_{out}[k]B[k]| & \text{if } A_{out}[k] < 0, B[k] < 0 \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

$$M_{incon}[k] = \sum_{i=1}^{2} M^i_{incon}[k], \quad M_{con}[k] = \sum_{i=1}^{2} M^i_{con}[k], \tag{8}$$

$$P^{CD}_{raw}[k] = \begin{cases} \operatorname{sgn}(B[k]) \frac{M_{incon}[k]}{M_{incon}[k]+M_{con}[k]} & \text{if } A_{out}[k] < 0, \\ & A_{rew}[k] > 0, \\ & \text{and } B[k] \neq 0, \\ \frac{|A_{out}[k]|}{|A_{out}[k]|+|A_{rew}[k]|} & \text{if } A_{out}[k] < 0, \\ & A_{rew}[k] > 0, \\ & \text{and } B[k] = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

$$P^{CD}[k] = r_2 P^{CD}[k-1] + (1-r_2) P^{CD}_{raw}[k], \tag{10}$$

$$\Delta A^{CD}_{out}[k] = K_1 P^{CD}[k]. \tag{11}$$

Overjustification Effects (Block B):

$$P^{OJ}_{raw}[k] = \begin{cases} A_{out}[k]A_{rew}[k] & \text{if } A_{out}[k]>0, B[k]>0, \\ & \text{and } A_{rew}[k]>0; \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

$$P^{OJ}[k] = r_3 P^{OJ}[k-1] + (1-r_3) P^{OJ}_{raw}[k], \tag{13}$$

$$\Delta A^{OJ}_{out}[k] = \begin{cases} -K_2 P^{OJ}[k] & \text{if } P^{OJ}[k] > 0 \\ & \text{and } K_2 P^{OJ}[k] \leq A_{out}[k]; \\ -A_{out}[k] & \text{if } P^{OJ}[k] > 0 \\ & \text{and } K_2 P^{OJ}[k] > A_{out}[k]; \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

Output of Block B:

$$\Delta A_{out}[k] = \Delta A^{CD}_{out}[k] + \Delta A^{OJ}_{out}[k]. \tag{15}$$

III. Controller Design

Although the plant equations (1)–(15) are fairly complicated and nonlinear, a key observation is that the system can be manipulated so that it fits within the classic linear time-invariant control framework with all nonlinear terms accounted for by a "disturbance." Figure 3 shows the system after such a manipulation, with the controller, yet to be designed, denoted $C[z]$ and the following transfer functions for the plant and disturbance filter:

$$P[z] = \frac{\mu_1(1-r_1)}{z-r_1}, \quad F[z] = \frac{1}{z-1}. \tag{16}$$

The plant transfer function follows from (2) and the disturbance filter transfer function follows from (1).

Our approach is to design the controller $C[z]$ to satisfy three criteria: (i) closed-loop stability of the loop in Figure 3; (ii) perfect steady-state tracking of the reference B_d (assumed

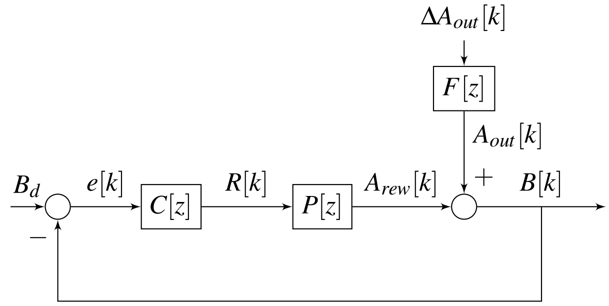

Fig. 3. Classic linear control feedback loop with plant $P[z]$, controller $C[z]$, and disturbance filter $F[z]$.

to be a constant); and (iii) rejection of the "disturbance" $A_{out}[k]$. The first two criteria are standard and accommodated using a linear controller with a single integrator:

$$C[z] = \frac{K_c}{z-1}. \tag{17}$$

Classic tools can be used to find suitable values for K_c to guarantee stability of the feedback loop in Figure 3. However, simulations reveal that, with the "disturbance" $A_{out}[k]$ included, the closed-loop system can be unstable. Consequently, further consideration of the nature of the "disturbance" is necessary. A starting point is to recognize that, according to (15), both cognitive dissonance pressure and overjustification pressure contribute to the signal $\Delta A_{out}[\cdot]$.

Focus first on the contribution of cognitive dissonance effects to $\Delta A_{out}[\cdot]$. By the form of the expressions in (9), $P^{CD}_{raw}[\cdot]$ is bounded above by 0.5 (see [5]). Moreover, the low-pass filter (10) is BIBO stable, so therefore the output of the filter, $P^{CD}[\cdot]$, is also bounded. It follows from (11) that $\Delta A^{CD}_{out}[\cdot]$ is bounded. Hence, the contribution of cognitive dissonance to $\Delta A_{out}[\cdot]$ is necessarily bounded. Next consider the overjustification pressure contribution to $\Delta A_{out}[\cdot]$. If raw overjustification pressure exists at times $k \geq \tilde{k}$, then, according to (12), necessarily $A_{out}[k]$, $B[k]$, and $A_{rew}[k]$ are all positive for $k \geq \tilde{k}$, which implies $M_{incon}[k] = 0$ for $k \geq \tilde{k}$, which in turn implies $P^{CD}_{raw}[k] = 0$ for $k \geq \tilde{k}$ and, by (10), $P^{CD}[k]$ decays to zero exponentially for $k \geq \tilde{k}$. Moreover, due to the signs in (12)–(14), $\Delta A^{OJ}_{out}[k] \leq 0$ for $k \geq \tilde{k}$, which has the effect of decreasing the positive signal $A_{out}[\cdot]$ according to (1). Putting together these two results, we deduce that $A_{out}[\cdot]$, for $k \geq \tilde{k}$, is bounded from above. By the structure of (14), it follows that $|\Delta A^{OJ}_{out}[\cdot]|$, for $k \geq \tilde{k}$, is also bounded above (in fact, it has the same bound as that of $A_{out}[\cdot]$). Hence, by (15), $\Delta A_{out}[\cdot]$, for $k \geq \tilde{k}$, is bounded.

Using the above discussion and recognizing that $F[z]$ in Figure 3 is an integrator, we have better understanding of the nature of the "disturbance" $A_{out}[\cdot]$: whether cognitive dissonance pressure or overjustification pressure is active, $A_{out}[\cdot]$ climbs or falls no sharper than a ramp. Moreover, we expect that, if the controller parameters are tuned such that the poles of the linear feedback loop in Figure 3 are well damped to minimize oscillations, $A_{out}[\cdot]$ looks very much like a ramp (when cognitive dissonance pressure dominates) or a constant (when overjustification pressure dominates). With this motivation, we use classical-control thinking to hypothesize that including two integrators in the controller should result in excellent accommodation of the "disturbance" $A_{out}[\cdot]$. The simplest two-integrator controller that can stabilize $P[z]$ is

$$C[z] = \frac{K_c(z-a)}{(z-1)^2}, \tag{18}$$

where $a \in (0,1)$ and $K_c > 0$ are controller parameters to be tuned. The Jury test [12] can be used to determine the values of a and K_c, as a function of μ_1 and r_1, that lead to closed-loop stability. The general result is complicated, but under the simplifying assumption $r_1 = 0$, the formula for the stabilizing a and K_c values is [5]

$$0.5 < a < 1, \quad 0 < K_c < \min\left\{\frac{4}{\mu_1(1+3a)}, \frac{2a-1}{\mu_1 a^2}\right\}. \tag{19}$$

In the time domain, the control signal for the controller in (18), $R[k]$, is computed as follows (where $e[k] := B_d - B[k]$):

$$R[k] = K_c e[k-1] - K_c a e[k-2] + 2R[k-1] - R[k-2]. \tag{20}$$

One final consideration is needed: the controller (20) could, in principle, result in $R[k] < 0$ at some k. However, the psychology undergirding our model applies only for $R[k] \geq 0$, so the control signal needs to be saturated. We choose the following form of saturation, which has the advantage of "stopping" the double integration within the controller whenever $R[k]$ saturates (i.e., an anti-windup mechanism):

$$\begin{aligned} R[k] \;=\; & \max\{0, K_c e[k-1] - K_c a e[k-2] + \\ & 2R[k-1] - R[k-2]\}. \end{aligned} \tag{21}$$

IV. Simulations and Analysis

We now study the performance of the double-integrator anti-windup controller (21) when it is connected to the plant (1)–(15). The thin curves in Figure 4 show a typical simulation output, with values for the plant parameters, controller parameters, initial conditions, and B_d indicated in the figure caption. Note that the child's attitude and behavior start off highly negative ($A_{out}[0] = B[0] = -40$), but the controller drives $B[k] \to B_d$ as $k \to \infty$, demonstrating that the control strategy successfully induces the child to eat his vegetables (in fact, at day 3, when $B[k]$ becomes positive). The simulation output also shows that a non-zero daily reward is required in steady-state to maintain $B[k]$ at B_d, a situation that is doubly undesirable: first, it means the parent, in order to have her child eat vegetables at the desired level, must continue to offer rewards to the child *ad infinitum*, and, second, the continual use of rewards drives $A_{out}[\cdot]$ (the child's intrinsic desire to eat vegetables) to zero through overjustification pressure. Note, however, that *any* control scheme that tracks B_d in steady-state will necessarily exhibit the same steady-state characteristics because there are only two mechanisms that drive $B[\cdot]$ greater than zero, namely the external reward $R[\cdot]$ and the internal attitude $A_{out}[\cdot]$, and the maximum value that can be achieved by $A_{out}[\cdot]$ is $K_1/2$ (see [5]), which is not enough (using the particular parameter values in this simulation) for $B[\cdot]$ to reach B_d.

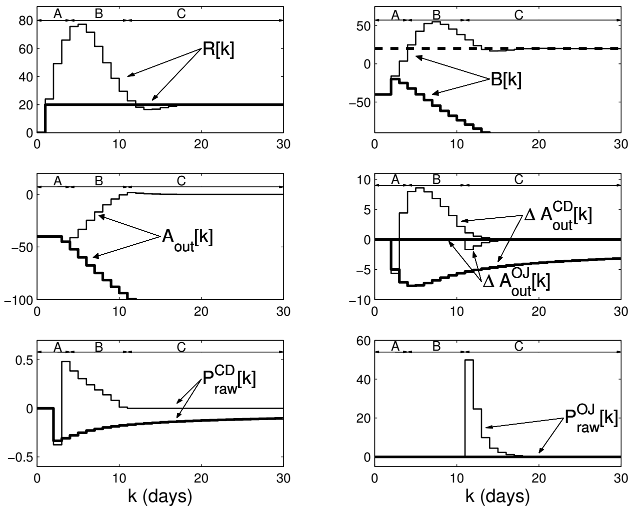

Fig. 4. Closed-loop simulation for the controllers (21) [thin curves] and (22) [thick curves] with the plant parameters $K_1 = 30$ attitude units, $K_2 = 0.1$ (attitude units)$^{-1}$, $r_1 = 0$, $r_2 = 0.5$, $r_3 = 0.5$, $\mu_1 = 1$ attitude units per dollar, the controller parameters $K_c = 0.4$ dollars per attitude units, $a = 0.95$, initial conditions $A_{out}[0] = -40$ attitude units, $A_{rew}[0] = 0$ attitude units, $P^{CD}[0] = 0$, $P^{OJ}[0] = 0$ (attitude units)^2, and the desired behavior $B_d = 20$ attitude units. The thick dashed curve indicates the desired behavior, B_d. The "A," "B," and "C" notation is explained in the text.

Figure 4 also shows, using thick lines, how the system behaves when the open-loop step-signal control law

$$R[k] = \frac{B_d}{\mu_1}, \quad \forall k \geq 0, \tag{22}$$

is used. This control law is much simpler than the feedback control law (21), and it provides the exact steady-state reward needed to keep $B[k]$ at B_d. However, as the figure shows, the controller (22) is totally ineffective, with $B[k]$ tending to $-\infty$ as $k \to \infty$. The problem with (22) is that the reward value is small enough that the child initially declines it, and the resulting cognitive dissonance pressure forces $A_{out}[\cdot]$ down, worsening the child's attitude towards eating vegetables; this effect worsens at every time step, as the child repeatedly declines the reward. In [5] we further pursue open-loop (and mixed open-loop and closed-loop) design approaches, and show how to make them more effective.

Mathematical analysis of the performance of controller (21) is challenging due to the nonlinearities of the plant, especially the switching that occurs when $B[\cdot]$ changes sign. Figure 5, in which the system is simulated six times with all parameters as listed in the caption of Figure 4, except for B_d which takes on different values, clearly demonstrates nonlinear characteristics.

Despite the complexity of the system, it is possible to carry out some analysis to support our conjecture that the controller (21) works well for any initial conditions as long as K_c and a are tuned so that the linear feedback system in Figure 3 is well damped. To this end, focus on the situation of most practical interest where the child starts with a negative attitude towards eating his vegetables and has no history of being offered rewards for such a task, i.e., the initial

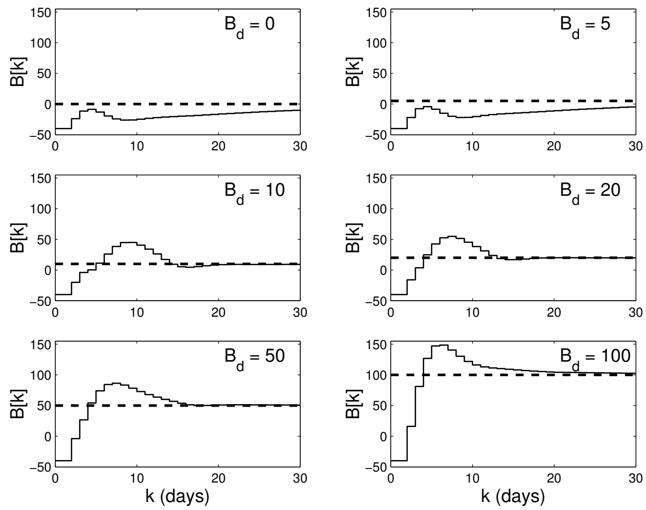

Fig. 5. Simulation output of the closed-loop system with controller (21). All settings are identical to those listed in the caption of Figure 4, except for B_d, which is set as indicated in the plots. The nonlinear nature of the system is evident.

conditions are

$$A_{out}[0] = A_0 < 0, A_{rew}[0] = 0, P^{CD}[0] = 0, P^{OJ}[0] = 0. \quad (23)$$

Under such initial conditions, the controller ideally drives the state trajectory through three stages with switching times denoted $\hat{k}$ and $\overline{k}$:

Stage A: for $0 \le k \le \hat{k} - 1$, $A_{out}[k] < 0$ and $B[k] \le 0$
Stage B: for $\hat{k} \le k \le \overline{k} - 1$, $A_{out}[k] < 0$ and $B[k] > 0$
Stage C: for $k \ge \overline{k}$, $A_{out}[k] \ge 0$ and $B[k] > 0$.

These stages are indicated on the plots in Figure 4. Implicit in Stage C is the desire for good tracking: either $B[k] \to B_d$ as $k \to \infty$, or $B[k]$ remains above B_d for sufficiently large k.

A. Analysis of Stage A

In Stage A, the controller ramps up the reward to try to pull $B[\cdot]$ positive—a substantial amount of reward is needed since cognitive dissonance pressure (which exists throughout Stage A) is pulling $A_{out}[\cdot]$ down, which, in turn, is trying to pull $B[\cdot]$ down. Simulations indicate that controller (21) always is able to pull the state out of Stage A into Stage B. It is also possible to prove that this occurs, at least if B_d is large enough, and under the simplifying assumption $r_1 = r_2 = 0$:

Lemma 1: Consider the controller (21) with zero initial conditions and parameters $a \in (0,1)$ and $K_c > 0$, and the plant with the initial conditions (23) and $r_1 = r_2 = 0$. If $B_d > K_1/(2\mu_1 K_c)$ then the state will necessarily be driven from Stage A to Stage B.

Sketch of proof (see [5] for details): The proof has three parts. First, it is established that, due to the integrators in the controller, $R[k]$ is an increasing function of k while in Stage A (i.e., for $0 \le k \le \hat{k} - 1$). This result, in turn, as well as the fact that $|P^{CD}[\cdot]|$ is bounded above by 0.5, implies that $B[k]$ cannot be a non-increasing function of k for all $k \ge 0$, i.e., there exists some $T^* \ge 0$ such that $B[T^*+1] > B[T^*]$.

Second, it is argued that, while in Stage A, $B[k]$ is an increasing function of k for all $k \ge T^*$. The proof again relies on the fact that $|P^{CD}[\cdot]|$ is bounded above by 0.5, and on the technical condition $B_d > K_1/(2\mu_1 K_c)$.

Finally, since $B[k]$ is an increasing function of k while in Stage A (with $k \ge T^*$), either $B[\cdot]$ eventually becomes positive (i.e., $\hat{k}$ is finite) or it converges (i.e., $\hat{k} = \infty$). However, once again using the fact that $|P^{CD}[\cdot]|$ is bounded above by 0.5, and using the result that $R[k]$ is an increasing function of k throughout Stage A, it is shown that, if $B[\cdot]$ does converge, it must converge to B_d, which is positive. So, in either case, $B[\cdot]$ eventually becomes positive, which means the state necessarily leaves Stage A at some finite time $\hat{k}$. Lastly, we argue that $A_{out}[k] < 0$ for all $0 \le k \le \hat{k}$, so, once leaving Stage A, the state must enter Stage B rather than Stage C. □

The condition on B_d in Lemma 1 is technical; simulations show that the result holds even if the condition is violated. Interestingly, the lemma is true even if K_c and a are tuned so that the linear feedback system in Figure 3 is unstable (although still subject to $a \in (0,1)$ and $K_c > 0$); such poorly-tuned controllers can lead to other problems, however, in Stage B and/or Stage C.

B. Analysis of Stages B and C

In Stage B, $B[\cdot]$ is positive, i.e., as of day $\hat{k}$, the child accepts the daily reward to eat his vegetables on the following day. The underlying psychology also switches from Stage A so that dissonance pressure now pulls $A_{out}[\cdot]$ up, which, in turn, boosts $B[\cdot]$ further. Consequently, the reward being offered by the mother can be substantially decreased, a trend that controller (21) exhibits, as demonstrated in Figure 4.

Simulations reveal that, once in Stage B, the state does not return to Stage A unless the controller is poorly tuned (in the sense that the closed-loop poles of the linear feedback system in Figure 3 have low damping or are unstable). Even if the state does return to Stage A, the integrators seem to always (possibly after multiple oscillations) pull it back to Stage B, but analysis of multiple switches becomes intractable. In cases where the state, once having left Stage A, stays in Stage B or switches from Stage B to Stage C, several results can be rigorously established:

Lemma 2: Consider the controller (21) with zero initial conditions and parameters $a \in (0,1)$ and $K_c > 0$, and the plant with the initial conditions (23) and $r_1 = r_2 = 0$. Assume $B_d > K_1/2$ and that the system enters Stage B at $k = \hat{k}$. If the state does not return back to Stage A, then:
(a) $A_{out}[k] \to 0$ as $k \to \infty$.
(b) If the system enters Stage C at some $k = \overline{k} > \hat{k}$, then it remains in Stage C for all $k \ge \overline{k}$.
(c) For all $\varepsilon > 0$, there does not exist a $\tilde{k}$ such that $B[k] \ge B_d + \varepsilon$ for all $k \ge \tilde{k}$.
(d) For all $\varepsilon > 0$, there does not exist a $\tilde{k}$ such that $B[k] \le B_d - \varepsilon$ for all $k \ge \tilde{k}$.

Sketch of proof (see [5] for details): The proof has five parts. First, result (b) follows immediately from (3) and the facts

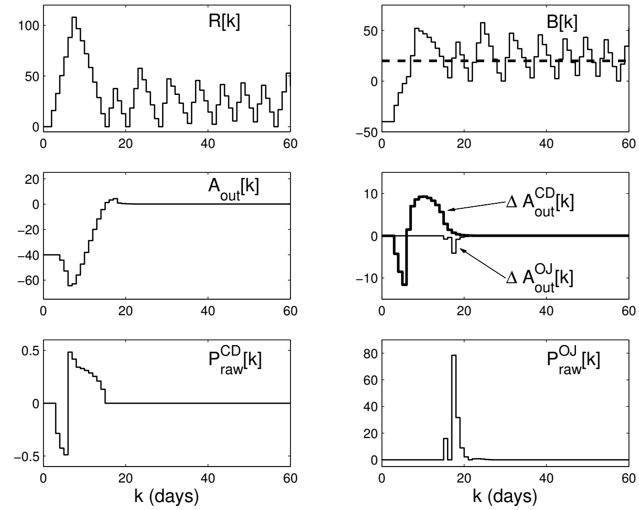

Fig. 6. Simulation output of the closed-loop system with controller (24). The sensor gain is $K_s = 40$, and all other settings are identical to those listed in the caption of Figure 4.

that $A_{rew}[k] \geq$ for $k \geq \bar{k}$ (due to the saturator in (21)) and $A_{out}[k] \geq 0$ for $k \geq \bar{k}$ (since overjustification pressure can never drive $A_{out}[\cdot]$ negative; see (14)).

Second, result (d) of the lemma follows, using a contradiction argument, because of the integrators in the controller (21). Specifically, suppose that for some $\varepsilon > 0$, there is a $\tilde{k}$ such that, for all $k \geq \tilde{k}$, $B[k] \leq B_d - \varepsilon$. Then the controller equation, (21), implies there exists an $N \geq \tilde{k}$ such that $R[N]$ is large enough to force $B[N+1]$ to be greater than $B_d - \varepsilon$, contradicting the original supposition. This proves (d).

Third, we argue that if the state remains in Stage B without switching to Stage C, then necessarily $P^{CD}[k] \to 0$ as $k \to \infty$. This, in turn, implies that either $R[k] \to \infty$ or $A_{out}[k] \to 0$ as $k \to \infty$, and we show that the former is impossible, so necessarily $A_{out}[k] \to 0$ as $k \to \infty$. This partly proves result (a).

Fourth, it is established that, if the state enters Stage C at $\bar{k}$, then $R[\cdot]$ does not converge to zero as $k \to \infty$. This result is used with a tedious argument involving the overjustification equations (12)–(15) to establish that $A_{out}[k] \to 0$ as $k \to \infty$. This completes the proof of result (a).

Finally, result (c) can be proved, like result (d), using contradiction and the fact that the controller contains two integrators. In particular, suppose that for some $\varepsilon > 0$, there is a $\tilde{k}$ such that, for all $k \geq \tilde{k}$, $B[k] \geq B_d + \varepsilon$. Equation (21) then implies that $R[\cdot]$ will, for some $N \geq \tilde{k}$, saturate at zero and remain at zero for all $k \geq N$. But this contradicts the earlier-proven fact that $R[\cdot]$ does not converge to zero as $k \to \infty$. This proves result (c). □

Lemma 2 implies that $B[\cdot]$ either converges to B_d or it oscillates around B_d without converging. Based on simulation investigations, we conjecture that the former (desirable) trend occurs whenever the controller parameters are tuned so that the linear feedback system in Figure 3 is stable and well damped; for all other K_c and a values, we conjecture, consistent with simulations, that $B[\cdot]$ oscillates around B_d without converging.

V. Practical Matters and Summary

We emphasize that (21) is easy to implement in practice: it requires measurement of only $B[\cdot]$, in contrast to the more complicated schemes in [4] that use state feedback. Practical quantitative measurement of psychological variables falls within the field of *psychometrics*, and psychometrists have developed well-tested methods to measure, specifically, attitudes and behaviors [7][8]. Having said this, a typical user of the controller (21) is unlikely to have the necessary expertise, time, or resources to measure $B[k]$ properly, which motivates the question as to how robust the controller is with respect to measurement errors. To help answer this question, we simulated the closed-loop system with various types of measurement errors. The most extreme type of error, arguably, is when the user senses only if $B[k]$ is above or below B_d, i.e., the controller effectively becomes

$$R[k] = \max\{0, K_c K_s \operatorname{sgn}(e[k-1]) - K_c a K_s \operatorname{sgn}(e[k-2]) + 2R[k-1] - R[k-2]\}, \quad (24)$$

where K_s is the "sensor gain." A typical simulation output is shown in Figure 6. Note that the performance is similar to that in Figure 4 other than the (expected) limit cycle that the binary sensing introduces.

To summarize, in this paper we have considered the dynamic control of a psychological problem that affects us all. Interestingly, due to the complexity of internal cognitive dynamics, the simplest output-feedback control scheme to effectively control behavior appears to require two integrators; simple proportional or single-integrator control will not work. We are currently extending this work to incorporate the influence of extra people (e.g., friends) on the person who is being controlled, and we are considering how Assumption 1 can be relaxed by introducing a high-level decision process to dictate the mechanism used to reduce dissonance pressure.

References

[1] J. W. Forrester, *Industrial Dynamics*. Productivity Press, 1961.
[2] J. D. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, 2000.
[3] W. T. Powers, *Behavior: The control of perception*. Aldine de Gruyter, 1973.
[4] 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), Dec. 2009.
[5] R.-A. Vanderwater, "A dynamic control approach to modeling and analysing the effects of rewards on behaviour and attitude change," MASc thesis, University of Waterloo, 2010.
[6] D. G. Myers and S. J. Spencer, *Social Psychology*. McGraw Hill Ryerson, 3rd Canadian ed., 2006.
[7] I. Ajzen, *Attitudes, Personality, and Behavior*. Open University Press, 2nd ed., 2005.
[8] M. Fishbein and I. Ajzen, *Predicting and Changing Behaviour*. New York: Psychology Press, 2010.
[9] L. Festinger, *A Theory of Cognitive Dissonance*. Evanston, 1957.
[10] J.-L. Beauvois and R. Joule, *A Radical Dissonance Theory*. Taylor and Francis, 1996.
[11] 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.
[12] C. L. Phillips and H. T. Nagle, *Digital Control System Analysis and Design*. Prentice Hall, 1995.