A Dynamic Control Approach to Modeling and Analysing the Effects of Rewards on Behaviour and Attitude Change

by

Elizabeth Vanderwater

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2010

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

Motivated by the dynamic feedback nature of human attitudes and behaviours, this work adopts a control systems engineering approach to studying a psychological system. In particular, two discrete-time nonlinear attitude-behaviour models are developed to describe how rewards influence a person's attitude and behaviour. The first model arises from three well-established social psychological theories: the theory of planned behaviour, cognitive dissonance theory and the overjustification effect. This model, called the *one-person system*, consists of a single person who is influenced (in a controlling manner) by a reward. The second model, called the *two-person system*, consists of two people, one of whom is influenced by a reward, and is obtained by extending the framework of the one-person system to incorporate the additional influence exerted by the behaviours of other people. Many interesting control problems arise for these two systems; this work focuses on the problem of finding a sequence of reward values that is capable of forcing the behaviour of a person (who has an initially negative attitude) to a certain, positive value. Open-loop and closed-loop controllers are designed to meet this control objective and simulations are used to confirm the functionality of these controllers. Additionally, throughout the system analysis, psychological predictions made by the models that may be of interest to social psychologists are highlighted.

Acknowledgements

I would like to thank my supervisor, Professor Daniel Davison, for his support and guidance. I would also like to thank Professors Kulić and Nielsen for their comments on this thesis. Finally, I would like to thank my family and friends for their endless patience and encouragement.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

To begin, the thesis motivation and problem statement are discussed. Then, the main contributions of the research are given, followed by the thesis outline.

1.1 Motivation

Psychology first emerged as a science at the end of the nineteenth century. By the beginning of the twentieth century, behaviourism become the predominant focus of major psychology research. Behaviourism, popularized by Ivan Pavlov and B.F. Skinner, flourished because it allowed researchers to use a scientific approach to study behaviour. Since this approach presumed internal mental states (e.g., attitudes, feelings, and emotions) were too subjective to evaluate with any degree of accuracy, researchers performed experiments measuring the observable behavioural response to a pre-determined stimulus. However, by the mid 1950s, criticism of behaviourism reached a tipping point, arguably through Noam Chomsky's critical review of Skinner's book "Verbal Behavior" [6]. In this review, Chomsky argues that language development in children does not follow the behaviourism model. This, and other criticisms, contributed to the *cognitive revolution*. The cognitive revolution was further promoted by the development of new, innovative computer technology. These new computers allowed scientists to view brain activity, thus providing evidence that internal mental states are indeed perceptible and, to some degree, quantifiable. Moreover, the information-processing approach used in communications engineering provided the foundation upon which researchers, such as Donald Broadbent, developed cognitive thinking models [5]. The cognitive revolution paved way to *cognitive psychology*, a term first used by Ulric Neisser [15], which considers people to be dynamic information-processing systems.

Since this time, other areas of psychology have developed cognitive-based theories, notably Festinger's contribution to social psychology, *cognitive dissonance theory* [9], in which aspects of the relationship between attitudes and behaviours are studied. Cognitive dissonance theory is not the only description of the attitude-behaviour relationship; other examples include Ajzen's *theory of planned behaviour* and Deci's *overjustification effect*. However, even though psychologists widely accept the dynamic nature of attitudes and behaviours and have developed single-stage, "cause and effect" attitude-behaviour models, a major deficiency in psychology is that these models predominantly neglect more involved dynamic phenomena such as feedback. The main premise of this thesis is that psychology as a science may benefit from dynamic attitude-behaviour models, much like other sciences have benefited from dynamic models (i.e., physics and chemistry). Other researchers agree with this premise and have made contributions towards dynamic modelling of psychological systems. *Systems dynamics theory* and *perceptual control theory* are examples of such contributions.

Systems dynamics originated through Jay Forrester's pioneering work at Massachusetts Institute of Technology. In [12], Forrester describes how the behaviour of large groups of people (i.e., society, organizations) can be dynamically modelled. Systems dynamics modelling use concepts such as first-order dynamics and feedback [17]. However, these models are not used for control purposes and most do not model dynamics of individuals within the group and instead, model the group as a whole. On the other hand, perceptual control theory, proposed by William T. Powers [16], does consider the dynamics associated with the behaviour of an individual. Specifically, this area of research focuses on modelling how an individual controls himself; thus, these models are not controlled by external factors, such as rewards and other people. A third research area that is somewhat related to modelling human behaviour is in the field of computer science, which typically focuses on large groups of people [13]. Generally, computer science models do not consider the underlying psychological dynamics that drive human behaviour and, instead, take a heuristic approach. Therefore, these computer science models are fundamentally different than those of systems dynamics and perceptual control theory.

This work has similarities to both the fields of systems dynamics and perceptual control theory. Like perceptual control theory, this work focuses on the underlying cognitive processes at the level of an individual person. Furthermore, like systems dynamics, this work models human behaviour as a key system output. The main difference between this work and these two research areas is that the models developed for this work are used for control purposes. Since this work considers human behaviour to be a key system output, from a psychological perspective, controlling the system output can be considered as influencing human behaviour. Finally, not only does this research contribute to dynamic attitude-behaviour modelling, it provides evidence that interesting control problems arise from this new control systems application.

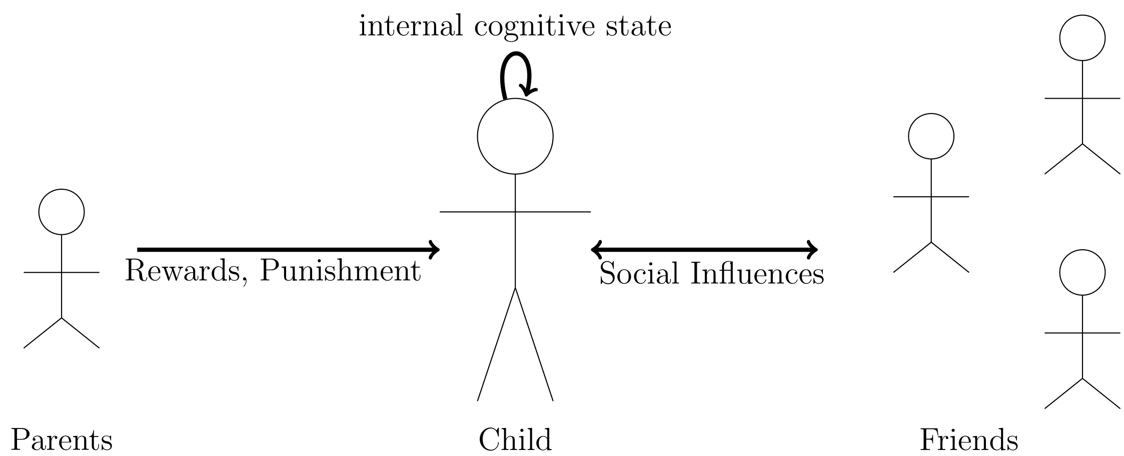

Figure 1.1: Three factors affecting a person's attitude and behaviour dynamics: internal cognitive state(s), rewards or punishments, and social influences.

1.2 Problem Statement

This thesis aims to construct and analyse two dynamic attitude-behaviour models, and thus, a *social psychological* approach is taken due to its focus on determining how a person's thoughts, beliefs, opinions, attitudes and behaviours are changed by external influences. The nonlinear, discrete-time dynamic attitude-behaviour models are applicable in any situation in which a tangible reward is used by one person in an attempt to overtly control the behaviour and/or attitude of another person. To assist with the development of our models, a specific example is considered: a child who does not want to play the piano. He has a negative attitude towards the act of playing the piano, and has no intention of carrying out this behaviour. How can this child's parents influence their son to practice playing the piano? Intuitively, one thinks of rewards and punishments. Psychology suggests that, indeed, behaviourist psychologists were right in that these controlling methods can be effective at influencing the child's behaviour. However, other factors such as social influences from friends and peers, in addition to the child's cognitive state, affect the child's behaviour [14]. Figure 1.1 outlines these possible external influences. At the center of Figure 1.1 is the child, who does not want to play the piano. The child's attitude and behaviour are dynamically influenced by several factors, represented by the arrows: rewards and punishments from parents, social influences from friends, and the child's cognitive state. The rewards and punishments are one-way influences, whereby rewards and punishments directly influence the child's behaviour and attitude (but not vice versa). The social influences are two-way, whereby a child both influences and is influenced by the behaviours of friends.

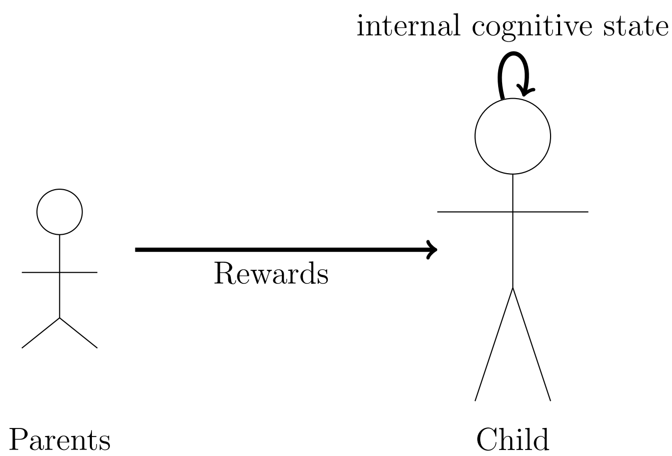

Figure 1.2: Conceptual representation of the one-person system.

This thesis considers two systems derived from Figure 1.1. Both models assume the parents offer their child a reward in an attempt to explicitly control him. The first assumes the child does not experience social influences. Other important psychological assumptions are also made but are discussed when the relevant concepts are introduced in future chapters. The first system can be considered within the typical plant-controller framework of control theory. The "plant" is a dynamic attitude-behaviour model of the child and the "controller" is the parents. Apart from the child's initial attitude and behaviour, which are negative, all of the plant's initial conditions are zero. Furthermore, since the parents influence their child through a sequence of rewards, the "control signal" is the reward. Finally, the child's behaviour is the "system output". Given that the plant contains the dynamics of one person, this system, conceptually shown in Figure 1.2, is referred to as the *one-person* system.

The conceptual representation of the second system, referred to as the *two-person* system, is shown Figure 1.3. This system drops the assumption that the child does not experience social influences. Instead, the child is influenced by the behaviour of a friend and likewise, the friend is influenced by the behaviour of the child. Therefore, the two-person system plant contains dynamic attitude-behaviour models of both the child, the friend and the effect these models have on each other. With the exception of each person's initial attitude and behaviour, all initial conditions are assumed to be zero. The initial attitude and behaviour of the child are negative and the initial attitude and behaviour of the friend are arbitrary. Similar to the one-person system, the control signal is the reward offered to the child (the friend is not offered a reward). Finally, the system output is the child's behaviour.

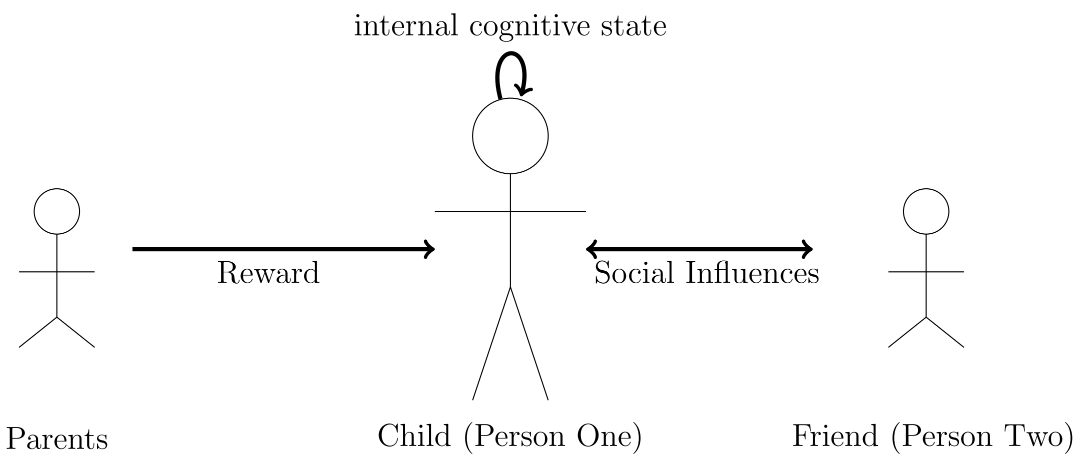

Figure 1.3: Conceptual representation of the two-person system.

While the formal problem definition is given after the relevant psychology is introduced and modelled, the fundamental question under consideration is whether or not it is possible to drive the child's behaviour to a desired, positive amount by offering a sequence of rewards. This problem is examined in the context of both the one-person system and the two-person system using both open-loop and closed-loop control strategies.

1.3 Contributions

This thesis contains three main contributions:

1. A new discrete-time dynamic attitude-behaviour model is proposed. The model is at least qualitatively consistent with known psychology.
2. The model predicts trends that have not yet been (fully) studied by psychologists. For example, the model predicts that
   (a) In both the one-person and the two-person systems, the known conclusion that a medium-sized reward produces the most amount of immediate attitude change is extended to predict that this reward subsequently causes the most amount of future attitude and behaviour change.
   (b) In the two-person system, known psychological phenomena arising when multiple people influence each other is produced by a mechanism that is traditionally not linked to these phenomena.

3. Studying a dynamic attitude-behaviour model demonstrates the existence of interesting control problems in this new control systems application.

1.4 Thesis Overview

Following this introduction, the thesis continues with an overview of the relevant social psychological concepts associated with the one-person problem. Using these concepts, a discrete-time, nonlinear model is developed and verified to be qualitatively consistent with these concepts. In Chapter 3, the formal problem statement is given, followed by an investigation of several open-loop and closed-loop control strategies. Chapter 4 extends the one-person plant to a two-person plant by introducing and applying social psychological concepts describing how individuals influence each other; then model verification and control strategies are discussed in Chapter 5. Finally, concluding remarks summarize the key results and highlight possible directions for future work.

Chapter 2

One-Person System: Psychology and Modelling

The thesis begins by exploring the one-person system: a child who does not want to play the piano. The child's parents wish to use a reward to influence their child to practice playing the piano. The problem at hand is whether or not there exists a sequence of rewards that will induce the child to play the piano. Before answering this question, a dynamic attitude-behaviour model of the child is developed from the relevant social psychology.

The key "signals" of the one-person system are the child's attitudes and behaviour. A well-known result of social psychology, the *theory of planned behaviour*, is used to model how the child's attitudes affect his behaviour. The theory of planned behaviour is a popular description of an intuitive conclusion of psychology research: attitudes influence behaviour. Social psychologists have found sufficient evidence to suggest the converse is also true: behaviour influences attitudes. Support for this less-intuitive result is found in *cognitive dissonance theory* and the *overjustification effect*. Clearly, these two conclusions suggest the attitude-behaviour relationship is dynamic. Indeed, psychology literature supports this suggestion [11]. Attitudes influence behaviours and vice versa; hence, from a control systems engineering perspective, the attitude-behaviour relationship forms a feedback loop. The feedback system block diagram is shown later (see Figure 2.3) after the necessary psychology and signals have been defined.

Sections 2.1 and 2.2 present the relevant psychology theories, which are used in Section 2.3 to develop our model. The end of Section 2.3 includes an interpretation of the signals in our model. After the model of the one-person system is developed, it is verified to be qualitatively consistent with the relevant psychology.

2.1 Psychology Describing How Attitudes Drive Behaviour

The first step to understanding the attitude-behaviour relationship is studying how attitudes combine to form a behaviour. In this section, justification is given to explain the psychological theory chosen for this component of our model. Following this justification, the theory is explained.

2.1.1 Background

Most people find it intuitive that actions follow attitudes, i.e., an individual's internal beliefs, thoughts, opinions and values will combine in some way to form his/her behaviour. However, this seemingly obvious result faced severe criticism in the 1970s, when many experimental studies suggested that behaviours do not necessarily follow attitudes. These criticisms sparked a flurry of research on determining *when* attitudes predict behaviour. In [14], researchers conclude that attitudes predict behaviour when

(i) external influences are minimized,

(ii) attitudes are specific to the behaviour under consideration, and

(iii) the individual considers the attitudes as being important.

Researchers such as Icek Ajzen and Martin Fishbein extended these results by developing models that incorporate the effect external influences have on behaviour formation. Indeed, their *theory of planned behaviour* considers several external factors affecting behaviour, including rewards and social influences [11]. Not only is this theory well established in the psychology literature, it has a wide application range and is simple enough for modelling purposes. For these reasons, the theory of planned behaviour is chosen to describe the influence attitudes have on behaviour.

2.1.2 Theory of Planned Behaviour

Icek Ajzen's theory of planned behaviour [2], [11] stems from his earlier work with Martin Fishbein: the *theory of reasoned action* [3]. Ajzen's theory has been verified through many experimental results and thus, is well established in the psychology literature. The theory of planned behaviour is a model of how people make reasoned decisions about performing or not performing a behaviour, based on various attitudes. For our model, attitudes are not affected by irrational influences on attitude, such as emotion and motivation (but could be considered in future work).

The theory of planned behaviour is composed of five elements: *behaviour*, *intent to behave*, *attitude towards the behavioural outcome*, *external influences* and *perceived behavioural control*. These five elements are described in detail below:

1. *Behaviour:* The behaviour is some action that an individual may or may not perform. For the piano example, the behaviour the child is considering is "I will accept the reward from my parents and practise the piano today." For our model, this decision is made at each sample k, measured in days.

2. *Intent to behave:* Before a behaviour can be realized, there must be some intention to perform the behaviour. Essentially, the intent to behave is the individual's assessment of the likelihood that he/she will engage in a behaviour. Statements such as "I will try to execute behaviour x," "I plan to execute behaviour x," and "I expect I will execute behaviour x" represent the intent to execute behaviour x [11]. An individual's behavioural intent may or may not yield the behaviour, but is a good indicator that the behaviour will be performed. An intent to behave may not lead to the behaviour due to unforeseen circumstances that could make the behaviour impossible. For example, suppose the child intends to accept the reward offered to him by his parents and practise the piano. If, after this intention has been formed, the child realizes the piano has been stolen, then the child is not able play the piano. Since, however, intention is a good predictor of behaviour, the following assumption is made:

   Assumption 2.1. *Behaviour and intent to behave are equal.*

   For this thesis, the terms *behaviour*, *behavioural intent* and *intent to behave* are used interchangeably and are denoted $B[k]$. An intent to perform the behaviour occurs when $B[k] \geq 0$ and an intent to not perform the behaviour occurs when $B[k] < 0$. Furthermore, the strength of this intention is represented by the magnitude $|B[k]|$. The notion of a positive behaviour and negative behaviours depends on the behaviour itself. For most behaviours, $B[k] \geq 0$ is demonstrated by carrying out the behaviour, while the behaviour strength is an indicator of how much time and/or effort is put into carrying out the behaviour. On the other hand, $B[k] < 0$ may not make sense for particular behaviours (as could be the case for the piano situation). However, for simplicity, we allow for negative behaviours. Investigating this issue in more detail is an item for future work. Finally, in practice there are various ways to measure $B[k]$ and the attitude elements ($A_{out}[k]$, $A_{rew}[k]$, described below), for example a 5-point Likert scale [11]; thus, it is possible to quantify these values. For our model, we begin with continuous values for $B[k]$. Considering discrete values for $B[k]$ is an item of future work.

3. *Attitude towards the behavioural outcome:* This element is the first of three contributors to behavioural intent and is related to how a person feels about performing the behaviour. Performing a behaviour results in one or more outcomes. Examples of behavioural outcomes of a child playing the piano may include: 1) feeling a sense of accomplishment; 2) becoming a more well-rounded person; and 3) becoming a social outcast. To keep our model simple, the following assumption is made:

   Assumption 2.2. *There is exactly one behavioural outcome.*

   Thus, suppose the child is concerned about only the first of the three aforementioned behavioural outcomes. At each sample k, the child forms an attitude towards this behavioural outcome, denoted $A_{out}[k]$. Following a popular approach to attitude modelling, the attitude consists of two factors: the likelihood the outcome will occur (the *expectancy*, E) and the importance the child places on the outcome (the *value*, V). Both the expectancy and the value range from positive to negative values. For example, if the child believes that it is highly likely that playing the piano leads to a sense of accomplishment, then $E > 0$. On the other hand, if the child's believes that playing the piano is unlikely to lead to a sense of accomplishment, then $E < 0$. Finally, if the child believes that it is neither likely nor unlikely that playing the piano leads to a sense of accomplishment, then $E \approx 0$. Moreover, if feeling a sense of accomplishment is very important to the child, then $V > 0$, whereas $V < 0$ if the child feels that it is entirely unimportant that he feels accomplished. Finally, $V \approx 0$ when a sense of accomplishment is neither important nor unimportant to the child. The internal attitude, A_{out}, is formed from the product EV, i.e., $A_{out} = EV$; thus from Table 2.1, a positive internal attitude could have two interpretations (likewise for a negative attitude). For simplicity, assume the child values a sense of accomplishment, i.e., $V > 0$. If, at sample k, the child believes that practising the piano certainly leads to a sense of accomplishment, then he has a high, positive attitude towards playing the piano (i.e., $A_{out}[k] \gg 0$). On the other hand, if, at sample k, the child feels the behavioural outcome is extremely unlikely, then his attitude towards the behavioural outcome is negative (i.e., $A_{out}[k] \ll 0$). Furthermore, $A_{out}[k]$ is more moderate for less extreme beliefs about the likelihood that performing the behaviour leads to the outcome in question. Finally, for simplicity, the terms *attitude* and *internal attitude* are used interchangeably when referring to $A_{out}[k]$.

4. *External influences:* This element is the second of three contributors to behavioural intent and is related to a key result of psychology research: external influences arising from rewards and other people affect an individual's behaviour. Although this

E	V	A_{out}
negative	negative	positive
positive	negative	negative
negative	positive	negative
positive	positive	positive

Table 2.1: Attitude formation using the expectancy-value model. A negative expectancy (E) means the behavioural outcome is highly unlikely, whereas a positive expectancy means the behavioural outcome is highly likely. A negative value (V) means the behavioural outcome is regarded as being totally unimportant, whereas a positive value means the behavioural outcome is regarded as being very important.

result appears obvious, most people underestimate the extent to which these external influences affect their behaviour, a phenomena termed by psychologists as *the fundamental attribution error* [14]. For the one-person system, one external influence is considered: the effect of a reward on the individual's behaviour. (For the two-person system, other external influences are also considered, as discussed in Section 4.1.) For our model, a monetary reward is used, where $R[k]$ denotes the number of dollars offered to the child, at time k, to entice him to play the piano. Furthermore, the extent to which the child values money, denoted μ_1, also contributes to the effect the reward has on the behavioural intent. This effect, is denoted $A_{rew}[k]$ and is referred to as the *attitude towards the reward* or *reward attitude*. Note that reward values are non-negative and thus $R[k] \geq 0$.

5. *Perceived behavioural control:* This element is the third of three contributors to behavioural intent and is related to the extent to which a person believes the behaviour can be executed. Two factors contribute to the perceived behavioural control: doubts a person may have regarding his/her ability to perform the behaviour and known obstacles that may prevent the behaviour. For example, the child may doubt his ability to play the piano if he hurt his hand, and thus, may have a lower behavioural intent. On the other hand, the child may recognize an actual obstacle, the absence of a piano for example, and thus have a lower behavioural intent. For simplicity, the following assumption is made:

   Assumption 2.3. *No doubts or obstacles arise that would reduce the behavioural intent.*

The theory of planned behaviour, conceptually shown in Figure 2.1, states that internal attitudes, external influences and perceived behavioural control together contribute to the formation of an intent to behave. Experimental results have shown that stronger internal

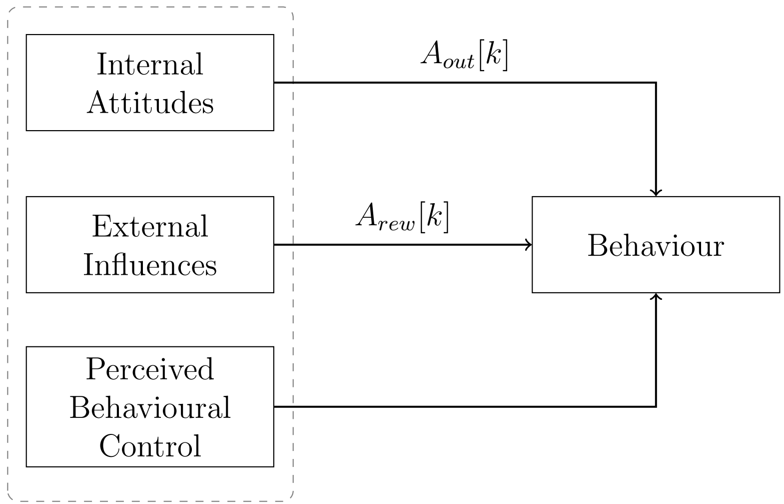

Figure 2.1: Conceptual representation of Ajzen's theory of planned behaviour [2].

attitudes and external influences contribute to stronger intentions, whereas perceived behavioural control has a "moderating" effect (i.e., a strong perceived behavioural control does not necessarily mean that a person will engage in the behaviour but a lower perceived behavioural control means a person is less likely to engage in the behaviour). Furthermore, the amount of weight placed on internal attitudes and external pressures affects the behavioural intent. For example, when a large reward is offered to a child who values money, i.e., $A_{rew}[k] \gg 0$, and who also has a large negative internal attitude, i.e., $A_{out}[k] \ll 0$, the behavioural intent that is formed may or may not be positive. If the internal attitude carries more weight than the reward attitude, then the behavioural intent will be negative. Likewise, if the reward attitude carries more weight than the internal attitude, then the behavioural intent will be positive. This weighting of attitudes and external pressures is known to vary by individual and situation. Thus, for our model, this weighting is included in the parameter values (see Table 2.2 for a summary of key parameters and signals that appear in our model).

2.2 Psychology Describing How Behaviour Drives Attitudes

The second step to understanding the attitude-behaviour relationship is studying ways in which behaviour can influence and change attitudes. These influences are characterized in our model's second component, thus closing the attitude-behaviour feedback loop. Several well-known, experimentally verified, psychological results support the seemingly counter-intuitive effect that behaviours influence attitudes, including *cognitive dissonance theory* developed by Festinger [9], and the *overjustification effect*, proposed by Deci [7]. These two theories are selected to form the basis of our model's "behaviour-driving-attitude" component. Cognitive dissonance theory is selected due to its wide application range and the overjustification effect is used due to its relevance to how the one-person system is controlled: rewards.

2.2.1 Cognitive Dissonance Theory: Overview

Cognitive dissonance theory is one of social psychology's most widely known and established theories. First proposed by Leon Festinger in 1957, cognitive dissonance has been experimentally verified in thousands of studies across dozens of settings, thus resulting in many different paradigms [4], [9]. Among other things, the theory contains a non-intuitive result: behaviour can influence attitudes. Before discussing cognitive dissonance theory in the context of our model and the piano example, a general understanding of the theory's concepts is required. Cognitive dissonance theory has two important concepts: *dissonance pressure* and *dissonance pressure reduction*.

Dissonance pressure is an uncomfortable psychological feeling that arises when a person holds two opposing *cognitions* (where a cognition is defined as knowledge of a thought, feeling, attitude or behaviour). For example, suppose a person who smokes cigarettes on a daily basis knows the harmful side-effects of his behaviour but has several friends who also smoke. Since this person's behaviour (choosing to smoke) is inconsistent with his knowledge that smoking his bad for his health, he experiences dissonance pressure. However, the knowledge that he has several friends who also smoke is consistent with his smoking behaviour. The smoker compares his behaviour cognition (termed the *generative cognition* by Beauvois and Joules [4]) with each of his other cognitions related to the act of smoking, thus forming *cognitive pairs*. If the behaviour cognition follows naturally from the other element in the cognitive pair, then the cognitive pair is said to be *consistent*. On the other hand, an *inconsistent* cognitive pair is one in which the two cognitions contradict each other. Inconsistent cognitive pairs lead to dissonance pressure.

The extent to which a cognitive pair is consistent (or inconsistent) is given a magnitude, denoted M_{con} (or M_{incon}). The consistency (or inconsistency) magnitude indicates how strongly the cognitive pair is consistent (or inconsistent). Festinger argues that the more strongly a cognition is held, the greater will be the consistency (or inconsistency) magnitude. The exact details of this calculations are given in Section 2.3.2. At this time, it is only important to understand that each cognitive pair has an associated consistency or inconsistency magnitude (whatever the case may be) that is proportional to the strength of each element in the cognitive pair.

Since dissonance pressure occurs when an inconsistent cognitive pair exists and such a pair has an associated magnitude, it follows that dissonance pressure also has a magnitude, i.e., a person can experience different amounts of dissonance pressure. Indeed, Festinger provides a description of dissonance pressure magnitude (P) which Beauvois and Joule transform into the following mathematical equation:

$$P = \frac{\sum M_{incon}}{\sum M_{incon} + \sum M_{con}}. \tag{2.1}$$

Since, Festinger argues, dissonance pressure is an uncomfortable feeling, people want to reduce or eliminate it. This motivates the second concept of cognitive dissonance theory: *dissonance pressure reduction*. Festinger suggests three dissonance reduction techniques, outlined below:

(i) Decreasing (or eliminating) the inconsistency magnitude sum, $\sum M_{incon}$. For the smoking example, the smoker could, for instance, decrease the inconsistency magnitude by believing that smoking is not as harmful as he once thought.

(ii) Increasing the consistency magnitude sum, $\sum M_{con}$. For the smoking example, the smoker could accomplish this by changing the strength of the consistent cognition, i.e., he could believe a greater proportion of his friends smoke, or he could acquire new smoking friends to increase the fraction who smoke.

(iii) Introducing a new, consistent cognition, thus increasing $\sum M_{con}$. For the smoking example, the smoker could, for instance, begin to convince himself that smoking has more, and stronger, benefits (stress reduction, weight loss, etc.) than it really does.

These three techniques are consistent with how the magnitude given in (2.1) could be decreased. Finally, the more dissonance pressure a person experiences, the more the person will tend to perform these three dissonance reduction techniques.

2.2.2 Cognitive Dissonance Theory: Induced Compliance Paradigm

Cognitive dissonance theory is now discussed within the context of our model and the piano example. The setup of the one-person problem, given in Chapter 1, indicates that the child initially has a negative attitude and behaviour, i.e., $A_{out}[0] < 0$ and $B[0] < 0$. The child's parents offer him a reward at some time $k \geq 0$, thus producing a positive reward attitude, $A_{rew}[k]$ for $k > 0$. Thus, two cognitive pairs are formed: $(A_{out}[k], B[k])$ and $(A_{rew}[k], B[k])$. This setup follows the classic *induced compliance paradigm* of cognitive dissonance theory. In the induced compliance paradigm, a person is offered a reward to do something he/she does not want to do. Moreover, the person does not experience any other external influences. Classic induced compliance studies have shown that subjects experience dissonance pressure because one of their two cognitive pairs is inconsistent. Furthermore, in these studies, subjects were shown to reduce dissonance pressure by changing their internal attitude [10]. Thus, for our purposes, the following assumption is made:

Assumption 2.4. *Dissonance pressure reduction occurs through change in $A_{out}[k]$.*

Note, for future work, our model could be extended to allow for other dissonance reduction methods.

The induced compliance paradigm considers two different cases, one in which the reward is large enough to induce a positive behaviour and the one in which the reward is insufficient. In each of these cases, dissonance pressure arises. First, the factors contributing to dissonance pressure are discussed, followed by an investigation into the relationship between how the amount of dissonance pressure varies with respect to the reward value. Then, for the two cases, dissonance pressure reduction methods are presented. Let Case A be the situation in which a sufficiently large reward is offered, thus inducing a positive behaviour. Let Case B be the situation in which the reward is insufficient at producing a positive behaviour.

In Case A, the sufficiently large reward produces a positive behaviour and therefore, the reward attitude and the behaviour are consistent. On the other hand, the positive behaviour and the negative internal attitude are inconsistent and hence, dissonance pressure arises. When a person is given a reward that is just large enough to produce a positive behaviour, some amount of dissonance pressure occurs. If, instead, a larger reward is offered, then less dissonance pressure occurs because the person essentially rationalizes their behaviour as being due to the reward (instead of their internal attitude). Therefore, in Case A, larger rewards produce less dissonance pressure.

In Case B, the reward is insufficient and thus, produces a negative behaviour, which forms an inconsistency between the reward attitude and the behaviour. Due to this inconsistency, dissonance pressure arises. Note, however, that in this case, the negative internal attitude is consistent with the negative behavioural intent. Like Case A, the amount of dissonance pressure a person experiences varies with respect to the reward value. When a person is offered a very small reward, a small amount of dissonance pressure occurs because he/she is rejecting a small reward. However, the dissonance pressure is larger in the case

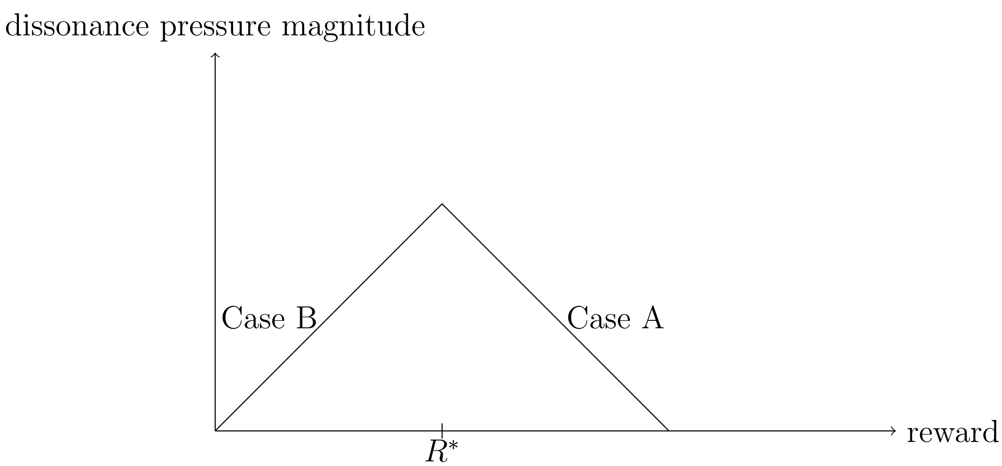

Figure 2.2: Short-term relationship between dissonance pressure magnitude and amount of reward when initial attitude is negative [9]. R^* is the smallest reward value that is sufficiently large to induce a positive behaviour. As the reward value varies from zero to R^*, dissonance pressure increases. As the reward value varies from R^* to infinity, dissonance pressure decreases. This figure looks only at the short-term relationship between dissonance pressure magnitude and reward, as the dissonance pressure magnitude can change over time.

of larger (but still insufficient) rewards. This increased dissonance pressure is due to the person rejecting a larger reward, which is a more uncomfortable experience. Therefore, in Case B, larger rewards produce more dissonance pressure.

In Figure 2.2, the relationship describing how dissonance pressure varies with respect to the reward is shown in graphical form. In this thesis, this relationship is termed the *dissonance triangle*. Festinger notes that the dissonance triangle does not represent the exact mathematical equation for calculating dissonance pressure. Instead, Festinger argues that the exact nature of the relationship has yet to be determined. As will be shown in Chapter 3, our model is consistent with Figure 2.2.

Now, dissonance pressure reduction is considered for the two situations. In Case A, the internal attitude is negative and the behaviour is positive, i.e., for some sample k, $A_{out}[k] < 0$ and $B[k] \geq 0$. Clearly, $(A_{out}[k], B[k])$ form the inconsistent cognitive pair. By Assumption 2.4, dissonance pressure is reduced by increasing or decreasing $A_{out}[k]$. Since $A_{out}[k]$ contributes to the inconsistency magnitude and from before, dissonance pressure can be reduced by decreasing $\sum M_{incon}$, the inconsistency magnitude must decrease. There are two ways to achieve this:

1. The internal attitude can become less negative (but still remain negative); thus,

reducing the strength of $A_{out}[k]$ and consequently, $\sum M_{incon}$.

2. The internal attitude can become positive; thus, eliminating $\sum M_{incon}$ and introducing a new, consistent cognition, which increases $\sum M_{con}$. In this situation, two of the three techniques outlined in Section 2.2.1 are employed simultaneously.

In both cases, $A_{out}[k]$ basically increases to reduce dissonance pressure.

On the other hand, in Case B, the behaviour is negative and therefore, $(A_{out}[k], B[k])$ form the consistent cognitive pair. By Assumption 2.4, dissonance is reduced by increasing or decreasing $A_{out}[k]$. Since $A_{out}[k]$ contributes to the consistency magnitude and from before, dissonance pressure can be reduced by increasing $\sum M_{con}$, the consistency magnitude must increase. For our model, this increase can only occur by the internal attitude becoming more negative; hence, for Case B, $A_{out}[k]$ decreases to reduce dissonance pressure.

To summarize, in the induced compliance paradigm, dissonance pressure arises when a reward is offered to induce a "counter-attitudinal" behaviour. The amount of dissonance pressure varies with respect to the reward value, as shown in Figure 2.2. When the reward is sufficiently large, dissonance pressure reduces through the internal attitude becoming less negative, i.e., $A_{out}[k]$ increases; whereas in the case of an insufficient reward, dissonance pressure reduces through the internal attitude becoming more negative, i.e., $A_{out}[k]$ decreases.

2.2.3 The Overjustification Effect

Cognitive dissonance theory suggests how behaviour can affect attitudes when there are inconsistent cognitions. If all attitudes are consistent with behaviour, then cognitive dissonance theory concludes no dissonance pressure arises and thus predicts no attitude change occurs. However, researchers including Edward Deci and colleagues have shown that attitude can still change when all cognitions are consistent [8]. Specifically, when a person has a positive internal attitude towards a behaviour and is offered a reward in a controlling manner, then the person's internal attitude will decrease. Hence, it is counter-productive to offer a person a reward if his attitude is already positive. This is called the *overjustification effect*.

The overjustification effect can be explained through *attribution theory*: the person attributes their performance of the behaviour to the reward instead of their internal attitude, i.e., the person thinks "I must be performing the behaviour because I'm getting a reward and therefore, I am less motivated to perform this behaviour without a reward." In a study on the overjustification effect, [7], subjects who were were interested in an activity were given a reward to continue performing the activity. Upon removal of the reward, the subjects still continued the activity, but to a much lesser extent. Thus, from the theory of planned behaviour, the internal attitude must have decreased to account for the difference between pre-reward and post-reward behaviour strength. Important to note is the fact that even though the internal attitude decreases, it does not stop the behaviour altogether, thus, the attitude does not become negative. Furthermore, the initial strength of the attitude and the reward amount were both factors in the amount of attitude change.

The dynamic attitude-behaviour model proposed in this thesis includes the overjustification effect due to the possibility in the induced compliance paradigm that the child's internal attitude may become positive (see Case A of the induced compliance paradigm in Section 2.2.2). In this case, the child's behaviour, internal attitude and reward attitude are all positive, so *overjustification pressure* arises, causing attitude to decrease.

2.3 A Dynamic Attitude-Behaviour Model

The three social psychological theories introduced in Sections 2.1 and 2.2 are now combined to form a discrete-time dynamic attitude-behaviour model. Even though it seems natural to consider human attitude and behaviour as a varying continuously with time, a discrete-time model is developed for two reasons. First, attitude measurement cannot be done instantaneously. However, the area of research pertaining to attitude measurement, *psychometrics*, provides methods to sample an individual's attitude. Thus, it is natural to follow this discrete-time approach. Second, a continuous-time reward signal makes little sense from a practical perspective. When a reward is offered, it is offered at a certain point in time and thus, the system input is a discrete-time signal.

The second modelling decision that has been made involves how dynamics are incorporated into our model. Two factors contribute to the system's dynamics: feedback and mental processing. The first factor, as previously discussed, incorporates feedback into the attitude-behaviour model to account for the feedback nature of the attitude-behaviour relationship. The second factor accounts for the mental processing that occurs when a new attitude is formed or pressure is experienced. Researchers in the field of cognitive psychology have concluded that humans take time to mentally process information; thus, following standard research on mental processing dynamics [17], first-order lag mental processing dynamics are assumed throughout our model.

Figure 2.3 provides a high-level diagram of our model. Component A essentially models the theory of planned behaviour, while Component B models how cognitive dissonance theory and the overjustification effect result in changes to A_{out}. The overall change, as indicated in the figure, is denoted $\Delta A_{out}[k]$. Component A and Component B are examined more carefully below. See Table 2.2 for a summary of key parameters and signals.

Symbol	**Description [*value or units*]**
$A_{out}[k] \in (-\infty, \infty)$	attitude towards the behaviour outcome [*attitude units*]
$\Delta A_{out}[k] \in (-\infty, \infty)$	change in attitude towards the behaviour due to cognitive dissonance and overjustification effects [*attitude units*]
$A_{rew}[k] \in [0, \infty)$	attitude towards accepting a reward to perform the behaviour [*attitude units*]
$R[k] \in [0, \infty)$	reward value [*dollars*]
$B[k] \in (-\infty, \infty)$	behaviour [*attitude units*]
$P^{CD}_{raw}[k], P^{CD}[k] \in (-0.5, 0.5)$	pressure arising from cognitive dissonance effects [*unitless*]
$P^{OJ}_{raw}[k], P^{OJ}[k] \in [0, \infty)$	pressure arising from the overjustification effect [(*attitude units*)^2]
$K_1 \in [0, \infty)\dagger$	gain reflecting how the dissonance pressure affects attitude change [$K_1 = 30$ *attitude units*]
$K_2 \in [0, \infty)\dagger$	gain reflecting how the overjustification pressure affects attitude change [$K_2 = 0.1$ *1/(attitude unit)*]
$\mu_1 \in (0, \infty)$	value assigned to one dollar [1 *attitude unit per dollar*]
$r_1 \in [0, 1)\dagger$	mental processing pole location for reward attitude formation in (2.3) [$r_1 = 0$]
$r_2 \in [0, 1)\dagger$	mental processing pole location for dissonance pressure in (2.11) [$r_2 = 0.5$]
$r_3 \in [0, 1)\dagger$	mental processing pole location for overjustification effect in (2.14) [$r_3 = 0.5$]

Table 2.2: Key signals and parameters that appear in the model. The parameter values listed are those used for analysis in Chapter 3. Parameters marked † depend on sampling period, taken here to be one day.

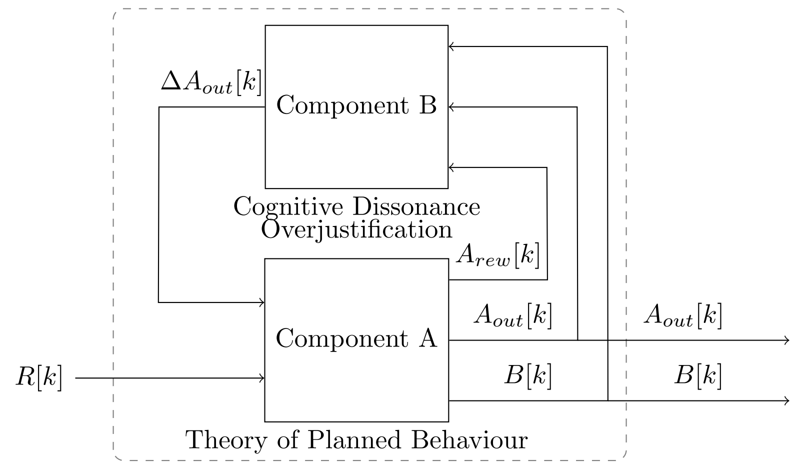

Figure 2.3: Block diagram of the one-person system (dotted box), decomposed into Components A and B. The system input is $R[k]$, and the system outputs are $A_{out}[k]$ and $B[k]$.

2.3.1 Details of Component A

Component A is based on the theory of planned behaviour, discussed in Section 2.1.2, which states that an individual's behaviour is formed from his/her internal attitude and external influences (assuming full perceived behavioural control). The internal attitude depends on how much attitude change is formed through dissonance and overjustification pressures. The reward attitude depends on the reward in dollars ($R[k]$), how much the reward is valued (μ_1), and a first-order mental processing model (with r_1 denoting the pole location). From the theory of planned behaviour, stronger internal attitudes and external influences cause stronger behavioural intent. Moreover, a weighting factor influences the extent to which each of these elements affects $B[k]$. For our model, this weighting is contained in μ_1. Thus, Component A is modelled by the following three equations:

$$A_{out}[k] = A_{out}[k-1] + \Delta A_{out}[k-1] \tag{2.2}$$

$$A_{rew}[k] = r_1 A_{rew}[k-1] + \mu_1 (1 - r_1) R[k-1] \tag{2.3}$$

$$B[k] = A_{out}[k] + A_{rew}[k]. \tag{2.4}$$

2.3.2 Details of Component B

Component B, shown in more detail in Figure 2.4, is based on cognitive dissonance theory and the overjustification effect. These two psychological effects create pressure, producing attitude change. Thus, the output of Component B is $\Delta A_{out}[k]$, the change in attitude

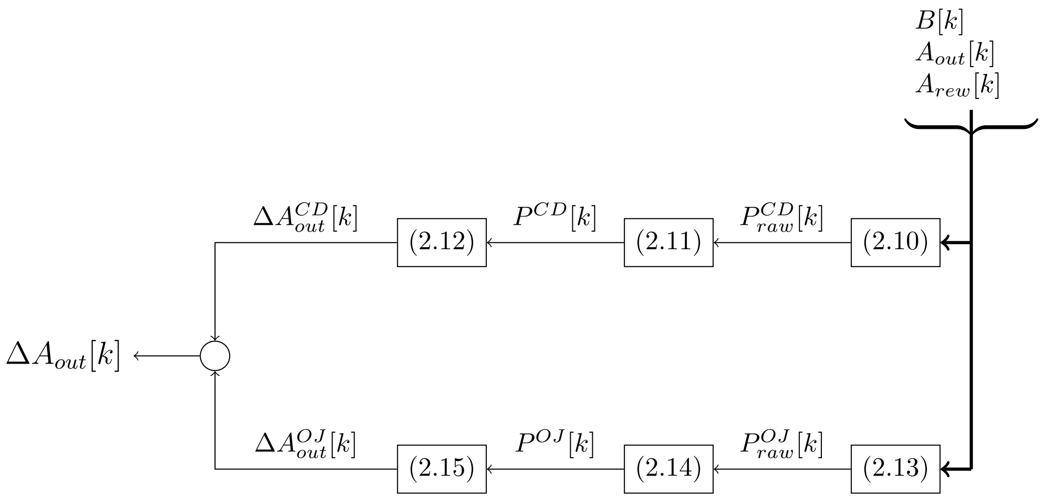

Figure 2.4: Details of Component B. The thick line is used to indicate multiple signals.

towards the behavioural outcome:

$$\Delta A_{out}[k] = \Delta A_{out}^{CD}[k] + \Delta A_{out}^{OJ}[k]. \tag{2.5}$$

Each of the terms in (2.5) will now be examined.

Cognitive Dissonance Theory Model

As discussed in Section 2.2.2, the effects due to the induced compliance paradigm of cognitive dissonance theory arise when $A_{out}[k] < 0$ and $A_{rew}[k] > 0$. If these two inequalities hold, then each of these cognitions form a cognitive pair with the generative cognition, $B[k]$, one of which is inconsistent, thus causing dissonance pressure and attitude change.

The dissonance pressure magnitude equation in (2.1) depends on $\sum M_{con}$ and $\sum M_{incon}$. For the one-person system, there are two cognitive pairs, $(A_{out}[k], B[k])$ and $(A_{rew}[k], B[k])$, each of which have an associated $M_{con}[k]$ and $M_{incon}[k]$. Festinger states that these magnitudes are proportional to the strength of each element in the cognitive pair; thus, define

$$M_{incon}^{1}[k] = \begin{cases} |A_{rew}[k]B[k]| & \text{if } A_{rew}[k] > 0, B[k] < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.6}$$

$$M_{incon}^{2}[k] = \begin{cases} |A_{out}[k]B[k]| & \text{if } A_{out}[k] < 0, B[k] > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.7}$$

and

$$M_{con}^1[k] = \begin{cases} |A_{rew}[k]B[k]| & \text{if } A_{rew}[k] > 0, B[k] > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{2.8}$$

$$M_{con}^2[k] = \begin{cases} |A_{out}[k]B[k]| & \text{if } A_{out}[k] < 0, B[k] < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.9}$$

Let $M_{incon}[k] = \sum_{i=1}^{2} M_{incon}^i[k]$ and $M_{con}[k] = \sum_{i=1}^{2} M_{con}^i[k]$. Then, the (raw, unprocessed) dissonance pressure at time k is:

$$P_{raw}^{CD}[k] = \begin{cases} \operatorname{sgn}(B[k]) \frac{M_{incon}[k]}{M_{incon}[k]+M_{con}[k]} & \text{if } A_{out}[k] < 0, A_{rew}[k] > 0, \text{ and } B[k] \neq 0, \\ \frac{|A_{out}[k]|}{|A_{out}[k]|+|A_{rew}[k]|} & \text{if } A_{out}[k] < 0, A_{rew}[k] > 0, \text{ and } B[k] = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

The first case in (2.10) captures the cognitive dissonance effects in all situations except the special situation $B[k] = 0$, where a division by zero error would occur. The special situation $B[k] = 0$ is handled in the second case in (2.10). The "$\operatorname{sgn}(B[k])$" factor ensures that the sign of $\Delta A_{out}^{CD}[k]$ is consistent with cognitive dissonance theory in that it results in a decrease in dissonance pressure at the next time instant.

Having derived an expression for the raw dissonance pressure, the first-order mental processing model is applied as follows. Let $P^{CD}[k]$ represent the actual dissonance pressure experienced at time k. Then,

$$P^{CD}[k] = r_2 P^{CD}[k-1] + (1 - r_2) P_{raw}^{CD}[k]. \tag{2.11}$$

Finally, the dissonance pressure results in attitude change, denoted $\Delta A_{out}^{CD}[k]$, given by

$$\Delta A_{out}^{CD}[k] = K_1 P^{CD}[k]. \tag{2.12}$$

In (2.12), it is assumed that attitude change is proportional to the dissonance pressure; psychologists have yet to identify the exact relationship.

The Overjustification Effect Model

As discussed in Section 2.2.3, the overjustification effect applies only when $A_{rew}[k] > 0$, $A_{out}[k] > 0$ and $B[k] > 0$. The basic pressure arising from the overjustification effect is modelled as

$$P_{raw}^{OJ}[k] = \begin{cases} A_{out}[k]A_{rew}[k] & \text{if } A_{out}[k] > 0, B[k] > 0, \text{ and } A_{rew}[k] > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{2.13}$$

In (2.13), it is assumed that the pressure depends on the product of the relevant attitudes. The use of a product is somewhat arbitrary; psychologists have not yet determined the exact relationship, so any function that increases in magnitude with each of $A_{out}[k]$ and $A_{rew}[k]$ would be equally justified.

Assuming first-order mental processing of the pressure $P^{OJ}_{raw}[k]$, the following equation for the processed pressure is obtained:

$$P^{OJ}[k] = r_3 P^{OJ}[k-1] + (1 - r_3)\, P^{OJ}_{raw}[k]. \tag{2.14}$$

Finally, to avoid the situation where $P^{OJ}[k]$ is large enough to actually change the sign of $A_{out}[k]$ (an effect that is inconsistent with overjustification theory), $P^{OJ}[k]$ is saturated as follows:

$$\Delta A^{OJ}_{out}[k] = \begin{cases} -K_2 P^{OJ}[k] & \text{if } P^{OJ}[k] > 0, K_2 P^{OJ}[k] \leq A_{out}[k], \\ -A_{out}[k] & \text{if } P^{OJ}[k] > 0, K_2 P^{OJ}[k] > A_{out}[k], \\ 0 & \text{otherwise.} \end{cases} \tag{2.15}$$

As in (2.12), attitude change is assumed to be proportional to the experienced psychological pressure.

The terms $\Delta A^{CD}_{out}[k]$ and $\Delta A^{OJ}_{out}[k]$ handle the basic one-person configuration. The two-person system, discussed in Chapter 4, contains additional equations modelling the effect another significant person has on the attitude-behaviour model.

2.3.3 Simplifying Assumptions and Initial Conditions

Along with the problem statement given in Chapter 1, the psychology behind this discrete-time model suggests initial conditions. In particular, the child in the one-person system initially has a negative internal attitude. Furthermore, it can be reasonably assumed that he has not previously been offered a reward; thus, initially, his behavioural intent matches his negative attitude. A second reasonable assumption is that the child has not previously experienced any pressures due to dissonance or overjustification effects and thus, experiences no initial attitude change.

In the context of our model, these initial conditions are given by:

$$\begin{aligned} P^{CD}[0] = P^{OJ}[0] &= 0, \\ \Delta A^{CD}_{out}[0] = \Delta A^{OJ}_{out}[0] &= 0, \\ A_{rew}[0] &= 0, \\ B[0] = A_{out}[0] &= A_o, \end{aligned} \tag{2.16}$$

where $A_o < 0$.

Additionally, the following assumptions are made on the parameter values:

Assumption 2.5. *Gains reflecting how the dissonance and overjustification pressures affect attitude change are strictly positive, i.e.,* $K_1 > 0$ *and* $K_2 > 0$.

Assumption 2.6. *The value assigned to one dollar is strictly positive, i.e.,* $\mu_1 > 0$.

Assumption 2.7. *The mental processing pole location for reward attitude formation in (2.3) is zero, i.e.,* $r_1 = 0$.

Assumption 2.7 simplifies the expressions for $A_{rew}[k]$ in (2.3) to

$$A_{rew}[k] = \mu_1 R[k-1]. \tag{2.17}$$

Assumption 2.8. *The mental processing pole locations for dissonance and overjustification pressures in (2.11) and (2.14) respectively, are contained in the range* $[0, 1)$*, i.e.,* $0 \leq r_2, r_3 < 1$.

These assumptions are used throughout the remainder of this chapter, as well as in Chapter 3.

2.3.4 Interpreting the Discrete-Time Model

From a practical perspective, it is useful to understand how to interpret $A_{out}[k]$, $A_{rew}[k]$, $R[k]$, $B[k]$ and $\Delta A_{out}[k]$. Since a discrete-time approach is taken, it is natural to also discuss the timing sequence of these signals, summarized in Table 2.3. This discussion considers the motivating piano situation with a sample period of one day, using the values in Table 2.2.

On the morning of the first day ($k = 0$), the child's internal attitude towards playing the piano is negative (say, $A_o = -15$ *attitude units*) and since he has not previously been offered a reward, his reward attitude is zero (say, $A_{rew}[0] = 0$ *attitude units*). These two attitudes combine to form the child's intent to play the piano at some point during the first day ($B[0] = -15$ *attitude units*). Since the child does not intend to practise piano on day one, his parents offer him a reward at the end of the day (say, $R[0] = \$16$) to induce their child to play piano on the next day. From (2.3), the child needs time to mentally process this reward; thus it is clear this reward does not influence $B[0]$.

On the morning of the second day ($k = 1$), the child's attitude is the same (because no dissonance or overjustification pressures arose on the first day, i.e., $\Delta A_{out}[0] = 0$). However, the child has had sufficient time to mentally process the reward offered by his parents on the previous evening. This processed reward has formed an attitude towards the reward that was offered to play the piano at some point on day two ($A_{rew}[1] = 16$ *attitude units*). In this example, the internal attitude and reward attitude combine to form a positive

	Day 1 ($k = 0$)	**Day** 2 ($k = 1$)	**Day** 3 ($k = 2$)
$A_{out}[k]$	-15	-15	-12.58
$A_{rew}[k]$	0	16	16
$B[k]$	-15	1	3.419
$\Delta A_{out}[k]$	0	2.4194	3.4106
$R[k]$	16	16	16

Table 2.3: Timing chart portraying sequence of events at each time step. At the beginning of each day, the child wakes up with an internal attitude, $A_{out}[k]$, and an attitude towards accepting a reward and playing the piano, $A_{rew}[k]$. These two attitudes form a behavioural intent, $B[k]$, which translates into the child engaging (or not engaging) in the behaviour to some degree throughout the day. Once this behaviour is realized, the child may or may not experience pressure (due to effects from dissonance and/or overjustification). This pressure creates an attitude change, which must undergo mental processing before influencing $A_{out}[k]$ (see (2.2)). At the end of the day, the parents offer their child some reward, $R[k]$, to induce their child to practise the piano on the subsequent day.

behavioural intent, i.e., the child decides to accept the reward and practise piano on day two ($B[1] = 1$ *attitude units*). The amount of time and effort he spends practising the piano is related to the strength of his behavioural intent, $|B[k]|$. At the end of the day, the child has played the piano and thus, since $B[1] > 0$ and $A_{out}[1] < 0$, the child experiences dissonance pressure. However, due to mental processing, this dissonance pressure does not affect $A_{out}[1]$. Instead, the dissonance pressure causes the child to wake up on the morning of the third day with a new attitude (because $A_{out}[2] = A_{out}[1] + \Delta A_{out}[1]$ and $\Delta A_{out}[1] > 0$ due to dissonance pressure). Finally, suppose the parents choose to offer their son the same reward on the evening of day two to induce him to play the piano on day three.

The same sequence of events occurs on the third day. However, where on the morning of day two the internal attitude remained the same, on the morning of day three, the child wakes up with a changed attitude, due to the dissonance pressure experienced on the previous day. From (2.12), the dissonance pressure experienced on day two produces some amount of attitude change. This attitude change undergoes mental-processing by the child (see (2.2)) and thus influences attitude such that on the beginning of day three, the child has a new attitude. This sequence of events continues for each subsequent day.

2.4 Model Consistency Verification

Typically, after a system is modelled, model validation is performed. However, for our model, the amount of time required for adequate validation is quite large. Additionally, further expertise and resources are needed to perform the psychology experiments necessary for formal model verification. Finally, it is unclear if experts in the field of *psychometrics*, which among other things, studies the measurement of attitudes and behaviours, are able to reliably measure all of the key signals in our model, such as dissonance pressure. For these reasons, a *model consistency verification* is conducted instead of formal model validation. For this verification, the consistency between the one-person system, developed in Section 2.3, and the relevant psychological theories, introduced in Sections 2.1 and 2.2, is studied. This model consistency verification is accomplished by comparing key qualitative characteristics and trends of the relevant psychology theories (the theory of planned behaviour, cognitive dissonance theory and the overjustification effect) with the qualitative characteristics and trends of signals within our model.

The theory of planned behaviour, which describes how attitudes combine to generate a behavioural intent, has two qualitative trends:

1. When all attitudes are positive (or negative), behavioural intent is positive (or negative).
2. An increase (or decrease) in any one attitude causes an increase (or decrease) in behavioural intent.

By inspection of the behavioural intent equation given in (2.4), it is clear that our model is consistent with these two characteristics.

The induced compliance paradigm of cognitive dissonance theory suggests, among other things, how behaviour can influence attitude through dissonance pressure and exhibits two distinct qualitative trends:

1. From general cognitive dissonance theory, people tend to reduce the dissonance pressure magnitude, should it arise.
2. From the induced compliance paradigm, the relationship between the reward and the dissonance pressure magnitude is similar to the dissonance triangle, shown in Figure 2.2.

In addition to these two qualitative trends, Festinger proposed a quantitative feature of cognitive dissonance theory. In [9], Festinger states "that the weighted proportion of [inconsistent cognitions] cannot be greater than 50 per cent." This suggests a maximum dissonance pressure of 50 per cent, or 0.5; hence, verification is performed to determine whether or not our model is consistent with this conclusion. Now our model is shown to exhibit these three characteristics.

To verify if the one-person system exhibits the first trend, recall that by Assumption 2.4, dissonance reduction occurs through attitude change. Now, consider the modelling equations associated with cognitive dissonance pressure and the resulting attitude change, i.e., (2.6)–(2.12). Verifying that dissonance pressure is reduced through attitude change amounts to showing that when $P^{CD}_{raw}[k] \neq 0$, following relationship holds:

$$|P^{CD}_{raw}[k+1]| < |P^{CD}_{raw}[k]|$$

when $A_{out}[k+1] \neq A_{out}[k]$. The raw, unprocessed dissonance pressure is considered here because the mental processing dynamics make it more challenging to compare the amount of dissonance pressure arising at each sample. If the above relationship is shown to be true, then, since the mental processing dynamics are simply modelled by a first-order lag transfer function, the processed dissonance pressure magnitude also decreases eventually. Furthermore, the magnitude is considered because it represents the amount or strength of the dissonance pressure.

From (2.10), the internal attitude must be negative and the reward attitude must be positive for raw, unprocessed dissonance pressure to arise. To understand why, consider a non-negative internal attitude. This internal attitude, combined with a positive reward, generates a positive behaviour and therefore, both cognitive pairs are consistent, i.e., dissonance pressure does not arise. Suppose instead, that the internal attitude is negative but the reward attitude is zero; then, the behavioural intent is the same at the internal attitude and again, both cognitive pairs are consistent. Therefore, in both the case when $A_{out}[k] \geq 0$ and when $A_{rew}[k] = 0$, the raw, unprocessed dissonance pressure is zero.

Supposing $A_{out}[k] < 0$ and $A_{rew}[k] > 0$, two cases must be considered. These two cases match Case A and Case B from the induced compliance paradigm discussed in Section 2.2.2, i.e., the first case considers when $B[k] \geq 0$ and the second case considers when $B[k] < 0$. For the situation in which $B[k] \geq 0$, from (2.6)–(2.9), $M_{incon}[k] = |A_{out}[k]B[k]|$ and $M_{con}[k] = |A_{rew}[k]B[k]|$. Assuming $B[k] \neq 0$, from (2.10),

$$P^{CD}_{raw}[k] = \frac{|A_{out}[k]B[k]|}{|A_{out}[k]B[k]| + |A_{rew}[k]B[k]|} = \frac{|A_{out}[k]|}{|A_{out}[k]| + |A_{rew}[k]|}. \tag{2.18}$$

Note that if $B[k] = 0$, then $P^{CD}_{raw}[k] = \frac{|A_{out}[k]|}{|A_{out}[k]|+|A_{rew}[k]|}$, which is the same as the above expression. From (2.11), this raw, unprocessed dissonance pressure produces a positive $P^{CD}[k]$ and thus from (2.12) a positive attitude change, i.e.,

$$A_{out}[k+1] = A_{out}[k] + K_1 P^{CD}[k] > A_{out}[k].$$

Note that $\Delta A_{out}^{OJ}[k] = 0$ since the conditions required for overjustification effect to hold are not met. The above expression is consistent with Case A of the induced compliance paradigm. A sufficiently large reward that is able to produce a positive behavioural intent has the effect of causing a negative attitude to increase. To see how this affects the raw, unprocessed dissonance pressure at sample $k + 1$, assume that this increased attitude, $A_{out}[k + 1]$ is still negative (otherwise, $P_{raw}^{CD}[k + 1] = 0$ and thus, dissonance pressure is reduced). Furthermore, since the reward attitude also contributes to the magnitude of $P_{raw}^{CD}[k]$, assume the reward attitude is constant to ensure only the effects of changes to the internal attitude are studied. Since the internal attitude increases and the reward remains constant, $B[k]$ necessarily increases and is therefore, still positive. Hence, from (2.10),

$$\begin{aligned} P_{raw}^{CD}[k+1] &= \frac{|A_{out}[k+1]B[k+1]|}{|A_{out}[k+1]B[k+1]| + |A_{rew}[k+1]B[k+1]|} \\ &= \frac{|A_{out}[k+1]|}{|A_{out}[k+1]| + |A_{rew}[k+1]|}. \end{aligned}$$

Since $A_{rew}[k + 1] = A_{rew}[k]$ and $|A_{out}[k + 1]| < |A_{out}[k]|$ (because $A_{out}[k] < A_{out}[k + 1]$ and both are negative), it follows that $P_{raw}^{CD}[k + 1] < P_{raw}^{CD}[k]$ as required. Note that since $P_{raw}^{CD}[k] > 0$, the magnitude signs are omitted.

For the situation in which $B[k] < 0$, from (2.6)–(2.9), $M_{con}[k] = |A_{out}[k]B[k]|$ and $M_{incon}[k] = |A_{rew}[k]B[k]|$. From (2.10),

$$P_{raw}^{CD}[k] = -\frac{|A_{rew}[k]B[k]|}{|A_{rew}[k]B[k]| + |A_{out}[k]B[k]|} = -\frac{|A_{rew}[k]|}{|A_{rew}[k]| + |A_{out}[k]|}. \tag{2.19}$$

From (2.11), this raw, unprocessed dissonance pressure produces a negative $P^{CD}[k]$ and thus from (2.12) a negative attitude change, i.e.,

$$A_{out}[k+1] = A_{out}[k] + K_1 P^{CD}[k] < A_{out}[k].$$

Similar to before, $\Delta A_{out}^{OJ}[k] = 0$. The above expression is consistent with Case B of the induced compliance paradigm. An insufficient reward that is not able to produce a positive behavioural intent has the effect of causing a negative attitude to decrease. This decrease effectively decreases $B[k]$ (again, assuming a constant reward attitude). Hence, from (2.10),

$$\begin{aligned} P_{raw}^{CD}[k+1] &= -\frac{|A_{rew}[k+1]B[k+1]|}{|A_{rew}[k+1]B[k+1]| + |A_{out}[k+1]B[k+1]|} \\ &= -\frac{|A_{rew}[k+1]|}{|A_{rew}[k+1]| + |A_{out}[k+1]|}. \end{aligned}$$

Since $A_{rew}[k+1] = A_{rew}[k]$ and $|A_{out}[k+1]| > |A_{out}[k]|$ (because $A_{out}[k] > A_{out}[k+1]$ and both are negative), it follows that $|P_{raw}^{CD}[k+1]| < |P_{raw}^{CD}[k]|$ as required.

For illustrative purposes, two simulations are performed, shown in Figure 2.5. In both simulations, a step reward is applied to the one-person system with the initial conditions given in (2.16) and the parameter values given in Table 2.2. The first simulation, shown in the left-hand side plots, is related to Case A of the induced compliance paradigm, whereas the right-hand side plots are the results of the second simulation, related to Case B. As both simulations show, when dissonance pressure arises, the raw, unprocessed dissonance pressure decreases through attitude change. Furthermore, the processed dissonance pressure, $P^{CD}[k]$ also decreases, as expected. Therefore, our model is consistent with the first qualitative characteristic of the induced compliance paradigm of cognitive dissonance theory.

Next, the second characteristic related to cognitive dissonance on page 26, which describes the relationship between the reward and the dissonance pressure magnitude, is considered. Again, the unprocessed dissonance pressure is studied due to the additional complexity arising from the mental processing dynamics. From Section 2.2.2, the reward and the dissonance pressure magnitude should exhibit the dissonance triangle relationship, given in Figure 2.2. The dissonance triangle indicates that the larger of two insufficient rewards produces more dissonance pressure than the smaller of these two rewards, whereas the larger of two sufficiently large rewards produces less dissonance pressure. To check if our model is consistent with this conclusion, the expression for the raw, unprocessed dissonance pressure is examined. From before, there are two possible expression for $P^{CD}_{raw}[k]$: (2.18) and (2.19). To determine how the magnitude of these two expressions varies with respect to $A_{rew}[k]$, fix $A_{out}[k]$ to some constant value. In Case A, when $B[k] \geq 0$, $P^{CD}_{raw}[k]$ is given by (2.18). For some fixed $A_{out}[k]$, increasing $A_{rew}[k]$ produces a smaller $P^{CD}_{raw}[k]$ (since $A_{rew}[k]$ is only in the denominator term of the expression). Therefore, for sufficiently large rewards, greater rewards produce smaller dissonance pressure. On the other hand, in Case B, when $B[k] < 0$, $P^{CD}_{raw}[k]$ is given by (2.19). This expression can be rearranged to

$$|P^{CD}_{raw}[k]| = \frac{1}{1 + \frac{|A_{out}[k]|}{A_{rew}[k]}}.$$

Again, the magnitude is considered because the sign simply indicates the attitude change direction. From the above expression it is evident that for a fixed $A_{out}[k]$, larger values of $A_{rew}[k]$ produce larger values of $|P^{CD}_{raw}[k]|$. Therefore, for insufficient rewards, greater rewards produce greater dissonance pressure. The relationship between the reward attitude and the dissonance pressure is illustrated in Figure 2.6. In the figure, the expressions for $P^{CD}_{raw}[k]$ are plotted with respect to $A_{rew}[k]$. For insufficient rewards, the expression given in (2.19) is used, whereas (2.18) is used for sufficiently large rewards. Figure 2.6 clearly demonstrates that our model exhibits the same dissonance triangle relationship predicted by Festinger.

In addition to the two qualitative trends of cognitive dissonance theory, the final characteristic, which is quantitative in nature, is considered by analysing the maximum dissonance

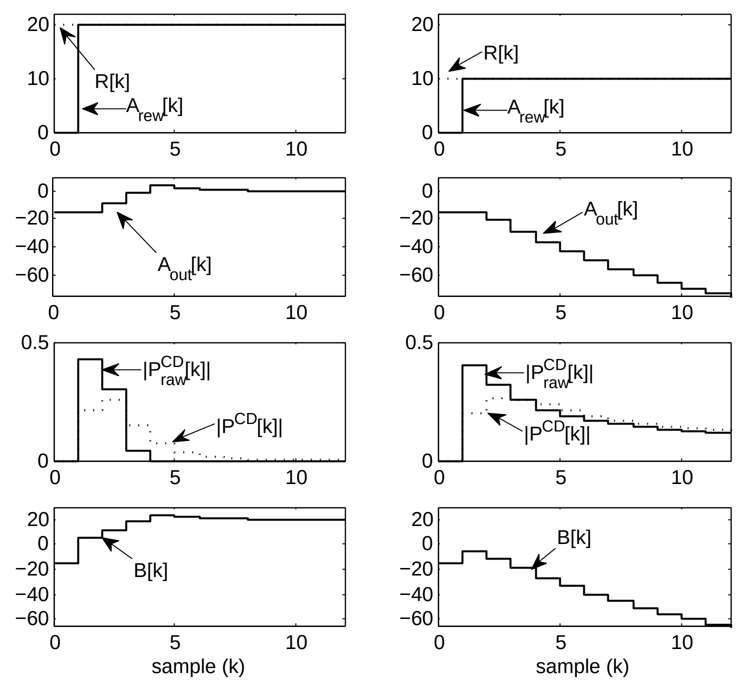

Figure 2.5: Simulation demonstrating how the one-person system with the initial conditions given in (2.16) responds to a step reward. Set $A_o = -15$ *attitude units* and consider two step reward magnitudes: a reward that is sufficiently large enough to produce positive behavioural intent ($R[k] = \$20$, left plots), and a reward that is not sufficiently large ($R[k] = \$10$, right plots). For a sufficiently large reward, the magnitude of $A_{out}[k]$ first decreases and, as a result, $A_{out}[k]$ becomes less negative. This has the effect of reducing $P^{CD}_{raw}[k]$ and $P^{CD}[k]$. Note though, that once attitude becomes positive, overjustification pressure arises and affects $A_{out}[k]$. For an insufficient reward, $A_{out}[k]$ decreases, thus increasing its magnitude. This has the effect of reducing $P^{CD}_{raw}[k]$ and $P^{CD}[k]$. $B[k]$, $R[k]$ and $A_{rew}[k]$ are also included in the above plots.

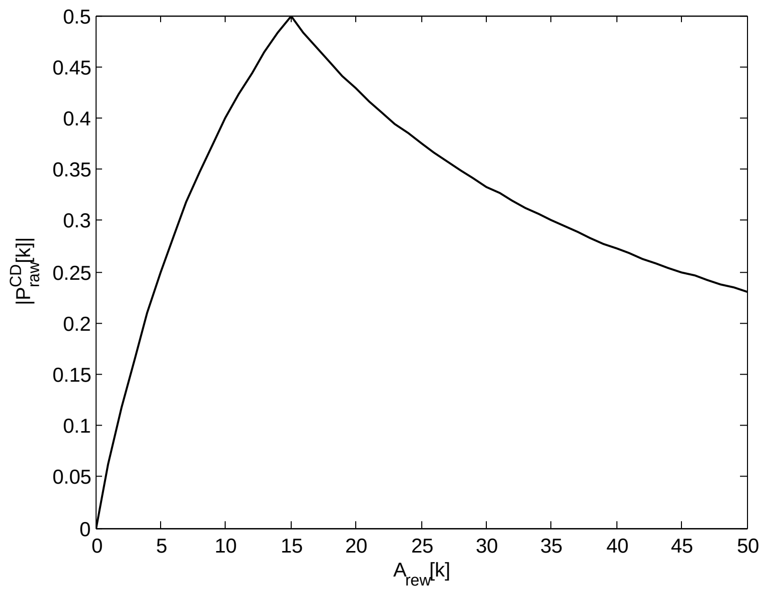

Figure 2.6: Relationship between raw, unprocessed dissonance pressure magnitude and reward attitude for the one-person system. Set $A_{out}[k] = -15$ *attitude units*; then, the raw, unprocessed dissonance pressure magnitude, $|P^{CD}_{raw}[k]|$, is calculated using (2.19) when $A_{rew}[k] \in [0 \textit{ attitude units}, 15 \textit{ attitude units})$ and (2.18) when $A_{rew}[k] \in [15 \textit{ attitude units}, 50 \textit{ attitude units}]$. The left-hand side of the plot is related to Case B, whereas the right-hand side of the plot is related to Case A.

pressure magnitude. Festinger's proposition that the weighted amount of inconsistent cognitive pairs must be less than 50 per cent suggests that the raw, unprocessed dissonance pressure has a maximum value. The raw, unprocessed dissonance pressure, $P^{CD}_{raw}[k]$, is a fraction; consequently, the maximum dissonance pressure magnitude should be $\frac{1}{2}$ or 0.5. Lemma 2.1 below confirms this proposition. Determining the maximum dissonance pressure magnitude is not only important for verifying whether or not our model is consistent with the psychology, it is also useful for controller design. For simplicity, the maximum dissonance pressure analysis only considers the one-person system with the initial conditions given in (2.16).

Lemma 2.1. *For the one-person system with the initial conditions given in (2.16), the maximum raw, unprocessed dissonance pressure magnitude ($|P^{CD}_{raw}[k]|$) is* 0.5.

Proof. See Appendix C.1. □

In conclusion, our model exhibits all three characteristics of cognitive dissonance.

The final component of our model, the overjustification effect, is now shown to be qualitatively consistent with the psychology introduced in Section 2.2.3. The overjustification effect has has two qualitative trends describing attitude change direction and magnitude:

1. When all cognitions are positive, i.e., $B[k] > 0$, $A_{out}[k] > 0$ and $A_{rew}[k] > 0$, the attitude towards the behavioural outcome decreases but does not change sign.

2. The attitude change magnitude increases with an increase in reward and/or attitude towards behavioural outcome.

To show that our model is consistent with the first trend, consider the overjustification modelling equations: (2.13), (2.14) and (2.15). From (2.13), if $B[k] > 0$, $A_{out}[k] > 0$ and $A_{rew}[k] > 0$, then $P^{OJ}_{raw}[k] > 0$. Note that $P^{OJ}_{raw}[k]$ can never be negative. Since $P^{OJ}_{raw}[k] > 0$, from (2.14), $P^{OJ}[k] > 0$; consequently, from (2.15) there are two possible expressions for $\Delta A^{OJ}_{out}[k]$:

(i) $\Delta A^{OJ}_{out}[k] = -K_2 P^{OJ}[k]$, and

(ii) $\Delta A^{OJ}_{out}[k] = -A_{out}[k]$.

Since $P^{OJ}[k] > 0$ and $A_{out}[k] > 0$, it follows that $\Delta A^{OJ}_{out}[k] < 0$. From (2.2) and (2.5),

$$A_{out}[k+1] = A_{out}[k] + \Delta A^{CD}_{out}[k] + \Delta A^{OJ}_{out}[k].$$

From the above expression, it is evident that $\Delta A^{OJ}_{out}[k]$ has a tendency to reduce attitude. Even though the $\Delta A^{CD}_{out}[k]$ term may actually be positive, its effect is small because the dissonance pressure that exists can only be residual dissonance pressure from previous samples. Extensive simulation results indicate that the attitude change arising from overjustification pressure is greater than that of any residual dissonance pressure. Figure 2.7 is an example of such a simulation. In this simulation, a sufficiently large step reward is applied to the system, thus increasing $A_{out}[k]$ to a positive value. At this point, overjustification pressure arises and, and shown in Figure 2.7 is likely to be greater in magnitude than the residual dissonance pressure arising from previous, raw, unprocessed dissonance pressure. Although the overjustification pressure is larger in magnitude and has a decreasing effect on $A_{out}[k]$, the simulation results suggest that the internal attitude remains positive and thus, the overjustification pressure does not cause attitude to change signs. This result is formalized in the lemma below.

Lemma 2.2. *If, for the one-person system, there is a* $\bar{k}$ *such that*

(i) $P^{CD}[\bar{k}] \geq 0$*,*

(ii) $A_{rew}[\bar{k}] \geq 0$*, and*

(iii) $A_{out}[\bar{k}] \geq 0$*,*

then for $k \geq \bar{k}$*,* $A_{out}[k] \geq 0$*.*

Proof. See Appendix C.2. □

The requirement that $P^{CD}[\bar{k}] \geq 0$ is merely a technical condition arising from combining dissonance pressure and overjustification pressure into the same model, as negative values of $P^{CD}[k]$ may case attitude to become negative. Therefore, when $A_{out}[k] > 0$, $A_{rew}[k] > 0$ and $B[k] > 0$, overjustification pressure arises, resulting in a decreased internal attitude that, from Lemma 2.2, remains non-negative.

To show that our model exhibits the second trend, again, the overjustification effect modelling equations are studied. Consider two attitudes, $A_{out1}[k]$ and $A_{out2}[k]$, where $A_{out1}[k] > A_{out2}[k] > 0$. Then, from (2.13), $P^{OJ}_{raw1}[k] > P^{OJ}_{raw2}[k]$. Assuming $P^{OJ}_1[k-1] = P^{OJ}_2[k-1]$, from (2.14), the overjustification pressure arising from the larger attitude is greater than that arising from the smaller attitude, i.e.,

$$P^{OJ}_1[k] > P^{OJ}_2[k].$$

The same argument can be applied in the case of two different reward attitude values. To see how these two overjustification pressures affect the attitude change magnitude, the attitude change equation, given in (2.15), is used.

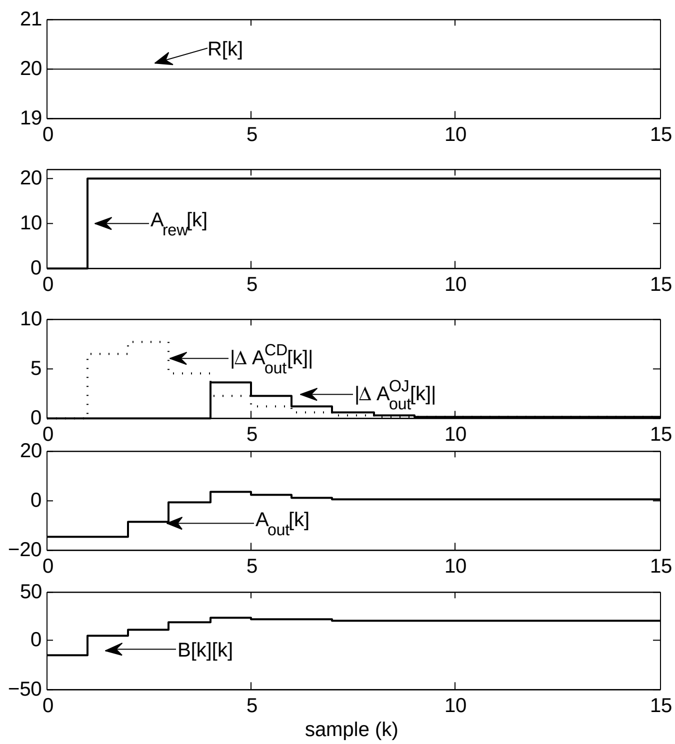

Figure 2.7: A step reward ($R[k] = \$20$) is applied to the one-person system with the initial conditions given in (2.16) with $A_o = -15$ *attitude units*. This reward generates a reward attitude that is large enough to cause $B[k] \geq 0$ for $k \geq 1$. This positive behaviour produces dissonance pressure that causes attitude to increase and, at $k = 4$, become positive. Once attitude becomes positive, overjustification pressure arises. Even though residual dissonance pressure exists, the overjustification pressure attitude change, $\Delta A_{out}^{OJ}[k]$, overpowers the positive effect of the dissonance pressure attitude change, $\Delta A_{out}^{CD}[k]$, and, therefore, causes attitude to decrease but remain non-negative.

First, suppose $K_2 P_2^{OJ}[k] \leq A_{out2}[k]$; then from (2.15), $|\Delta A_{out2}^{OJ}[k]| = K_2 P_2^{OJ}[k] < K_2 P_1^{OJ}[k]$. If $K_2 P_1^{OJ}[k] \leq A_{out1}[k]$, then from (2.15), $|\Delta A_{out1}[k]| = K_2 P_1^{OJ}[k]$, implying

$$|\Delta A_{out2}^{OJ}[k]| < |\Delta A_{out1}^{OJ}[k]|.$$

If, on the other hand, $K_2 P_1^{OJ}[k] > A_{out1}[k]$, then from (2.15), $|\Delta A_{out1}[k]| = A_{out1}[k]$. Since, $A_{out1}[k] > A_{out2}[k]$ and from above, $A_{out2}[k] \geq K_2 P_2^{OJ}[k] = |\Delta A_{out2}^{OJ}[k]|$, it follows that

$$|\Delta A_{out2}^{OJ}[k]| < |\Delta A_{out1}^{OJ}[k]|.$$

Therefore, if $K_2 P_2^{OJ}[k] \leq A_{out2}[k]$, then the attitude change magnitude resulting from the larger attitude is greater than the attitude change magnitude resulting from the smaller attitude.

Now, suppose $K_2 P_2^{OJ}[k] > A_{out2}[k]$; then from (2.15), $|\Delta A_{out2}^{OJ}[k]| = A_{out2}[k] < A_{out1}[k]$. If $K_2 P_1^{OJ}[k] > A_{out1}[k]$, then from (2.15), $|\Delta A_{out1}[k]| = A_{out1}[k]$, implying

$$|\Delta A_{out2}^{OJ}[k]| < |\Delta A_{out1}[k]|.$$

If, on the other hand, $K_2 P_1^{OJ}[k] \leq A_{out1}[k]$, then from (2.15), $|\Delta A_{out1}[k]| = K_2 P_1^{OJ}[k]$. However, since $K_2 P_1^{OJ}[k] > K_2 P_2^{OJ}[k]$ and from above $K_2 P_2^{OJ}[k] > A_{out2}[k]$, it follows that

$$|\Delta A_{out2}^{OJ}[k]| < |\Delta A_{out1}[k]|.$$

Therefore, if $K_2 P_2^{OJ}[k] > A_{out2}[k]$, then again, the attitude change magnitude resulting from a larger attitude is greater than that arising from the smaller attitude and thus, our model exhibits the second qualitative trend of the overjustification effect.

In conclusion, the one-person system, modelled in Section 2.3 from the psychology concepts introduced in Sections 2.1 and 2.2, exhibits the key qualitative characteristics of theory of planned behaviour, cognitive dissonance theory and the overjustification effect. Not only is the one-person system qualitatively consistent with the relevant psychology, two lemmas provide some quantitative consistency. The first lemma quantifies the maximum amount of dissonance pressure magnitude; the second formally shows the overjustification effect cannot cause attitude to become negative. These two results prove useful in the next chapter, when open-loop and closed-loop controllers are designed.

Chapter 3

One-Person System: Simulation and Control Strategies

This chapter explores open-loop and closed-loop control strategies to achieve a control objective for the one-person system with the initial conditions given in (2.16): the control objective is to determine whether or not, for any $B_d > 0$, there exists a reward sequence, $R[k]$, such that $B[k] \geq B_d$ as k tends to infinity. For this analysis, Assumptions 2.5–2.8 are used, as are two technical lemmas found in Appendix A.

3.1 Open-Loop Investigation

To begin investigating whether or not there exists a reward sequence, $R[k]$, such that $B[k] \geq B_d$ as $k \to \infty$, open-loop control strategies are considered. Motivated by the psychology presented in Chapter 2, an impulse reward is first examined. However, Section 3.1.1 reveals that this control strategy is not sufficient to ensure the control objective is met for any initial attitude, A_o, and desired behaviour, B_d. Nevertheless, the impulse-reward results are promising and suggest that by simply offering the same reward at each sample k, the control objective may be achievable. Indeed, Section 3.1.2 demonstrates that a sufficiently large step reward is able to meet the control objective for any initial attitude and desired behaviour.

3.1.1 Impulse-Reward Controller Design

Recall that the control objective is to find a sequence of rewards such that $B[k] \geq B_d$ as k tends to infinity. Given that our model is qualitatively consistent with the relevant psychology, a natural starting point for finding such a sequence of rewards is the psychology literature, specifically, the psychological experiments related to the induced compliance paradigm of cognitive dissonance theory. In these experiments, psychologists offer subjects a reward at a single instance in time in an attempt to influence behaviour. From a controls point of view, this translates into an impulse control signal; hence, consider a control signal given by

$$R[k] = \begin{cases} R_o & \text{if } k = 0, \\ 0 & \text{otherwise,} \end{cases} \tag{3.1}$$

where $R_o > 0$. From (2.17), this control signal generates a reward attitude signal given by

$$A_{rew}[k] = \begin{cases} \mu_1 R_o & \text{if } k = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3.2}$$

In the context of the piano example, applying an impulse reward to the one-person system means the child's parents are offering him a one-time reward, with the hope that the child *eventually* produces the desired behavioural intent, denoted here by some $B_d > 0$.

The reward attitude given in (3.2) simplifies the equation for $B[k]$ given in (2.4). The behaviour is simply a sum of two attitudes, $A_{out}[k]$ and $A_{rew}[k]$. However, for $k \geq 2$, the reward attitude is zero and therefore for $k \geq 2$,

$$B[k] = A_{out}[k]. \tag{3.3}$$

A direct result of (3.3) is that changes to $A_{out}[k]$ produce changes to $B[k]$.

Since the initial conditions state $B[0] < 0$, it follows that to meet the control objective, the behaviour must increase; consequently, $A_{out}[k]$ must necessarily increase. From the psychology presented in Section 2.2, changes to the internal attitude are caused by dissonance pressure and overjustification pressure. However, due to the nature of the reward signal and the initial conditions, overjustification pressure does not arise in the case of an impulse reward. The lemma below provides the foundation upon which this claim is made.

Lemma 3.1. *For all $\bar{k} \geq 0$, if, for each k in the interval $0 \leq k \leq \bar{k}$, one or more of the following conditions hold:*

(i) $A_{out}[k] < 0$,

(ii) $A_{rew}[k] = 0$, *and/or*

(iii) $B[k] < 0$,

then for $0 \leq k \leq \bar{k}$, $\Delta A_{out}^{OJ}[k] = 0$.

Proof. If any of the conditions hold for each k in the interval $0 \leq k \leq \bar{k}$, then the raw, unprocessed overjustification pressure, given in (2.13), equals zero, i.e., $P^{OJ}_{raw}[k] = 0$. Since the raw, unprocessed overjustification pressure is zero, then for $0 \leq k \leq \bar{k}$, the processed dissonance equation, given in (2.14), becomes $P^{OJ}[k] = r_3 P^{OJ}[k-1]$. The initial conditions given in (2.16) imply $P^{OJ}[k] = 0$ for $0 \leq k \leq \bar{k}$. Therefore, from (2.15), $\Delta A^{OJ}_{out}[k] = 0$ for $0 \leq k \leq \bar{k}$. □

Notice that $A_{out}[0] < 0$ and from the initial conditions, $A_{out}[1] = A_{out}[0] + \Delta A_{out}[0] = A_{out}[0]$; hence the first relationship of Lemma 3.1 holds for $k = 0$ and $k = 1$. Moreover, from (3.2), $A_{out}[k] = 0$ for $k \geq 2$ and thus, the second relationship of Lemma 3.1 holds for $k \geq 2$. Combining these two facts shows that for $k \geq 0$, the conditions of Lemma 3.1 hold and hence, the above claim that overjustification pressure does not arise follows. Consequently, in the case of an impulse reward,

$$\Delta A_{out}[k] = \Delta A^{CD}_{out}[k], \tag{3.4}$$

i.e., the only factor contributing to attitude (and thus, behaviour) change is dissonance pressure.

As shown in Section 2.4, the amount of reward offered to induce a behaviour affects the attitude change direction. In particular, an insufficient reward causes an initially negative attitude to become more negative, whereas a sufficiently large reward produces an increase in the internal attitude. For an impulse reward, there is only one sample at which the reward directly influences the behaviour: $k = 1$. From (2.4), (2.16) and (3.2), $B[1] = A_o + \mu_1 R_o$, which means that an insufficient reward that produces a negative $B[1]$ is given by

$$R_o < \frac{-A_o}{\mu_1}, \tag{3.5}$$

whereas a sufficiently large reward that produces a non-negative $B[1]$ is given by

$$R_o \geq \frac{-A_o}{\mu_1}. \tag{3.6}$$

Note that due to the initial conditions, $-A_o > 0$ and therefore the term on the right-hand side of both (3.5) and (3.6). This is important because the reward must at least satisfy $R[k] \geq 0$, as the model is not valid for negative reward values.

From Section 2.4, (3.5) should cause the attitude to decrease because it is unable to produce a positive behaviour. On the other hand, Section 2.4 suggests that (3.6) causes the attitude to increase due to its ability to drive behaviour positive. The theorem below combines these two hypotheses with the fact that $B[k] = A_{out}[k]$ for $k \geq 2$ to demonstrate that the impulse response of $B[k]$ depends on which of the two inequalities is satisfied.

Theorem 3.1. *If a control signal of the form (3.1) is applied to the one-person system with the initial conditions given in (2.16), then*

(a) if $R_o < \frac{-A_o}{\mu_1}$*, then* $B[k]$ *is a decreasing function of* k *for* $k \geq 2$*; and*

(b) if $R_o \geq \frac{-A_o}{\mu_1}$*, then* $B[k]$ *is an increasing function of* k *for* $k \geq 2$*.*

Proof. See Appendix D.1. □

Theorem 3.1 states that if an impulse reward is too small, then not only does it fail to produce a positive behaviour, it also has the unintended consequence of causing the behaviour to become more negative. From a practical point of view, this conclusion indicates that a small reward is counter-productive, as the aim of applying a reward is to increase the behaviour.

Furthermore, Theorem 3.1 implies that $R_o \geq \frac{-A_o}{\mu_1}$ is a necessary (though not sufficient) condition to meet the control objective. Since, for $k \geq 2$, $B[k] = A_{out}[k]$, the only way to meet the control objective is if $A_{out}[k] \geq B_d$ as k tends to infinity. From (3.4), $A_{out}[k]$ changes when dissonance pressure arises. From (2.10) an impulse reward produces no raw, unprocessed dissonance pressure for $k \geq 2$. Nevertheless, $P^{CD}_{raw}[1] \neq 0$ and therefore, from (2.11), dissonance pressure arises for $k \geq 1$. Unfortunately, for $k \geq 2$, no new, raw dissonance pressure occurs and therefore, $P^{CD}[k]$ is simply a decaying function. As a result, the ensuing attitude change decays at each sample, which suggests $A_{out}[k]$ converges as k tends to infinity. As the lemma below states, not only does $A_{out}[k]$ converge as k tends to infinity, but an impulse reward has a limited ability to increase the attitude. This result should not be surprising given that from Lemma 2.1, there is a maximum possible raw dissonance pressure, and from (2.10) and (3.2), the raw, unprocessed dissonance pressure is non-zero only at $k = 1$. As the maximum value for $P^{CD}_{raw}[k]$ is 0.5, the results of the following lemma appear reasonable.

Lemma 3.2. *If a control signal of the form (3.1) is applied to the one-person system with the initial conditions given in (2.16), then, provided (3.6) is satisfied,* $A_{out}[\infty]$ *exists and is given by*

$$A_{out}[\infty] = A_o + K_1 \frac{|A_o|}{|A_o| + |\mu_1 R_o|}. \tag{3.7}$$

Moreover, $A_{out}[\infty]$ *is maximized over values of* R_o *when* $R_o = \frac{-A_o}{\mu_1}$*, and the maximum value is given by*

$$\max_{\substack{Ro \\ R_o \geq -\frac{A_o}{\mu_1}}} A_{out}[\infty] = A_o + \frac{K_1}{2}. \tag{3.8}$$

Proof. See Appendix D.2. □

A direct consequence of Lemma 3.2 is that if B_d is greater than the maximum possible steady-state value of $A_{out}[k]$, then the control objective is impossible to meet with an impulse reward. Theorem 3.2 summarizes this result and, for the situation in which achieving the control objective is possible, provides the set of impulse-reward control signals that is able to meet the control objective.

Theorem 3.2. *If a control signal of the form (3.1) is applied to the one-person system with the initial conditions given in (2.16), then it is possible to drive $B[k] \geq B_d$ as k tends to infinity if and only if*

$$A_o + \frac{K_1}{2} \geq B_d. \tag{3.9}$$

Assuming (3.9) is satisfied, the range of R_o values that are able to meet the control objective is given by

$$\frac{|A_o|}{\mu_1} \leq R_o \leq \frac{|A_o|}{\mu_1}\left(\frac{K_1}{B_d - A_o} - 1\right). \tag{3.10}$$

Proof. See Appendix D.3. □

Theorem 3.2 states that meeting the control objective may be impossible for some combinations of A_o, B_d and K_1. Within the context of the piano example, this conclusion means that, for a one-time reward, it may not be possible to drive the child's behaviour to the desired strength if his attitude is quite negative and his sensitivity/susceptibility to dissonance pressure is quite small. To support the results of Theorem 3.2, two simulations are performed. The first simulation, shown in Figure 3.1, demonstrates that when (3.9) is not met, the control objective cannot be achieved. The second simulation, shown in Figure 3.2, demonstrates that when (3.9) is met, the control objective is achievable.

To summarize, for an impulse-reward controller, given in (3.1), if $R_o < \frac{-A_o}{\mu_1}$, then from Theorem 3.1, $B[k]$ is a decreasing function for $k \geq 2$ and therefore, the control objective cannot be met. To meet the control objective, it is necessary that the impulse-reward magnitude satisfies $R_o \geq \frac{-A_o}{\mu_1}$. This reward generates dissonance pressure, which is reduced through increasing $A_{out}[k]$. This increase has the effect of increasing $B[k]$, a prediction of the model that may be of interest to psychologists. Unfortunately, from Lemma 3.2, the improvements gained by the attitude and behaviour are limited, i.e., as k tends to infinity, $A_{out}[k]$ and $B[k]$ converge to the value given by (3.7). Furthermore, for a given A_o, B_d and K_1, there is a maximum steady-state value for $A_{out}[k]$ (and thus, $B[k]$), given by (3.8). If B_d is greater than this maximum value, then meeting the control objective is impossible with an impulse reward. On the other hand, if B_d is less than this maximum value, then it is possible to meet the control objective with a reward magnitude within the range given

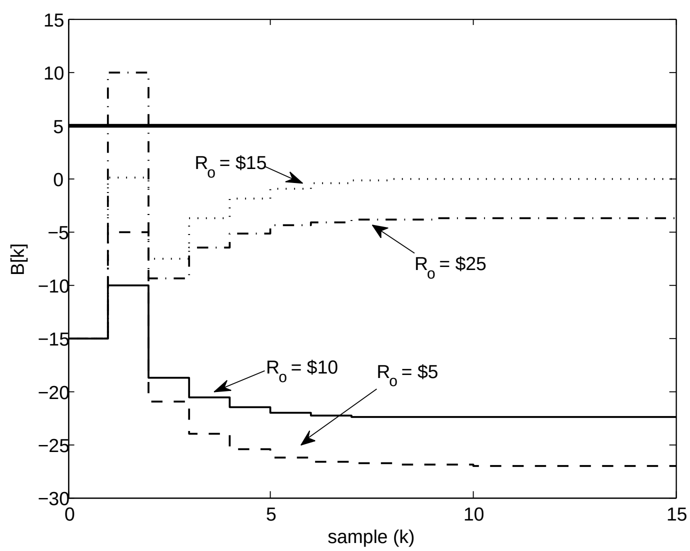

Figure 3.1: Impulse response of $B[k]$ when $A_o = -15$ *attitude units*. The control signal given in (3.1) is applied to the one-person system with the initial conditions in (2.16) and parameter values given in Table 2.2 for various values of R_o. Since the parameter values do not satisfy (3.9), from Theorem 3.2, the reward is not able to produce enough attitude change to ensure $B[k]$ reaches $B_d = 5$ *attitude units* (denoted by the thick solid line) as $k \to \infty$. In particular, in the case when $R_o < 15$, the behaviour decreases, whereas when $R_o \geq 15$, the behaviour increases. In the latter scenario, larger values of R_o produce smaller changes to behaviour and therefore, the simulation results suggest that meeting the control objective is impossible for this combination of initial attitude (A_o), desired behaviour (B_d) and susceptibility to dissonance pressure (K_1).

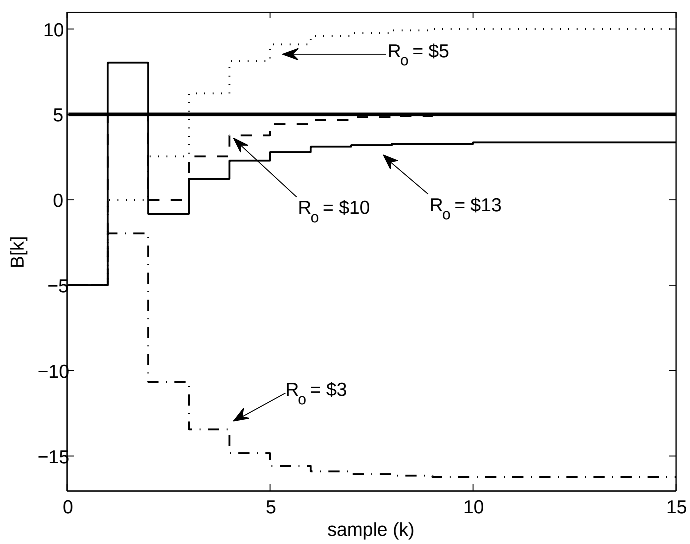

Figure 3.2: Impulse response of $B[k]$ when $A_o = -5$ *attitude units*. The control signal given in (3.1) is applied to the one-person system with the initial conditions in (2.16) and parameter values given in Table 2.2 for various values of R_o. When $R_o < \frac{-A_o}{\mu_1} = \$5$, $B[k]$ decreases and is unable to reach the desired behaviour, $B_d = 5$ *attitude units*, denoted on the graph by the thick solid line. On the other hand, when $\$5 \leq R_o \leq \frac{|A_o|}{\mu_1}\left(\frac{K_1}{B_d - A_o} - 1\right) = \$10$, $B[k]$ not only increases, but is able to meet the control objective. Lastly, when $R_o > \$10$, the dissonance pressure is not sufficient to produce the required change to the internal attitude (and thus, behaviour).

by (3.10). Since the impulse-reward controller is unable to meet the control objective for any initially negative A_o, desired behaviour B_d, and susceptibility to dissonance pressure, K_1, a different control strategy is considered next.

3.1.2 Step-Reward Controller Design

An impulse reward that is sufficiently large has the effect of increasing $A_{out}[k]$ and subsequently $B[k]$. Unfortunately, the amount of attitude and behaviour change that is possible may not be enough to meet the control objective. As discussed in the previous section, the attitude change is produced by dissonance pressure, which, after the reward is removed, is a decaying function. If, however, new dissonance pressure arises at each sample, then the resulting overall attitude change may be greater. Given that a reward generates dissonance pressure, it is natural to predict that a reward offered at every sample creates new dissonance pressure at every sample until $A_{out}[k] \geq 0$. The simplest scheme of this type is a step reward and its effectiveness at meeting the control objective is studied in this section.

Applying a step reward to the one-person system means the child's parents offer him a fixed amount of money each day, with the hope that the child *eventually* produces the desired behavioural intent, denoted here by some $B_d > 0$. Formally speaking, this problem asks whether or not there exists some $R_o > 0$ and control signal

$$R[k] = R_o, \; k \geq 0 \tag{3.11}$$

such that $B[k] \geq B_d$ as k tends to infinity. From (2.17), this control signal generates a reward attitude signal for $k \geq 1$ given by

$$A_{rew}[k] = \mu_1 R_o. \tag{3.12}$$

Since, for $k \geq 1$,

$$B[k] = A_{out}[k] + \mu_1 R_o, \tag{3.13}$$

a change in $A_{out}[k]$ produces some change in $B[k]$. Since changes in $A_{out}[k]$ cause $B[k]$ to change, understanding the factors contributing to $\Delta A_{out}[k]$ is important. From Section 2.2, the internal attitude changes can be generated through dissonance pressure and overjustification pressure. Unlike the impulse reward, a step reward may produce overjustification pressure. This detail is explored later in this section because, initially, the only pressure arising from a step reward is dissonance pressure. From the initial conditions, $A_{out}[1] = A_{out}[0] + \Delta A_{out}[0] = A_o < 0$, and therefore, by Lemma 3.1, $\Delta A_{out}^{OJ}[1] = 0$. Since the only pressure that can arise at $k = 1$ is dissonance pressure, the effect of the reward on this dissonance pressure is first explored.

As previously discussed, when $A_{out}[k] < 0$, a reward that is sufficiently large to produce a positive $B[k]$ causes dissonance pressure. This dissonance pressure is reduced through attitude change, specifically, an increased attitude. On the other hand, an insufficient reward causes dissonance pressure that is decreased by the internal attitude becoming more negative. From (3.13), it appears that if the reward is insufficient at driving $B[1] \geq 0$, then, due to the decreasing $A_{out}[k]$, the behaviour will always remain negative. The range of R_o values that generate this decreasing effect on $A_{out}[k]$ and thus, $B[k]$, is found through the expression for $B[1]$ and the inequality $B[1] < 0$, i.e., $B[1] = A_o + \mu_1 R_o < 0$. Rearranging for R_o gives the reward that causes $A_{out}[k]$ (and hence, $B[k]$) to decrease:

$$R_o < \frac{-A_o}{\mu_1}.$$

On the other hand, if a reward is large enough to drive $B[1] \geq 0$, then the attitude increases, resulting in an increased behaviour. This sufficient reward is given by

$$R_o \geq \frac{-A_o}{\mu_1}.$$

Thus, similar to the impulse reward, a larger reward shows promise in its ability to achieve the control objective. Given that this goal requires $B[k]$ to increase from its initially negative value, it is evident that $R_o \geq \frac{-A_o}{\mu_1}$ must hold. From this discussion, it follows that $B[k]$ can exhibit one of two trends, depending on the value of R_o. The theorem below characterizes these two trends.

Theorem 3.3. *If a control signal of the form (3.11) is applied to the one-person system with the initial conditions given in (2.16), then*

(a) if $R_o < \frac{-A_o}{\mu_1}$, then $B[k]$ tends to negative infinity as k tends to infinity; and

(b) if $R_o \geq \frac{-A_o}{\mu_1}$, then $B[k]$ is an increasing function of k for $0 \leq k \leq T$ where $T := \max\left\{\bar{T} : A_{out}[k] < 0 \text{ for } 0 \leq k \leq \bar{T}\right\}$.

Proof. See Appendix D.4. □

Up until this point, the results of the step reward are quite similar to those of the impulse reward. In the case of a small reward, i.e., $R_o < \frac{-A_o}{\mu_1}$, $B[k]$ decreases for both controllers. For an impulse reward, $B[k]$ converges as k tends to infinity, whereas $B[k]$ tends to negative infinity in the case of a step reward. On the other hand, when $R_o \geq \frac{-A_o}{\mu_1}$, the characteristics exhibited by $B[k]$ are different. In particular, $A_{out}[k]$ exhibits a different trend when it becomes positive. For an impulse reward, overjustification pressure never occurs. However, if, in the case of a step reward, the internal attitude becomes positive, then both $A_{out}[k] > 0$

and $A_{rew}[k] > 0$, implying $B[k] > 0$ and as a result, overjustification pressure arises, decreasing the internal attitude that, by Lemma 2.2, remains positive. Nevertheless, since the internal attitude remains positive (and the reward attitude is positive), $B[k]$ also remains positive. As previously argued, $B[k] \geq 0$ for $1 \leq k \leq T$ and since $B[k]$ remains positive after $A_{out}[k] \geq 0$, i.e., for $k > T$, it follows that if $R_o \geq \frac{-A_o}{\mu_1}$, then $B[k] \geq 0$ for $k \geq 1$. When the behaviour is positive for all $k \geq 1$, several key signals converge as $k \to \infty$, as stated in Lemma 3.3 below.

Lemma 3.3. *For the one-person system with the initial conditions given in (2.16), if $B[k] \geq 0$ for $k \geq 1$, then, as k tends to infinity,*

(a) $A_{out}[k]$ *converges to some constant, c,*

(b) $\Delta A_{out}^{CD}[k]$ *converges to zero, and*

(c) $\Delta A_{out}^{OJ}[k]$ *converges to zero.*

Proof. See Appendix D.5. □

If $R_o \geq \frac{-A_o}{\mu_1}$, then Lemma 3.3 and Theorem 3.3 imply that $A_{out}[\infty]$ exists; consequently, from (3.13), $B[\infty]$ exists. To determine the steady-state value of $B[k]$ the value for $A_{out}[\infty]$ is given in the following lemma.

Lemma 3.4. *If a control signal of the form (3.11) is applied to the one-person system with the initial conditions given in (2.16), and $R_o \geq \frac{-A_o}{\mu_1}$, then*

$$A_{out}[k] \to 0 \text{ as } k \to \infty.$$

Proof. See Appendix D.6. □

An interesting consequence of Lemma 3.4 is the undesirable effect of the reward once $A_{out}[k] \geq 0$. When the internal attitude becomes positive, the reward, through overjustification pressure, causes the attitude to decrease to zero. Consequently, to ensure $B[k] \geq B_d$ as k tends to infinity, the reward must be large enough to generate a reward attitude such that $A_{rew}[k] \geq B_d$. Indeed, by combining the requirement that $R_o \geq \frac{-A_o}{\mu_1}$ with the control objective that $B[\infty] \geq B_d$, a necessary and sufficient condition on R_o is obtained, as stated in the follow theorem.

Theorem 3.4. *For all $B_d > 0$, if a control signal of the form (3.11) is applied to the one-person system with the initial conditions given in (2.16), then it is possible to drive $B[k] \geq B_d$ as k tends to infinity. Indeed, this goal is attained if and only if*

$$R_o \geq \max\left(\frac{-A_o}{\mu_1}, \frac{B_d}{\mu_1}\right).$$

Proof. See Appendix D.7. □

Theorem 3.4 says that a step-reward controller is always able to meet the control objective of driving $B[k] \geq B_d$ as $k \rightarrow \infty$. Within the context of the piano example, this conclusion means that provided the reward is large enough, parents can offer their child a fixed reward every day and the child will eventually play the piano to the desired behavioural intent strength. To support the results of Theorem 3.4, consider the simulations shown in Figure 2.5 of Section 2.4. For the first simulation, a step reward of amplitude $R_o = \$20$ is applied to the one-person system with the initial conditions given in (2.16) and initial attitude $A_o = -15$ *attitude units*. From Theorem 3.3, this reward is large enough to cause $B[k]$ to increase until $A_{out}[k]$ is positive. Once the internal attitude becomes positive, from Lemma 3.4 it converges to zero as k tends to infinity. Finally, from Theorem 3.4, the reward is able to produce the desired behavioural intent, $B_d = 20$ *attitude units*, as k tends to infinity. For the second simulation, a step reward of amplitude $R_o = \$10$ is applied to the one-person system with the initial conditions given in (2.16) and initial attitude $A_o = -15$ *attitude units*. From Theorem 3.3, this reward is too small and therefore, causes $B[k]$ to tend to negative infinity as k tends to infinity.

In conclusion, Theorem 3.4 provides an open-loop control scheme that is always able to meet the control objective. However, the interesting consequence of Lemma 3.4 provides a significant drawback of the step-reward control strategy. In particular, once attitude becomes positive, a reward has a negative effect on the internal attitude. Perhaps then, instead of a step reward, a reward that is constant from $0 \leq k \leq T$ and zero for $k > T$ can be used to capitalize on the contribution a positive internal attitude has on the behavioural intent. This controller, termed the *extended-impulse-reward* controller, is essentially a cross between the impulse reward considered in Section 3.1.1 and the step reward considered in this section. Unfortunately, T cannot be easily determined without feedback. This provides one reason for investigating closed-loop control strategies.

3.2 Closed-Loop Investigation

The particular open-loop controllers considered thus far have drawbacks: the impulse reward is not always able to meet the control objective and the step reward eventually eliminates the contribution a positive internal attitude has on the behavioural intent. Motivated by the benefits of closed-loop control (i.e., the ability to more effectively fight disturbances and uncertainty), a closed-loop approach is taken instead of continuing to investigate other open-loop control strategies. This section begins by investigating state-feedback controllers, the first of which is the extended-impulse-reward controller originally proposed in the previous section. Considering the impracticality of state measurement at each sample, an output-feedback controller is then designed.

3.2.1 State-Feedback Controllers

State-feedback is a natural place to start for closed-loop controller design. First, motivated by the open-loop control results, an extended-impulse-reward controller is considered. However, the analysis reveals that this controller is unable to meet the control objective for most values of B_d. For this reason, a second state-feedback controller is then designed.

Extended-Impulse-Reward Controller

The extended-impulse-reward controller is essentially a cross between an impulse reward and a step reward. Loosely speaking, this controller applies a step reward to the one-person system until the internal attitude becomes positive. Once the internal attitude becomes positive, the reward is removed to ensure its undesirable effect on a positive internal attitude does not arise. From Theorem 3.3, a step reward with magnitude $R_o < \frac{-A_o}{\mu_1}$ causes $B[k]$ to decrease to negative infinity, whereas a step reward with magnitude $R_o \geq \frac{-A_o}{\mu_1}$ generates dissonance pressure that forces $B[k]$ to increase for $0 \leq k \leq T$ where T is the last sample at which $A_{out}[k] < 0$. It is apparent that for the extended-impulse-reward controller to meet the control objective, $R_o \geq \frac{-A_o}{\mu_1}$ must hold when the reward is "turned on."

To determine the sample at which the reward "turns off," the overjustification pressure equations are studied. The overjustification pressure causes a positive internal attitude to decrease (but remain positive). This undesirable effect is eliminated by guaranteeing that overjustification pressure never arises. The overjustification pressure arises when, from (2.13), $A_{out}[k] > 0$, $A_{rew}[k] > 0$ and $B[k] > 0$; thus, by ensuring $A_{rew}[k] = 0$ when $A_{out}[k] > 0$, the overjustification pressure is forced to be zero, i.e., $P^{OJ}[k] = 0$, and subsequently, $\Delta A^{OJ}_{out}[k] = 0$ as desired.

To ensure overjustification pressure does not arise, $A_{rew}[k] = 0$ when $A_{out}[k] > 0$; but from (2.17), $A_{rew}[k]$ is a scaled, delayed version of $R[k]$. The condition that $R[k-1] = 0$ when $A_{out}[k] > 0$ is non-causal. However, from (2.2),

$$A_{out}[k] = A_{out}[k-1] + \Delta A_{out}[k-1].$$

As a result, the condition guaranteeing overjustification pressure does not arise becomes $R[k-1] = 0$ when $A_{out}[k-1] + \Delta A_{out}[k-1] > 0$. Hence, the control signal can be defined, for $R_o > 0$, as

$$R[k] = \begin{cases} R_o & \text{if } A_{out}[k] + \Delta A_{out}[k] < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.14}$$

Regarding the timing sequence needed to implement 3.14, recall from Section 2.3.4 that $A_{out}[k]$ is the child's internal attitude at the beginning of the sample period (taken for this thesis to be a day). The child begins the day with some internal attitude, $A_{out}[k]$. Throughout the day, some amount of attitude change may arise, $\Delta A_{out}[k]$, and at the end of the day, the parents offer their child a reward, $R[k]$. Therefore, the control signal at sample k is formed at the end of the sample period and depends on states that are evaluated at the beginning and middle of the sample period.

From (2.17), the control signal in (3.14) yields the following equation for the reward attitude:

$$A_{rew}[k] = \begin{cases} \mu_1 R_o & \text{if } A_{out}[k] < 0, \\ 0 & \text{otherwise.} \end{cases} \tag{3.15}$$

The above equation for $A_{rew}[k]$ implies that for $k > T$, $B[k] = A_{out}[k]$. To see why, note that at $k = T$, $A_{out}[k] < 0$ but $A_{out}[k+1] = A_{out}[k] + \Delta A_{out}[k] \geq 0$ and therefore, $A_{rew}[k+1] = 0$, i.e., $A_{rew}[k+1] = 0$ for $k > T$.

Practically speaking, using this reward scheme, parents offer their child a fixed reward each day to play the piano, until it becomes clear that the child's internal attitude is about to become positive. At this time, the reward is no longer offered. However, as shown in the theorem below, the benefits of this strategy are limited.

Theorem 3.5. *If a control signal of the form (3.14) is applied to the one-person system with the initial conditions given in (2.16), then it is possible to drive $B[k] \geq 0$ for $k \geq 1$. Indeed, this goal is attained if and only if*

$$R_o \geq \frac{-A_o}{\mu_1}.$$

Proof. See Appendix D.8. □

As indicated in Theorem 3.5, the extended-impulse-reward strategy only guarantees that a positive behaviour is obtained. To understand why $B[k]$ cannot be driven arbitrarily large, note that the only pressure driving attitude change is dissonance pressure. Given that the reward attitude is zero when the internal attitude is positive, all cognitions are positive and are thus, consistent; therefore, no new, raw, unprocessed dissonance pressure arises upon the internal attitude switching from being negative. Similar to the impulse-reward strategy, in which the dissonance pressure decays because no new, raw, unprocessed dissonance pressure arises after a certain sample, in the case of the extended-impulse-reward controller, the dissonance pressure is a decaying function upon the internal attitude becoming positive because the raw, dissonance pressure is zero. Since the extended-impulse-reward controller ensures $B[k] \geq 0$ for $k \geq 1$, by Lemma 3.3, $A_{out}[k]$ converges as k tends to infinity. Since $B[k] = A_{out}[k]$ for $k > T$, $B[k]$ also converges as

$k \to \infty$. In the case of small values of B_d, this control strategy may meet the control objective, but unfortunately, for high values of B_d, the value to which the behaviour converges may not reach this desired behavioural intent.

The simulation shown in Figure 3.3 is consistent with above discussion. The system's initial conditions and parameter values are such that the conditions given in Theorem 3.5 are met. As shown in these results, $B[k] \geq 0$ for $k \geq 0$ and settles as k tends to infinity. These results show that if B_d is small, say $B_d = 5$ *attitude units*, then the control objective of driving $B[k] \geq B_d$ as k tends to infinity is met. On the other hand, for large values of B_d, say $B_d = 20$ *attitude units*, the control objective is not met. Hence, an alternative feedback-controller is considered.

General State-Feedback Controller

Recall that the control objective is to drive $B[k] \geq B_d$ as k tends to infinity. The control objective focuses on $B[k]$ and therefore, a natural place to start is examining the equation for $B[k]$:

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

It follows from this equation that if, for all $k \geq 1$, $A_{out}[k] + A_{rew}[k] \geq B_d$, then the control objective is met, not only as $k \to \infty$, but for all $k \geq 1$. Observe that because $A_{rew}[k] = \mu_1 R[k-1]$, the reward enters the above inequality in a straight-forward manner. As a result, the above inequality can be arranged as

$$R[k-1] \geq \frac{1}{\mu_1} \left(B_d - A_{out}[k] \right).$$

Using the equation for $A_{out}[k]$ given in (2.2), this inequality can be re-written as

$$R[k-1] \geq \frac{1}{\mu_1} \left(B_d - A_{out}[k-1] - \Delta A_{out}[k-1] \right).$$

Accounting for the requirement that $R[k] \geq 0$ and shifting samples suggests a state-feedback controller given by

$$R[k] = \max \left\{ 0, \frac{1}{\mu_1} \left(B_d - A_{out}[k] - \Delta A_{out}[k] \right) \right\}, \tag{3.16}$$

which, from the theorem below, meets the control objective.

Theorem 3.6. *For all $B_d > 0$, if the control signal given in (3.16) is applied to the one-person system, then $B[k] \geq B_d$ for $k \geq 1$.*

Proof. See Appendix D.9. □

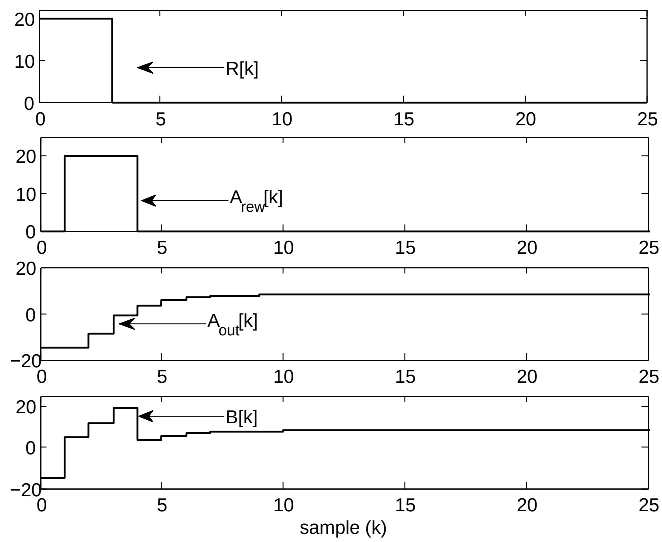

Figure 3.3: Simulation results of the one-person system when the extended-impulse-reward controller given in (3.14) is applied. The system's initial conditions are given by (2.16) and parameter values are given in Table 2.2. Moreover, the initial attitude is given by $A_o = -15$ *attitude units* and the reward magnitude is given by $R_o = \$20$. This reward satisfies the conditions given in Theorem 3.5 and produces the reward attitude shown in the second plot. This reward attitude is large enough to drive behaviour positive at $k = 1$, shown in the bottom plot, and therefore, $A_{out}[k]$ increases, as shown in the third plot. Immediately preceding $A_{out}[k]$ becoming positive, the reward equals zero, thus ensuring $A_{rew}[k] = 0$ when $A_{out}[k] \geq 0$. For the above simulation, $A_{out}[k] \geq 0$ at $k = 4$ and since $A_{out}[k] \geq 0$ for $k \geq 4$ (by Lemma 2.2), the controller ensures $A_{rew}[k] = 0$ for $k \geq 4$. Consequently, $B[k] = A_{out}[k]$, implying $B[k]$ is also non-negative. In other words, $B[k] \geq 0$ for $k \geq 1$. If B_d is small, then the control objective is met, whereas the control objective is not achieved for larger values of B_d.

Note that Theorem 3.6 does not depend on the initial conditions. That is, the state-feedback controller given in (3.16) meets the control objective for any set of initial conditions (due to how the reward enters the expression for $B[k]$). Additionally, the controller given in (3.16) suggests that not only can a constant B_d be achieved, but by modifying (3.16) to

$$R[k] = \max\left\{0, \frac{1}{\mu_1}\left(B_d[k] - A_{out}[k] - \Delta A_{out}[k]\right)\right\}. \tag{3.17}$$

for some bounded $B_d[k]$, it is possible to guarantee that $B[k] \geq B_d[k-1]$ for $k \geq 1$. This result is formalized in the theorem below.

Theorem 3.7. *For any bounded sequence $B_d[\cdot]$ satisfying, for all $k \geq 0$, $0 < B_d[k] \leq B_{d,max}$, if the control signal given in (3.17) is applied to the one-person system, then $B[k] \geq B_d[k-1]$ for $k \geq 1$.*

Proof. See Appendix D.10. □

The results of Theorem 3.7 are stronger than those of Theorem 3.6 since they allow the desired behaviour to vary at each sample. The simulation results shown in Figure 3.4 support Theorems 3.6 and 3.7. The plots on left-hand side of Figure 3.4 show a simulation demonstrating that the controller given in (3.16) ensures $B[k] \geq B_d$ for $k \geq 1$ when applied to the one-person system. On the other hand, the plots on the right-hand side of Figure 3.4 show a simulation demonstrating that the controller given in (3.17) ensures $B[k] \geq B_d[k-1]$ for $k \geq 1$. For both simulations, the system's initial conditions do not satisfy (2.16). If, however, the initial conditions given in (2.16) are met, then further conclusions can be drawn, given in the theorem below.

Theorem 3.8. *For any bounded sequence $B_d[\cdot]$ satisfying, for all $k \geq 0$, $0 < B_d[k] \leq B_{d,max}$, if the control signal given in (3.17) is applied to the one-person system with the initial conditions given in (2.16), then*

(a) $B[k] \geq B_d[k-1]$ for $k \geq 1$,

(b) $A_{rew}[\cdot]$ and $B[\cdot]$ remain bounded, and

(c) $A_{out}[k]$, $\Delta A^{CD}_{out}[k]$, $\Delta A^{OJ}_{out}[k]$, $P^{CD}[k]$ and $P^{OJ}[k]$ all converge as k tends to infinity.

Moreover, if $B_d[k]$ is constant for all $k \geq 0$, then the controller maintains internal stability in the sense that all signals converge as $k \to \infty$.

Proof. See Appendix D.11. □

The simulation results given in Figure 3.5 support the above theorem. The plots on the left-hand side demonstrate that when $B_d[k]$ is a constant, i.e., $B_d[k] = B_d$ for all $k \geq 0$, all signals converge as k tends to infinity and $B[k] \geq B_d$ for $k \geq 1$. On the other hand, the plots on the right demonstrate the more general result that for a bounded sequence, $B_d[k]$, $B[k] \geq B_d[k-1]$ for $k \geq 1$ and, $A_{rew}[k]$ and $B[k]$ remain bounded while all other signals converge as k tends to infinity. Furthermore, although Theorem 3.8 requires the initial conditions given in (2.16) be met to ensure conclusions (b) and (c), simulation results suggest that meeting this requirement is not necessary. The results shown in Figure 3.4 provide evidence supporting this hypothesis, as these simulations use initial conditions that do not satisfy (2.16), yet the conclusions of Theorem 3.8 above still hold.

Unfortunately, the controllers given in Theorems 3.6, 3.7 and 3.8 bear the same drawback as the step reward: a positive internal attitude may be eliminated due to overjustification pressures, and as a result, the reward must be large enough to generate the desired behaviour. At this point, it is evident that there is a trade-off between meeting the control objective of driving $B[k] \geq B_d$ as $k \to \infty$, and maintaining some amount of positive internal attitude. A reward may be necessary for the former to occur, thus eliminating any positive internal attitude that may arise. On the other hand, maintaining a positive internal attitude (should it arise) is possible using the extended-impulse-reward controller; but the behavioural intent, while positive, may not meet the desired strength, B_d.

Both state-feedback controllers suffer an additional drawback, which is practical in nature. Specifically, measuring the state information at each sample is, at best, quite difficult. In particular, at each sample, these controllers require knowledge of the internal attitude and the amount of attitude change arising from dissonance and overjustification pressures. For the piano example, this amounts to the parents measuring their child's internal attitude, the pressures he experiences, and the extent to which these pressures influence his attitude change. This is hardly a feasible task. Perhaps though, there is a controller that simply relies on the child's behaviour, i.e., the system output. The next section investigates whether or not such an output-feedback controller exists.

3.2.2 Output-Feedback Controller Design

Output-feedback control has the benefit of simply requiring information about the output, which from an implementation perspective, is easier than needing to know full state information. In the remaining sections, output-feedback control is studied: this section designs a suitable output-feedback controller and the next section analyses the system's response to this controller. First, a simple proportional controller is considered as an option for meeting the control objective. Although this approach does not work well, it motivates another linear control approach to controller design. This latter control approach exploits

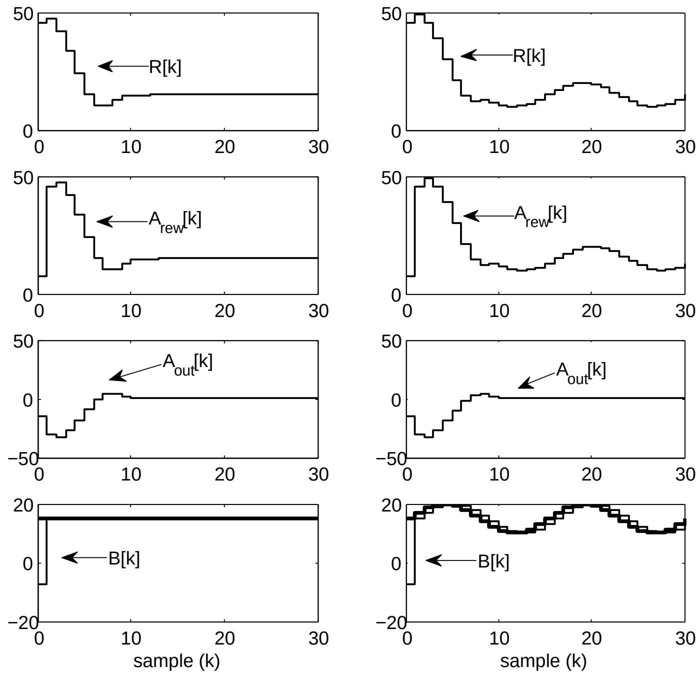

Figure 3.4: Simulation results supporting Theorems 3.6 and 3.7. For both simulations, the one-person system's parameters are given in Table 2.2, the initial conditions do not satisfy (2.16), and $A_o = -15$ *attitude units*. The plots on the left-hand side show simulation results when the controller given in (3.16) is applied to the one-person system, where $B_d = 15$ *attitude units*, denoted by the thick, solid line. The simulation results confirm that $B[k] \geq B_d$ for $k \geq 1$, as stated in Theorem 3.6. The plots on the right-hand side show simulation results when the controller given in (3.17) is applied to the one-person system, where $B_d[k]$ is given by the thick, solid line. The simulation results confirm that $B[k] \geq B_d[k-1]$ for $k \geq 1$, as stated in Theorem 3.7.

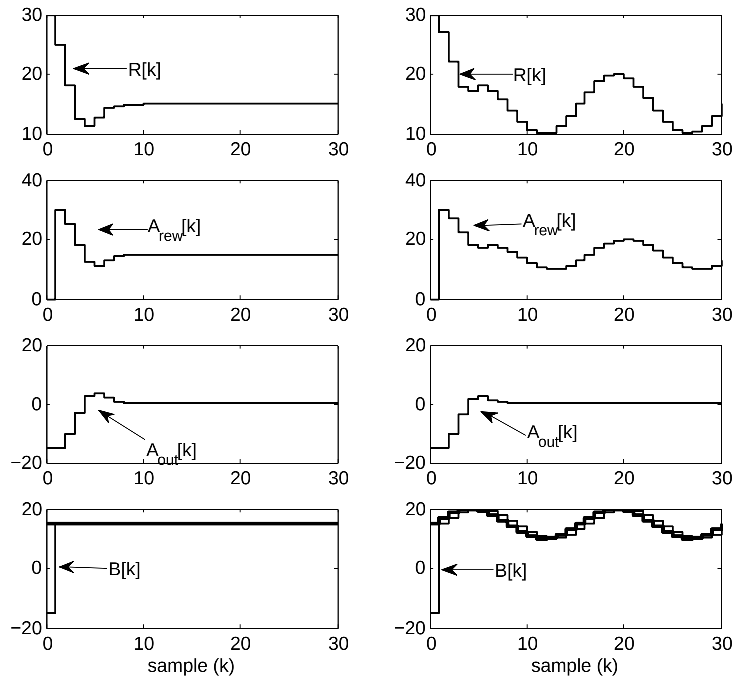

Figure 3.5: Simulation results supporting Theorem 3.8. For both simulations, the one-person system's parameters are given in Table 2.2, the initial conditions satisfy (2.16) and $A_o = -15$ *attitude units*. For the simulation results shown on the left, $B_d[k] = B_d$ for $k \geq 0$ and therefore, the controller in (3.17) simplifies to (3.16). This controller drives $B[k] \geq B_d$ for $k \geq 1$ and ensures the system maintains internal stability. For the simulation results shown on the right, the controller given in (3.17) is applied to the one-person system for some bounded sequence, $B_d[k]$. These plots support Theorem 3.8, as $A_{rew}[k]$ and $B[k]$ remain bounded, while the other signals converge as k tends to infinity.

the system characteristics discussed in previous sections to give an output-feedback controller that, under some mild technical conditions, appears to meet the control objective of driving $B[k] \geq B_d$ as k tends to infinity.

From a practical standpoint, a proportional controller allows parents to measure their child's behaviour at each sample and offer a reward proportional to the difference between the measured value and the desired behavioural intent, B_d. This is a simple approach and is thus the starting point of the output-feedback controller design. Consider the proportional controller given by

$$R[k] = K_c \left(B_d - B[k]\right), \tag{3.18}$$

where, $K_c > 0$, is the controller gain and $B[k]$ is given by (2.4). From a control engineering perspective, proportional control has potential drawbacks, including poor steady-state tracking. As the control objective is to generate a behaviour that is at least as large as B_d, control theory suggests that (3.18) may not be sufficient (since $B[k]$ could settle to some value less that B_d). Simulation results suggest that this is indeed the case. Figure 3.6 shows a simulation demonstrating that a proportional controller is not necessarily able to meet the control objective.

Although a proportional controller is not necessarily able to meet the control objective, the strategy of employing linear control techniques is appealing. Even though the one-person system is non-linear, it can be arranged in a manner similar to the framework of linear output-feedback systems. Figure 3.7 shows the standard linear feedback control system with an output disturbance. The reference ($r[k]$), control ($u[k]$) and output ($y[k]$) signals in Figure 3.7 can map to the desired behaviour (B_d), reward ($R[k]$) and behaviour ($B[k]$) of our model respectively. The expression for $B[k]$ given in (2.4) suggests a straightforward way to integrate the remainder of the one-person system into the framework shown in Figure 3.7. Specifically,

$$B[k] = A_{out}[k] + A_{rew}[k],$$

and therefore, one of the two attitude signals can be the plant signal $\bar{y}[k]$, while the other can be represented by the disturbance, $d[k]$. The control signal choice suggests that $A_{rew}[k]$ be $\bar{y}[k]$, whereas $A_{out}[k]$ be considered as $d[k]$.

Identifying the expression for $P[z]$ follows directly from the model equations. From (2.17), the plant output ($A_{rew}[k]$) is a scaled, delayed version of the plant input ($R[k]$). As a result,

$$P[z] = \frac{\mu_1}{z}. \tag{3.19}$$

The initial conditions given in (2.16) imply the plant above has zero initial conditions.

Incorporating the nonlinear expression for $A_{out}[k]$ is more involved. From the model equations, $A_{out}[k]$ can be expressed as the sum of the initial attitude and all previous

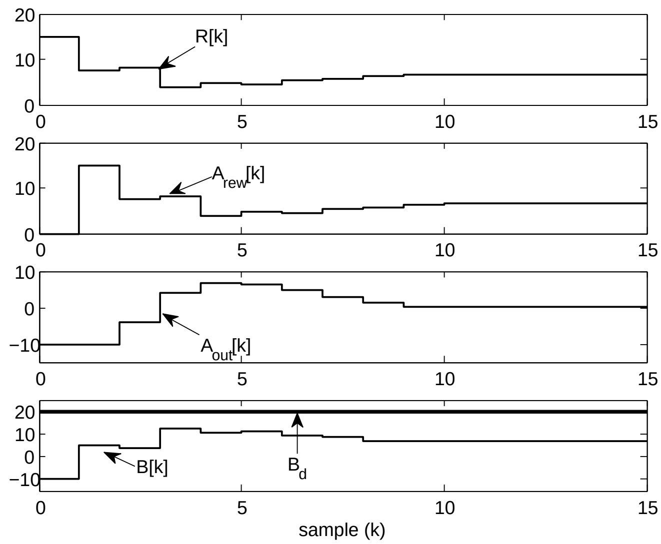

Figure 3.6: Simulation indicating that a proportional controller may not be able to meet the control objective of driving $B[k] \geq B_d$ as k tends to infinity due to its poor steady-state tracking capabilities. For this simulation, the proportional controller given in (3.18), with $K_c = 0.5$, is applied to the one-person output-feedback system with the initial conditions given in (2.16), with $A_o = -10$ *attitude units*, and parameter values given in Table 2.2. The simulation results show that $B[k]$ converges as $k \to \infty$, as do $A_{rew}[k]$ and $A_{out}[k]$. As predicted, there is a steady-state tracking error, since $B[k]$ settles to some value less than $B_d = 20$ *attitude units*(thick line).

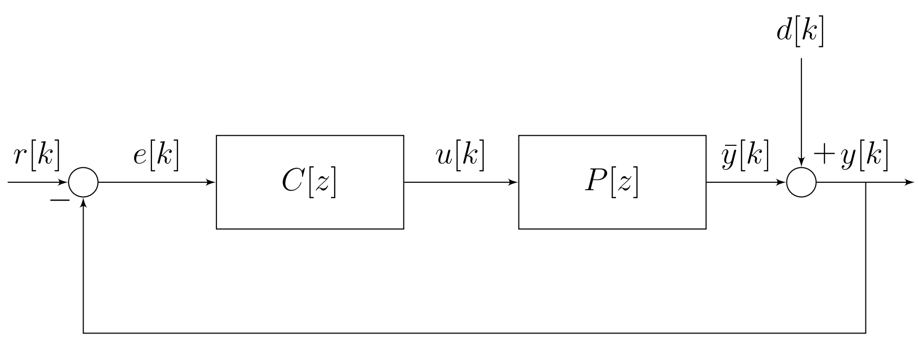

Figure 3.7: Block diagram of the standard linear output-feedback system with disturbance.

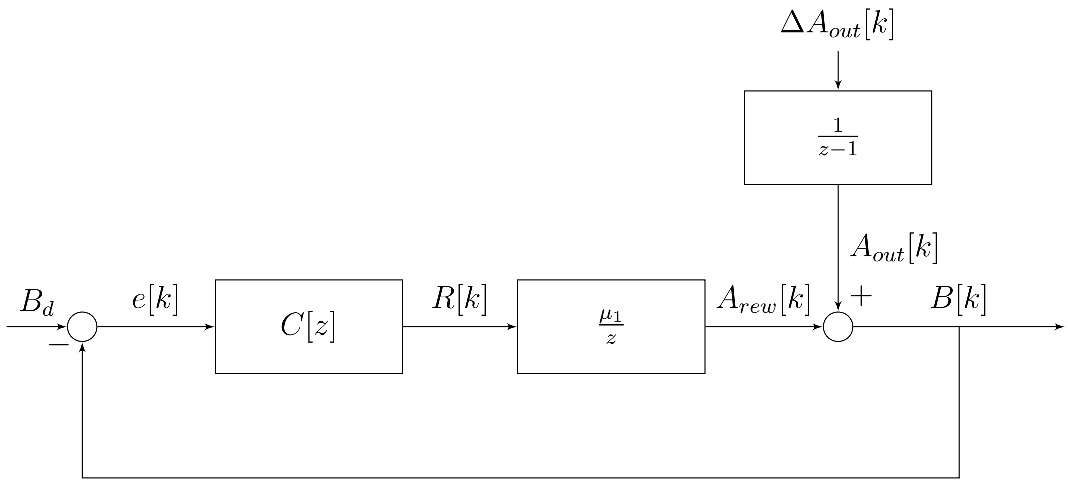

Figure 3.8: Block diagram of the one-person output-feedback system. The linear plant converts the control signal, i.e., the reward, to a reward attitude. The disturbance, which arises from the nonlinear plant dynamics, forms the internal attitude. Summing these two attitudes generates the behavioural intent, which is fed back and compared against the desired behaviour, B_d. The resulting error signal is used by the controller, $C[z]$, to update the reward. The control signal, $R[k]$, is given by (3.21).

attitude changes, i.e.,

$$A_{out}[k] = A_o + \sum_{i=0}^{k-1} \Delta A_{out}[i].$$

The summation term can be modelled with a discrete-time transfer function:

$$\frac{A_{out}[z]}{\Delta A_{out}[z]} = \frac{1}{z-1},$$

with the non-zero initial condition $A_{out}[0] = A_o$. Integrating $P[z]$ and the disturbance model into the standard output-feedback configuration yields the closed-loop system given in Figure 3.8. This configuration is termed the *one-person output-feedback system*. Now, the problem essentially becomes a step-tracking problem, where the reference input is a step representing the desired (positive) behavioural intent, $B_d > 0$.

At this point, the one-person system is framed in such a way as to fit into the standard linear output-feedback configuration with one complication: the disturbance, $A_{out}[k]$, depends (nonlinearly) on signals within the closed-loop system. To ensure perfect steady-state tracking in the presence of step-disturbances, the feedforward system must contain one summer, whereas the feedforward system must contain two summers in the case of ramp-disturbances. To determine if $A_{out}[k]$ can be approximated as a step-disturbance or a ramp-disturbance, the effect of the initial conditions on the system behaviour is examined and the following assumption is made:

Assumption 3.1. *The mental processing pole location for dissonance pressure in (2.11) is zero, i.e.,* $r_2 = 0$.

The above assumption is made to simplify the analysis of the output-feedback analysis and thus, is specific to this section. All simulations in this section still use the mental processing pole location for dissonance pressure given in Table 2.2. A direct consequence of Assumption 3.1 is that $P^{CD}[k] = P^{CD}_{raw}[k]$ and thus, from Lemma 2.1, $P^{CD}[k]$ has a maximum magnitude.

From the initial conditions on the attitudes, i.e., $A_{out}[0] = A_o$ and $A_{rew}[0] = 0$, the output, $B[k]$, is initially negative. Suppose the controller produces a reward that is too small to drive behaviour positive. Previous analysis has demonstrated that in this case, attitude change occurs due to dissonance pressure alone, causing $A_{out}[k]$ to decrease. Since in this case, dissonance pressure is the only pressure contributing to attitude change, it follows from Assumption 3.1 that this attitude change is bounded. To see why, notice that since no overjustification pressure arises, the attitude change is simply proportional to the dissonance pressure, which is bounded. In other words, the disturbance experienced in the one-person output-feedback system is bounded by a ramp-function with a slope corresponding to the maximum possible attitude change.

If, on the other hand, the controller produces a reward that is large enough to drive behaviour positive, then initially, the only pressure experienced is dissonance pressure (since $A_o < 0$). Like the previous case, this dissonance pressure produces an internal attitude change, which is bounded by a ramp-function. However, unlike the case of a small reward, the large reward potentially produces enough dissonance pressure to cause $A_{out}[k] \geq 0$, in which case overjustification pressure may arise. Should this situation occur, by Assumption 3.1, $P^{CD}[k] = 0$ and therefore, only overjustification pressure arises. As the overjustification pressure causes $A_{out}[k]$ to decrease but ensures $A_{out}[k] \geq 0$, it follows that $\Delta A^{OJ}_{out}[k]$ is bounded. In conclusion, changes to $A_{out}[k]$ at each sample are bounded and as a result, $A_{out}[k]$ can be bounded by a ramp function.

Before proceeding, a technical detail is addressed. Since $A_{out}[k]$ can both increase and decrease, saying that $A_{out}[k]$ can be bounded by a ramp function is ambiguous. In reality, this claim is composed of two different statements. First, if $A_{out}[k]$ is an increasing function of k, then it can be bounded from above by a ramp function. Alternatively, if $A_{out}[k]$ is a decreasing function of k, then it can be bounded from below by a ramp function. In the case when $A_{out}[k]$ switches from being an increasing function to a decreasing function (or vice versa), it does not make sense for one ramp function to bound $A_{out}[k]$; thus, this switch means the bounding ramp function also switches. From a control perspective, switching the slope of a ramp-disturbance influences the controller's ability to achieve perfect steady-state tracking. A double summer ensures perfect steady-state tracking for a constant ramp but does not guarantee that tracking is achieved if the slope of the ramp changes. Nevertheless, the boundedness of $\Delta A_{out}[k]$ is simply used as motivation for the controller design and therefore, this issue is not considered in more detail.

Since $A_{out}[k]$ is at worst a ramp-function, and from before, perfect steady-state tracking of a linear plant with a ramp-disturbance is possible only when two summers are in the feed-forward system, consider

$$\bar{C}[z] = \frac{K_c}{(z-1)^2}.$$

Before proceeding, two important details are examined: closed-loop stability and the requirement that $A_{rew}[k]$ cannot be negative.

To study the closed-loop stability of the one-person output-feedback system with the above controller, a simplification is made: the disturbance $A_{out}[k]$ is assumed to be exactly a ramp function. In particular, the simplifying assumption reduces the problem to considering the closed-loop stability of the system given in Figure 3.9, called the *approximated one-person output-feedback system* or simply the *approximated output-feedback system*. Closed-loop stability of the approximated output-feedback system is studied using root locus techniques. The root locus plot of $P[z]\bar{C}[z]$ is shown in Figure 3.10. This plot clearly demonstrates that for the given plant-controller combination, it is not possible to stabilize the approximated output-feedback system since any gain $K_c > 0$ is not able to ensure that the closed-loop poles, initially located at $z = 1$, are contained within the unit disk. To provide closed-loop stability of the approximated output-feedback system, a zero is included in the controller; thus, suppose

$$C_{nom}[z] = \frac{K_c\,(z-a)}{(z-1)^2}. \tag{3.20}$$

Finding the range of values for K_c and a that guarantee closed-loop stability of the approximated output-feedback system is performed later in this section (see Theorem 3.9).

The second detail to discuss is the requirement that $A_{rew}[k] \geq 0$. To meet this requirement, a saturator and anti-windup scheme are included in the controller. Note that, $C_{nom}[z] = \frac{R[z]}{E[z]}$, which, using the above equation for $C_{nom}[z]$, can be converted into a difference equation:

$$R[k] = K_c e[k-1] - K_c a e[k-2] + 2R[k-1] - R[k-2].$$

Applying the anti-windup saturator gives the follow control law:

$$R[k] = \max\left\{0, K_c e[k-1] - K_c a e[k-2] + 2R[k-1] - R[k-2]\right\}. \tag{3.21}$$

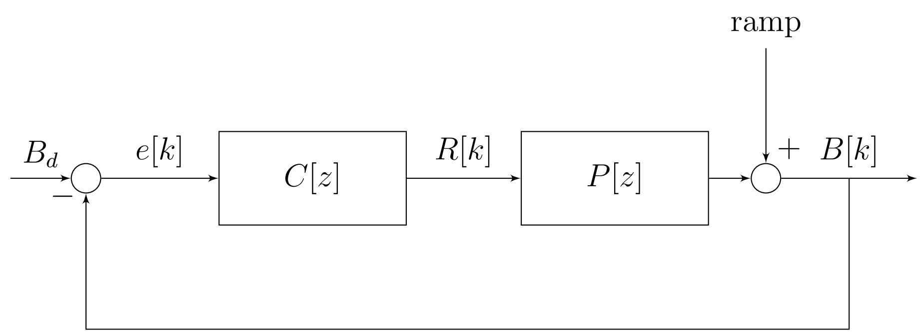

Figure 3.9: Block diagram of the approximated output-feedback system. The plant, $P[z]$, is given by (3.19) and the controller, $C[z]$, is given by (3.20). The ramp disturbance represents the worst possible amount of attitude change that can occur.

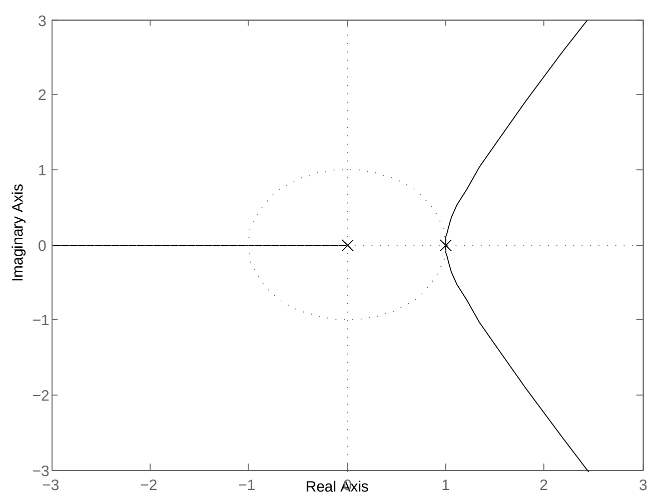

Figure 3.10: The root locus plot of $P[z]\bar{C}[z]$ demonstrates that the closed-loop system of $P[z]\bar{C}[z]$ is unstable for any controller gain. Parameter values used are those given in Table 2.2.

Before detailed analysis is performed, a high level examination of the trends exhibited by the one-person output-feedback system is presented. This high-level overview not only indicates that the controller in (3.21) may be able to meet the control objective of driving $B[k] \geq B_d$ as k tends to infinity, but also provides stepping-stones used in the analysis.

At a high-level, when the controller given in (3.21) is applied to the one-person output-feedback system, the closed-loop system undergoes three stages:

I: For $0 \leq k \leq \hat{k} - 1$, $A_{out}[k] < 0$ and $B[k] \leq 0$

II: For $\hat{k} \leq k \leq \bar{k} - 1$, $A_{out}[k] < 0$ and $B[k] > 0$, and

III: For $k \geq \bar{k}$, $A_{out}[k] \geq 0$ and $B[k] \geq 0$.

The switching from stage I to stage II occurs at $k = \hat{k}$, whereas the switching from stage II to stage III occurs at $k = \bar{k}$.

In stage I, the fact that $B[k]$ is non-positive means the control signal should work to increase $B[k]$, resulting in a positive reward, $R[k] > 0$. If this reward is large enough to produce a positive behaviour, then the system enters stage II. On the other hand, in the case of smaller rewards, the system may continue in stage I, in which case, the positive reward attitude generated by $R[k] > 0$ is inconsistent with the non-positive behaviour, causing dissonance pressure to arise. Note that no overjustification pressure arises and therefore dissonance is the only pressure contributing to attitude change. This attitude change is in the negative direction due to the negative behaviour. Although this situation does not seem promising, the controller is designed to provide steady-state tracking and thus, eventually the controller should generate a $R[k]$ that is sufficiently large to drive behaviour positive, in which case, the system enters stage II. In fact, Lemma 3.5 presents sufficient conditions for ensuring the system enters stage II.

In stage II, the behaviour is positive but the internal attitude is still negative and therefore, no overjustification pressure arises. However, owing to the sign change of $B[k]$, the dissonance pressure becomes positive, but again, is bounded. In other words, the internal attitude increases by at most a bounded rate and therefore, the assumption that $A_{out}[k]$ is a ramp-disturbance still holds and therefore perfect steady-state tracking still seems reasonable. Moreover, since attitude is increasing, it may become positive, thus placing the system in stage III.

In stage III, the fact that $A_{out}[k] \geq 0$, $B[k] \geq 0$ and $A_{rew}[k] \geq 0$ means overjustification pressure arises. Furthermore, by Assumption 3.1, $P^{CD}[k] = 0$ (because $P^{CD}_{raw}[k] = 0$ when all cognitions are positive); thus, overjustification pressure is the only pressure contributing to attitude change. Previous analysis demonstrates that overjustification pressure cannot drive attitude negative and therefore, attitude change is, again, bounded. Furthermore, since $A_{out}[k]$ remains positive, it follows that once the system enters stage III, it remains in stage III; hence, provided the system is closed-loop stable, the controller should achieve perfect steady-state tracking in this stage III.

At a high-level, the controller given in (3.21) seems promising in its ability to meet the control objective. To proceed with the analysis, closed-loop stability of the system is first examined; then, the system characteristics for each of the three aforementioned stages are investigated. Results ensuring the system transitions through these stages to ultimately reach stage III are given throughout. Once the system reaches stage III, four possible output trends are studied. Two of these trends are shown to be impossible, and one is shown to occur when the system is closed-loop unstable. A final conjecture is presented indicating that if the system is closed-loop stable, then upon entering stage III, the controller is able to provide perfect steady-state tracking of the desired behaviour, B_d.

3.2.3 Output-Feedback Controller Analysis

To begin the formal analysis, closed-loop stability is examined in the context of the approximated output-feedback system given in Figure 3.9. Since the controller is designed to reject ramp-like disturbances, the stability of the approximated output-feedback system depends on the characteristic polynomial of the closed-loop system composed of $P[z]$ and $C_{nom}[z]$. The closed-loop stability of the approximated one-person output-feedback system is considered over that of the actual one-person output-feedback system for simplification purposes. The motivation for this simplification is two-fold. The primary reason for this simplification is that if the controller given in (3.20) is unable to provide closed-loop stability for the approximated output-feedback system, then the likelihood that the more restrictive controller given in (3.21) is able to provide this stability is low. Secondly, the approximated output-feedback system assumes a ramp-disturbance, which is considered to be a worst-case scenario for the actual disturbance generated by $A_{out}[k]$. Using a modified Routh-Hurwitz approach, the theorem below gives the range of permissible gains (K_c) and zeros (a) that provide closed-loop stability of approximated output-feedback system.

Theorem 3.9. *For the approximated one-person output-feedback system containing the controller given in (3.20), if for* $a \in (0.5, 1)$*,* K_c *satisfies*

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

then the closed-loop system is stable.

Proof. The proof, found in Appendix D.12, uses the Routh-Hurwitz stability test on a bilinear transformation of the closed-loop system's characteristic polynomial. □

The conditions of Theorem 3.9 are sufficient to ensure closed-loop stability of the approximated output-feedback system. To determine if the controller given by (3.21) is able to meet the control objective, the dynamic performance of the system is studied through consideration of trends exhibited in each of the three aforementioned stages. In the first stage, both the behaviour and the internal attitude are negative. Since the goal is to meet a desired, positive behaviour, the controller must necessarily drive $B[k] > 0$. Suppose $B[k] > 0$ at $k = \hat{k}$ while for $0 \leq k < \hat{k}$, $B[k] \leq 0$. As previously discussed, the reward provided by the control signal either causes $B[k]$ to increase (if the reward is large enough to drive behaviour positive) or decrease (if the reward is too small). To guarantee $B[k]$ eventually switches from being negative, it is enough to find conditions to ensure that $B[k]$ becomes an increasing function of k until (at least) $B[k] > 0$. To this end, define $T^* < \hat{k}$ to be the first sample at which $B[T^* + 1] > B[T^*]$. The lemma below states T^* exists and gives sufficient conditions to ensure $B[k]$ is an increasing function of k for at least the interval $T^* \leq k \leq \hat{k}$.

Lemma 3.5. *If the one-person output-feedback system with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions, then T^* exists. Moreover, if*

$$B_d > \frac{K_1}{2\mu_1 K_c}, \tag{3.22}$$

then $B[k]$ is an increasing function of k for at least the interval $T^* \leq k \leq \hat{k}$, eventually becoming positive at $k = \hat{k}$.

Proof. See Appendix D.13. □

Lemma 3.5 states that as long as the desired behaviour, B_d, is large enough, then once $B[k]$ begins increasing, it remains an increasing function of k at least until $B[k]$ becomes positive. Since $P^{CD}[k] < 0$ for $0 \leq k < \hat{k}$ (due to the sign of $B[k]$ over this interval), $A_{out}[k] < 0$ for $0 \leq k \leq \hat{k}$. Therefore, $B[k]$ necessarily becomes positive before $A_{out}[k]$ can become positive. Consequently, provided B_d is large enough, the system is guaranteed to transition from stage I to stage II (and not to stage III). The simulation results shown in Figure 3.11 confirm this result. Figure 3.11 contains simulation results of three different situations. The left-hand side plots show the behaviour and the internal attitude for a simulation using the strategy of Lemma 3.5. In this simulation, the condition given in (3.22) is met and $B[k]$ not only starts increasing immediately, but continues increasing until after $B[k] > 0$. However, the sufficient condition given by Lemma 3.5 is conservative because it is derived assuming the maximum amount of dissonance pressure occurs at each sample, which is generally not the case. The simulation results shown in the middle plots of Figure 3.11 suggest that it is still possible for $B[k]$ to remain an increasing function when the condition given by Lemma 3.5 is not met. For this simulation, $B_d = 40$ *attitude units*,

$K_1 = 30$, $K_c = 0.3$ and $\mu_1 = 1$; thus, (3.22) is not met. Nevertheless, $B[k]$ not only begins to increase immediately, but $B[k]$ continues increasing until after it becomes positive.

Not only does Lemma 3.5 provide a conservative condition for guaranteeing $B[k]$ becomes positive, the strategy of ensuring $B[k]$ remains an increasing function is also conservative. In particular, it is still possible for the system to enter stage II in the case when $B[k]$ does not remain an increasing function after the first sample at which it increases. The simulation results shown in the right-hand side plots of Figure 3.11 support this claim. For this simulation, the condition given in (3.22) is not met. Moreover, even though $B[k]$ is an increasing function for $0 \leq k \leq 5$, it becomes a decreasing function of k for $5 < k \leq 11$ before again becoming an increasing function and, at $k = 26$, becoming positive. Therefore, the strategy to ensure $B[k]$ remains an increasing function once it begins increasing is not necessary for ensuring the system enters stage II.

Nevertheless, Lemma 3.5 provides conditions that guarantee the system enters stage II; thus, suppose the system enters stage II at $k = \hat{k}$, i.e., $B[k] > 0$ at $k = \hat{k}$ but $B[k] \leq 0$ for $0 \leq k < \hat{k}$. In the ideal situation, $B[k] > 0$ for $k \geq \hat{k}$. However, simulation results demonstrate that it is possible for the system to switch back into stage I. In fact, the closed-loop system can display three possible trends, shown in Figure 3.12. The first trend, displayed in the top plot in Figure 3.12, is the ideal situation in which, upon switching positive, behaviour remains positive. Once $B[k]$ becomes positive, it remains positive. The middle plot of Figure 3.12 shows the second trend, in which the behaviour eventually remains positive. This trend is acceptable, although the switching characteristics of $B[k]$ are undesirable due to additional complications arising in the analysis. The last trend, displayed in the bottom plot of Figure 3.12, demonstrates that the behaviour could remain oscillating between positive and negative values as k tends to infinity. Extensive simulations suggest that this final trend occurs only for smaller controller gains and small values for B_d or if the controller parameters do not satisfy the conditions of Theorem 3.9.

Finding conditions to guarantee $B[k] > 0$ for $k \geq \hat{k}$ is quite difficult. However, Lemma 3.5 suggests B_d err on the side of being large to ensure the system enters stage II, and since simulation results indicate larger values of B_d lead to less switching between stage I and stage II, it is reasonable to make the following assumption:

Assumption 3.2. *The controller given in (3.21) can be tuned such that upon entering stage II, the one-person output-feedback system never returns to stage I, i.e.,* $B[k] > 0$ *for* $\hat{k} \leq k < \bar{k}$ *and* $B[k] \geq 0$ *for* $k \geq \bar{k}$.

By Assumption 3.2, the system either remains in stage II or enters stage III. Recall that to enter stage III, $A_{out}[k]$ must become non-negative. Due to being in stage II, the behaviour is positive and therefore, the dissonance pressure that arises causes $A_{out}[k]$ to increase. This increase could drive $A_{out}[k] \geq 0$, or could potentially lead to attitude

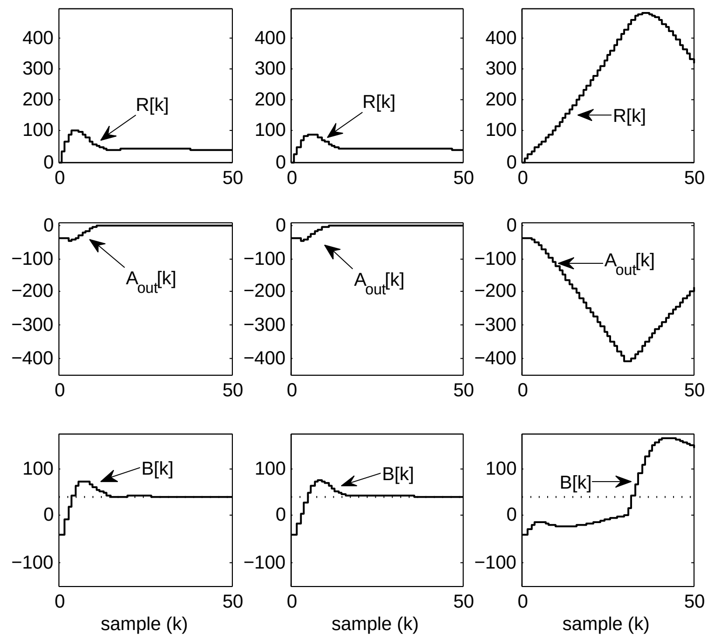

Figure 3.11: Simulation results demonstrating that the condition given in Lemma 3.5 is conservative. The one-person output-feedback system is simulated with the controller given in (3.21). The desired behavioural intent is set to $B_d = 40$ *attitude units* (dashed) and the controller zero is set to $a = 0.959$. The one-person system has the initial conditions given in (2.16) with $A_o = -40$ *attitude units* and parameter values given in Table 2.2. The simulation is performed for three different controller gains: $K_c = 0.4$ (left), $K_c = 0.3$ (middle) and $K_c = 0.15$ (right). The left simulation uses a controller that meets the sufficient condition of Lemma 3.5 and demonstrates that $B[k]$ remains an increasing function in stage I, as predicted by Lemma 3.5. The middle simulation demonstrates that the condition in Lemma 3.5 is conservative, as the parameter values used for this simulation do not meet the condition given in Lemma 3.5, yet $B[k]$ remains an increasing function of k. The right simulation demonstrates that the overall strategy of $B[k]$ remaining an increasing function is a conservative approach, since in the simulation $B[k]$ does not remain increasing function but still becomes positive.

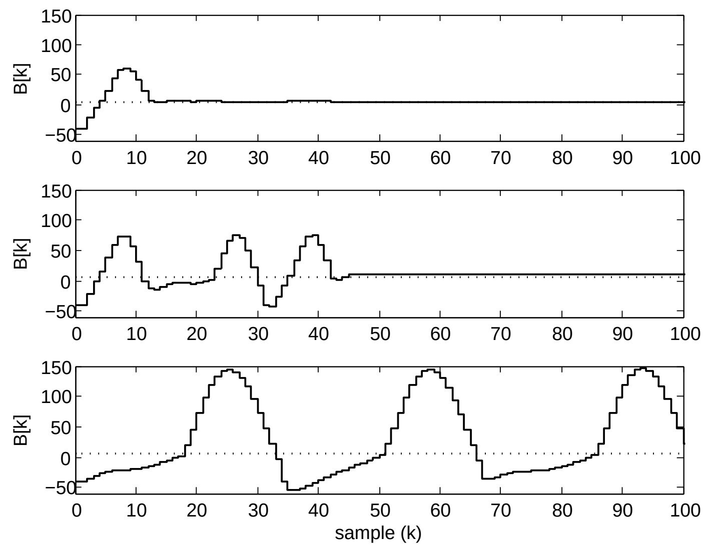

Figure 3.12: Simulation results indicating three possible trends for $B[k]$ when the system enters stage II. The one-person output-feedback system is simulated with the controller given in (3.21). The desired behavioural intent is set to $B_d = 5$ *attitude units* (dotted). The one-person system has the initial conditions given in (2.16) with $A_o = -40$ *attitude units* and parameter values given in Table 2.2. The simulation is performed for three controller parameter value sets: $K_c = 0.4$, $a = 0.75$ (top), $K_c = 0.4$, $a = 0.6$ (middle) and $K_c = 0.1$, $a = 0.6$ (bottom). The first controller demonstrates that once $B[k] > 0$, it remains positive. The second controller demonstrates that even though $B[k]$ switches between positive and negative values, the controller is eventually able to ensure $B[k] > 0$ after some specific sample. The final controller shows that it is possible for a controller that meets the conditions of Theorem 3.9 to cause $B[k]$ to never remain exclusively positive.

converging to a negative constant as k tends to infinity. The following lemma states that if $A_{out}[k] < 0$ for all k, then $A_{out}[k]$ converges to zero.

Lemma 3.6. *If the one-person output-feedback system, with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions, and, for all* $k \geq \hat{k}$*,*

(i) $B[k] > 0$*, and*

(ii) $A_{out}[k] < 0$*;*

then $A_{out}[k] \to 0$ *as* $k \to \infty$*.*

Proof. See Appendix D.14. □

Although Lemma 3.6 does not guarantee $A_{out}[k]$ becomes non-negative, thus ensuring the system enters stage III, it is later shown that if $A_{out}[k] \geq 0$ at some $k = \bar{k}$, then $A_{out}[k]$ still tends to zero as k tends to infinity (see Lemma 3.9). As a result, the characteristics of the system output, $B[k]$, as $k \to \infty$ are the same in the case when $A_{out}[k] < 0$ for all $k \geq 0$ as when $A_{out}[k]$ becomes non-negative at $k = \bar{k}$. Thus, the analysis continues by investigating the system characteristics for the case when $A_{out}[k] \geq 0$.

Suppose the system enters stage III at $k = \bar{k}$, i.e., $B[\bar{k}] \geq 0$ and $A_{out}[\bar{k}] \geq 0$ (but $A_{out}[k] < 0$ for $0 \leq k < \bar{k}$); then, the system remains in stage III for $k \geq \bar{k}$ because $A_{rew}[k] \geq 0$ (due to the controller equation given in (3.21)) and $A_{out}[k] \geq 0$ for $k \geq \bar{k}$. The latter claim holds from the fact that the overjustification effect cannot decrease the internal attitude to a negative value. Furthermore, since $P^{CD}_{raw}[0] = 0$ at $k = \bar{k}$, it follows from Assumption 3.1 that $P^{CD}[k] = 0$ at $k = \bar{k}$ and therefore, overjustification pressure is the only contributor to attitude change, i.e., $A_{out}[k] \geq 0$ for $k \geq \bar{k}$. Moreover, as behaviour is the sum of $A_{rew}[k]$ and $A_{out}[k]$, both of which are non-negative for $k \geq \bar{k}$, $B[k] \geq 0$ also and therefore, the system necessarily remains in stage III for $k \geq \bar{k}$.

Now, it only remains to determine whether or not the controller given in (3.21) is able to drive $B[k] \geq B_d$ as $k \to \infty$. At this point, $B[k]$ could display four possible trends as k tends to infinity.

- $B[k]$ remains below B_d in the sense that there exists an $\epsilon > 0$ and N such that, for all $k > N$, $B[k] < B_d - \epsilon$.
- $B[k]$ remains above B_d in the sense that there exists an $\epsilon > 0$ and N such that, for all $k > N$, $B[k] > B_d + \epsilon$.
- $B[k]$ converges to B_d.

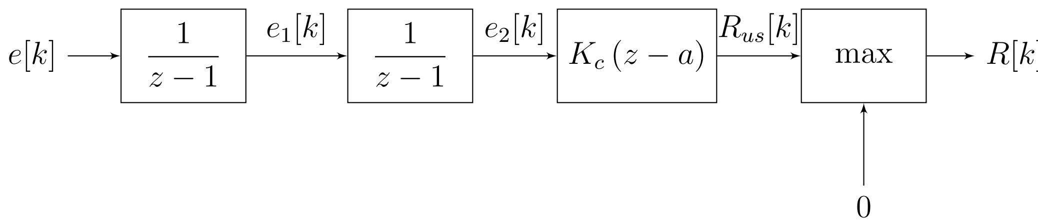

Figure 3.13: The controller given in (3.21) is expanded into several first-order components. Intermediate signals are given names to aid the analysis. The input to the controller is given by $e[k] = B_d - B[k]$ and the output of the controller is the reward signal, $R[k]$. This expanded block-diagram is used in Lemmas 3.7 and 3.10 to show that two of the four possible characteristics for $B[k]$ are impossible.

- $B[k]$ oscillates around B_d without converging, in the sense that there exists an $\epsilon > 0$ such that, for every N, there is a $k_1 > N$ and a $k_2 > N$ such that $B[k_1] > B_d + \epsilon$ and $B[k_2] < B_d - \epsilon$.

The trend of remaining below B_d is considered first. To aid the analysis, the controller is separated into various components and intermediate signals as shown in Figure 3.13. At a high level, this figure helps explain why, provided a mild technical condition holds, the system cannot exhibit the first trend, as stated in the lemma below.

Lemma 3.7. *For all $B_d > \frac{K_1}{2}$, if the one-person output-feedback system, with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions and enters stage III at $k = \bar{k}$, then, for all $\epsilon > 0$, there does not exist a $\tilde{k}$ such that for all $k \geq \tilde{k}$, $B[k] \leq B_d - \epsilon$.*

Proof. See Appendix D.15. □

Lemma 3.7 says that the one-person output-feedback behaviour signal cannot remain less than B_d as k tends to infinity provided $B_d > \frac{K_1}{2}$. To understand why this technical condition exists, consider the state-feedback controllers in the previous section. For relatively small values of B_d, the controller forces $R[k] = 0$, whereas larger B_d values require positive reward, as the internal attitude cannot alone produce the desired behaviour. In the same vein, the technical condition in Lemma 3.7 simplifies the analysis by forcing the control signal to always be positive. As a result, not only is a positive reward necessary for achieving the control objective, $R[k]$ cannot converge to zero as k tends to infinity, as stated in the lemma below.

Lemma 3.8. *For all $B_d > \frac{K_1}{2}$, if the one-person output-feedback system, with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions and enters stage III at $k = \bar{k}$, then $R[k]$ cannot converge to zero as k tends to infinity.*

Proof. See Appendix D.16. □

Not only does Lemma 3.8 conclude that, as $k \to \infty$, a reward is necessary for achieving the control objective, it can be used to show two more conclusions relevant to the system's characteristics. First, the fact that $R[k]$ cannot converge to zero as $k \to \infty$ helps show that $A_{out}[k]$ necessarily converges to zero as k tends to infinity as stated in the following lemma.

Lemma 3.9. *For all $B_d > \frac{K_1}{2}$, if the one-person output-feedback system, with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions and enters stage III at $k = \bar{k}$, then*

$$A_{out}[k] \to 0 \text{ as } k \to \infty.$$

Proof. See Appendix D.17. □

By combining the results of Lemmas 3.6 and 3.9, we conclude that if the system enters stage II, then the internal attitude converges to zero as k tends to infinity and, therefore, the trends exhibited by $B[k]$ as $k \to \infty$ are the same in the case when the system remains in stage II as in the case when the system reaches stage III.

The second conclusion arising from Lemma 3.8 is Lemma 3.10 below, which says that the second of the four aforementioned trends for $B[k]$ is impossible.

Lemma 3.10. *For all $B_d > \frac{K_1}{2}$, if the one-person output-feedback system, with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions and enters stage III at $k = \bar{k}$, then, for any $\epsilon > 0$, there does not exists a $\tilde{k}$ such that for all $k \geq \tilde{k}$, $B[k] \geq B_d + \epsilon$.*

Proof. See Appendix D.18. □

The results of Lemmas 3.7 and 3.10 are conditional on the system eventually entering stage III. However, Lemmas 3.6 and 3.9 show that in both the case when the system remains in stage II and when the system reaches (and remains in) stage III, $A_{out}[k]$ converges to zero as k tends to infinity. Consequently, the trends exhibited by $B[k]$ as $k \to \infty$ are the same in both cases. The theorem below summarizes these results.

Theorem 3.10. *For all $B_d > \frac{K_1}{2}$, if the one-person output-feedback system, with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions and enters stage II at $k = \hat{k}$, then, for any $\epsilon > 0$, there does not exist a $\tilde{k}$ such that*

(a) $B[k] \geq B_d + \epsilon$ *for all* $k \geq \tilde{k}$, *or*

(b) $B[k] \leq B_d - \epsilon$ *for all* $k \geq \tilde{k}$.

Proof. The theorem follows immediately from Lemmas 3.6, 3.7, 3.9 and 3.10. □

There are now two remaining possible trends that could be exhibited by $B[k]$: oscillation around B_d or convergence to B_d as k tends to infinity. We first consider the case when $B[k]$ oscillates.

Consider the parameters a and K_c, which are common to the controller of both the approximated and actual output-feedback systems ((3.20) and (3.21) respectively). Given that (3.21) is more restrictive than (3.20) (because of the additional saturator and anti-windup scheme in (3.21)), it is reasonable to suggest that if a and K_c do not provide stability for the approximated system, these same parameters are unlikely to provide stability for the actual system. Since Theorem 3.10 does not assume the system is closed-loop stable, two of the three potentially unstable trends have been eliminated, i.e., $B[k]$ cannot tend to positive or negative infinity as $k \to \infty$ because $B[k]$ cannot remain above or below B_d as k tends to infinity. Thus, the remaining unstable trend, oscillation around B_d, is likely to occur only if the conditions of Theorem 3.9 fail to hold, whereas the desirable outcome of $B[k]$ converging to B_d as k tends to infinity is likely to occur only if the conditions of Theorem 3.9 hold. This hypothesis is summarized in the conjecture below.

Conjecture 3.1. *For all $B_d > \frac{K_1}{2}$, suppose the one-person output-feedback system, with initial conditions given in (2.16), contains the controller given in (3.21) with zero initial conditions and enters stage II at $k = \hat{k}$. If the controller parameters do not satisfy the conditions of Theorem 3.9, then $B[k]$ oscillates around B_d as $k \to \infty$. On the other hand, if the controller parameters satisfy the conditions of Theorem 3.9, then $B[k]$ tends to B_d as $k \to \infty$.*

Figure 3.14 below presents simulations supporting Conjecture 3.1 and Theorem 3.10. In these simulations, the controller is tuned to ensure Assumption 3.2 is met. Both simulation results are consistent with Theorem 3.10 because both demonstrate that $B[k]$ does not remain above or below B_d as k tends to infinity. In the left simulation, the controller parameters used satisfy the conditions of Theorem 3.9. Since these simulation results show $B[k]$ converges to B_d as k tends to infinity, they provide evidence to support Conjecture 3.1. Furthermore, in the simulation shown on the right-hand side of Figure 3.14, the controller

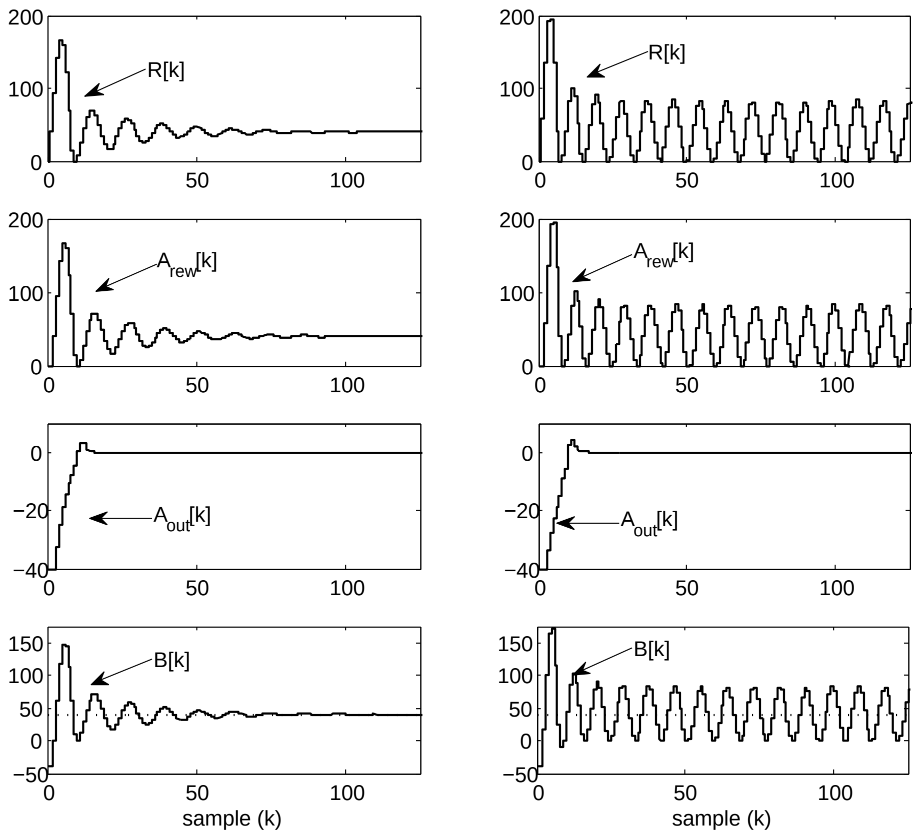

Figure 3.14: Simulation of the one-person output-feedback system with initial conditions given in (2.16), controller given in (3.21) and parameter values given in Table 2.2. Both simulations use an initial attitude $A_o = -40$ *attitude units*, desired behaviour $B_d = 40$ *attitude units*, and controller zero, $a = 0.65$. Using these parameter values, the approximated output-feedback system is stable when $K_c < 0.71$. In the left simulation, $K_c = 0.5$. The system enters stage III at $\bar{k} = 10$ because $B[k] \geq 0$ and $A_{out}[k] \geq 0$ for $k \geq 10$. As Conjecture 3.1 hypothesizes, as k tends to infinity, $B[k]$ converges to B_d (denoted by the dotted line). The right simulation provides additional support for Conjecture 3.1. In this simulation $K_c = 0.72$, and, therefore, the conditions of Theorem 3.9 are not met. The system enters stage III at $\bar{k} = 10$, and the output, $B[k]$, oscillates around the desired behaviour, B_d, again denoted by the dotted line. Both simulations demonstrate the results of Theorem 3.10 by showing $B[k]$ cannot remain above or below B_d as k tends to infinity.

parameters do not meet the conditions of Theorem 3.9, and the output signal, $B[k]$, is shown to oscillate around B_d as k tends to infinity, thus supporting Conjecture 3.1.

Although showing that the controller given in (3.21) can meet the control objective for any set of initial conditions is currently an open problem, there is evidence to suggest that this is the case. Note that the approximation on the nonlinear disturbance in Figure 3.9 still holds for any set of initial conditions. To see why, recall that this disturbance is the internal attitude, $A_{out}[k]$. To show that changes to $A_{out}[k]$ are bounded, we argue that $\Delta A_{out}[k]$ is bounded. From (2.5), $\Delta A_{out}[k]$ is the sum of two signals, $\Delta A_{out}^{CD}[k]$ and $\Delta A_{out}^{OJ}[k]$. By Lemma 2.1, $-0.5 \leq P_{raw}^{CD}[k] \leq 0.5$ and therefore, $P_{raw}^{CD}[k]$ is bounded. Since $P^{CD}[k] = r_2 P^{CD}[k-1] + (1-r_2)\, P_{raw}^{CD}[k]$ is BIBO stable, it follows that $P^{CD}[k]$ is bounded. As a result, $\Delta A_{out}^{CD}[k]$ is also bounded. Next, consider $\Delta A_{out}^{OJ}[k]$. If raw overjustification pressure exists at times $k \geq \ddot{k}$, then $A_{out}[k]$, $A_{rew}[k]$ and $B[k]$ are all positive for $k \geq \ddot{k}$. This result implies that $P_{raw}^{CD}[k] = 0$ for $k \geq \ddot{k}$, meaning $P^{CD}[k]$ tends to zero as k tends to infinity. Also, due to the signs in (2.13)–(2.15), $\Delta A_{out}^{OJ}[k] \leq 0$. As a result, $A_{out}[k]$ is bounded above for $k \geq \ddot{k}$. Finally, from (2.15), $A_{out}[k] \geq 0$ for $k \geq \ddot{k}$ and thus, $A_{out}[k]$ is bounded. Therefore, the nonlinear disturbance can be approximated with a ramp for any set of initial conditions; hence, the following conjecture is made.

Conjecture 3.2. *For all $B_d > 0$, if the one-person output-feedback system contains the controller given in (3.21), then the controller parameter values can be selected to ensure*

$$B[k] \rightarrow B_d \text{ as } k \rightarrow \infty.$$

To provide further evidence of Conjecture 3.2, a set of 200 simulations is performed. For each of the 200 simulations, random values are given to B_d, the initial conditions (both of the model and the controller), and the model parameters (according to the restrictions in Table 2.2). Then the controller parameters are chosen to meet the condition given in Theorem 3.9. In each simulation, the output, $B[k]$, converges to B_d as $k \rightarrow \infty$. Figure 3.15 shows the simulation results for one of the tests.

To summarize, this chapter studied various open-loop and closed-loop controllers to meet the control objective of driving $B[k] \geq B_d$ as k tends to infinity. Two open-loop controllers were considered: an impulse reward and a step reward. The impulse-reward controller was shown to be able to increase $B[k]$ but is unable to meet the control objective for all combinations of A_o, B_d and K_1, unlike the step-reward controller. Although the step-reward controller is able to meet the control objective, the undesirable consequence of the reward on the internal attitude motivated studying closed-loop controllers. The first, a state-feedback controller, was the extended-impulse-reward controller, where the reward signal is removed immediately before the internal attitude becomes non-negative. Although this controller is unable to always satisfy the goal of driving $B[k] \geq B_d$ as k tends to infinity, its use of state-feedback motivated a more general state-feedback controller.

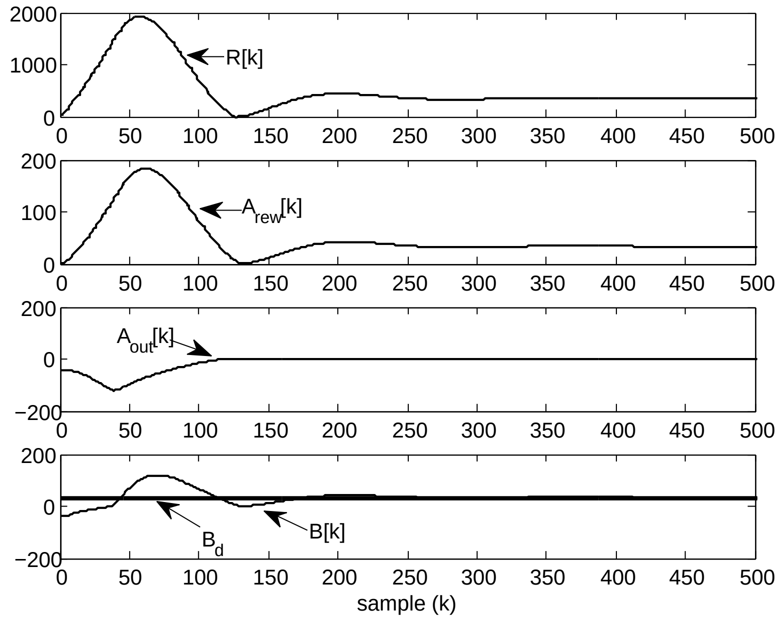

Figure 3.15: Simulation of the one-person output-feedback system demonstrating that the controller given in (3.21) can be tuned to ensure $B[k] \to B_d$ as $k \to \infty$, thus supporting the claim of Conjecture 3.2. For this simulation, $B_d = 33.8132$ and the following values were used for the model parameters: $\mu_1 = 0.0959$, $r_1 = 0.7475$, $r_2 = 0.7485$, $r_3 = 0.5433$, $K_1 = 6.5785$, $K_2 = 0.0326$. The following values were used for the controller parameters: $K_c = 0.4$, $a = 0.959$. The following initial conditions were used: $R[-2] = 13.3925$, $R[-1] = 1.9151$, $e[-2] = 1.6392$, $e[-1] = -2.8520$, $A_{out}[0] = -41.6167$, $A_{rew}[0] = 2.7629$, $B[0] = -15.19935$, $P^{CD}[0] = -0.01565$, $P^{OJ}[0] = 0$.

This second state-feedback controller is not only able to meet the goal as $k \to \infty$, it is able to meet the goal immediately, for $k \geq 1$, while also maintaining internal stability. However, due to the impractical nature of state-feedback for the given application, an output feedback controller was then considered. Since the one-person output-feedback system can be framed as a linear output-feedback system with a disturbance, linear control techniques were used to design this final controller. In particular, the disturbance was approximated to be a ramp. The output-feedback controller was then designed using linear control techniques to attain perfect steady-state tracking and perfect steady-state ramp-disturbance rejection. The analysis given showed that the system passes through three stages, eventually reaching a stage with only two possible outcomes. Supported by extensive simulation results, a final conjecture hypothesizes that $B[k]$ necessarily converges to B_d as $k \to \infty$ if the approximated system is stable, whereas $B[k]$ necessarily oscillates around B_d if the approximated system is unstable. The results presented in this chapter are not only useful for control of the one-person system, but also motivate design approaches taken in Chapter 5 when control of the two-person system is investigated.

Chapter 4

Two-Person System: Psychology and Modelling

In the one-person system, a person with an initially negative attitude is offered a reward in an attempt to overtly control the person's behaviour. This person does not experience any other external influences. This assumption is idealistic since an individual may experience social influences from a variety of sources including friends, peers, society and media; hence, this assumption is now relaxed. In particular, we now assume that, in addition to the external influence of a reward, an individual experiences a social influence from the behaviour of another person. In the context of the piano situation, the child's friend represents this additional person. The friend's behaviour influences the child's attitude and behaviour and, likewise, the child's behaviour influences his friend. Thus, the two-person system consists of attitude-behaviour models of two individuals, connected in some way through the behaviour of these two individuals.

To model this connection and its effect on attitude and behaviour, additional psychology background is required. From the theory of planned behaviour, three elements combine to generate a behaviour, one of which is external influences. The one-person system deals with only one external influence, rewards. However, the theory of planned behaviour extends beyond rewards and models how two general categories of external influences affect an individual's behaviour. These categories, termed *social pressure* and *conformity pressure*, are explained in Sections 4.1.1 and 4.1.2 respectively. Provided in these explanations is a discussion of how these pressures relate to cognitive dissonance theory and the overjustification effect. Following this discussion, a two-person discrete-time dynamic model is derived using the framework given in Chapter 2. The model is then verified to be qualitatively consistent with the psychology. Finally, other psychological phenomena arising from group dynamics are introduced and discussed within the context of the two-person system.

4.1 Psychology Describing the Influence of Other People

In Chapter 2, the effect of external influences is not fully captured. For the one-person system, the external influence is simply a reward. However, the theory of planned behaviour extends beyond the effect rewards have on an individual's behaviour. The theory captures the effects of two general categories of external influences, which for the thesis are termed social pressure and conformity pressure. These two influences, studied, among others, by Ajzen in [11], are defined below:

1. **Social pressure:** Influence arising from the perception of "what should or ought to be done with respect to performing a given behaviour."

2. **Conformity pressure:** Influence arising from the perception "that others are or are not performing the behaviour in question."

The external influence in the one-person system, a reward, is a special case of social pressure, as the presence of a reward is accompanied by the expectation of performing a given behaviour. The two-person system contains external influences from both of these categories; hence, each are now discussed in more detail.

4.1.1 Social Pressure

Social pressure refers to pressures arising from perceptions of the expectations other people have of an individual carrying out a particular behaviour. For example, a child may perceive that other people, such as his teachers, friends and/or parents, want him to play the piano. If the child perceives that his teachers strongly want him to play the piano, then the social pressure experienced by the child, denoted $P^{SP}[k]$, is high, i.e., $P^{SP}[k] \gg 0$. On the other hand, if the child perceives that his teachers strongly want him to not play the piano, then $P^{SP}[k] \ll 0$. The sign of $P^{SP}[k]$ indicates whether the expectation is to play or to not play the piano, whereas the magnitude $|P^{SP}[k]|$ indicates the social pressure strength. Note that $P^{SP}[k]$ is not the actual expectation of another person, but instead, is an individual's perception of this expectation.

Research suggests that even though social pressure may arise, it does not necessarily influence the behaviour of an individual. A second component is thus needed to transform this social pressure into an attitude and, consequently, into behaviour. This component reflects the value an individual places on the wishes of the person exerting the social pressure and is denoted in our model by μ_2. For example, suppose the child experiences strong social pressure from his teachers. If he does not care about the wishes of his teachers, then the influence this social pressure has on the child's attitude is small, whereas this influence is larger in the case when the child cares about the expectations of his teachers. The value placed on the wishes of the individual exerting this pressure combines with the social pressure, $P^{SP}[k]$, through a product, forming the attitude towards a behavioural outcome that arises from social pressure due to other people, denoted $A_{others}[k]$. If more than one person is applying social pressure, then $A_{others}[k]$ is the sum of the attitudes arising from the social pressure applied by each individual. That is, the social pressure exerted by person n, $P_n^{SP}[k]$, and the degree to which the individual cares about the expectations of person n, μ_{2n}, combine over N people to form $A_{others}[k]$ as follows:

$$A_{others}[k] = \sum_{n=1}^{N} P_n^{SP}[k]\mu_{2n}.$$

The above sum excludes social pressure that arises from rewards because our model already has a specific mechanism in place to generate a reward attitude. This mechanism is consistent with the above formation of $A_{others}[k]$ (see (2.3)). The two attitudes are differentiated because, later, rewards are the only social pressure considered in the two-person model and $A_{rew}[k]$ provides intuitive notation.

Not only does social pressure directly influence behaviour through the theory of planned behaviour, the related attitude, $A_{others}[k]$, is a cognition, much like $A_{out}[k]$ and $A_{rew}[k]$. As such, it forms a cognitive pair with the generative cognition, $B[k]$, and potentially contributes to dissonance pressure and overjustification pressure. From the psychology literature, overjustification pressure arises in a very specific paradigm: it assumes a reward is used to control the behaviour of an individual with a pre-existing positive attitude towards the behaviour, i.e., $A_{out}[k] > 0$, $A_{rew}[k] > 0$ and $B[k] > 0$. Since rewards are a special case of social pressure, the following assumption is made to extend the overjustification effect psychology to the more general social pressure:

Assumption 4.1. *Any type of social pressure can produce overjustification pressure.*

Following the way in which $A_{rew}[k]$ leads to overjustification pressure, if all cognitive pairs are consistent, we model that if $A_{others}[k] = 0$ then no overjustification pressure arises; on the other hand, if $A_{others}[k] > 0$, then overjustification pressure arises because the child attributes his behaviour to being overtly controlled by other people as opposed to his own internal attitude. If any cognitive pair is inconsistent, then overjustification pressure does not arise and, instead, dissonance pressure arises. In this case, the amount of dissonance pressure experienced is formed as described in Section 2.2.1, i.e.,

$$P = \frac{\sum M_{incon}}{\sum M_{incon} + \sum M_{con}}.$$

To summarize, one type of external influence on an individual's behaviour is social pressure, which forms a related attitude. In the case of rewards, this attitude is given by $A_{rew}[k]$, whereas for more general social pressure, given by $P^{SP}[k]$, the related attitude is given by $A_{others}[k]$. Both $A_{rew}[k]$ and $A_{others}[k]$ form cognitive pairs with $B[k]$ and can contribute to dissonance and overjustification pressures. It is through the reward that social pressure is considered in this thesis. The effect of $A_{others}[k]$ on the internal attitude and the behaviour of an individual is an item for future, related work and is included in the model for completeness.

4.1.2 Conformity Pressure

If other, significant people are performing the behaviour, then an individual is more likely to also perform the behaviour. This is the well-known psychological theory of conformity, which Myers and Spencer, in [14], describe as the "change of behaviour or belief to accord with others." This description raises an interesting observation: conformity can influence both the behaviour and the attitude of an individual. Indeed, psychology literature separates conformity into two cases [14]:

1. **Compliance:** When an individual carries out the given behaviour because other people carry out the behaviour, while privately disagreeing. That is, behaviour changes but not attitude.

2. **Acceptance:** When an individual carries out the given behaviour because other people carry out the behaviour, and this individual changes his internal attitude to match. That is, both attitude and behaviour change.

Both compliance and acceptance involve change in behaviour, and the theory of planned behaviour offers one mechanism through which such change can occur; thus, we use this theory as the way of modelling behaviour change arising from conformity pressures.[1]

From the theory of planned behaviour, when conformity pressure is experienced, a conformity attitude is formed, denoted $A_{conf}[k]$. Similar to the other attitudes in our model, the conformity attitude consists of two components. The first component is the behaviour of a second person. For our model, the behaviours of the two people are denoted $B_1[k]$ and $B_2[k]$, where the subscripts 1 and 2 differentiate between the person being offered a reward and the person not being offered a reward. In the context of the piano example, the child is person one and his friend is person two. The conformity pressure

[1] The original formulation of Ajzen's theory of planned behaviour did not consider the effect of other people's behaviour. However, in [11], Ajzen states that there is plenty of evidence to suggest that this influence should be included in his model.

experienced by the child depends on the behaviour of his friend, $B_2[k]$, and vice versa. The second component of conformity attitude is the extent to which a person can be influenced by another and is denoted μ_3. For our model, there are two people and μ_{32} represents how much the behaviour of person two influences the conformity attitude of person one, whereas μ_{31} represents the extent to which person one's behaviour influences person two's conformity attitude. These two components combine to generate a conformity attitude and, although psychologists have yet to determine this exact relationship, in [11] Ajzen proposes the following model for the conformity attitude arising from conformity pressure exerted by N other people:

$$A_{conf}[k] = \sum_{i=1}^{N} \mu_{3i} B_i[k].$$

In our model, there are two people and thus, each person experiences conformity pressure from one other person, i.e., $N = 1$. For each person in the model, the conformity attitude, $A_{conf,i}[k]$, then combines (through the theory of planned behaviour) with the other attitudes, $A_{out,i}[k]$, $A_{rew,i}[k]$ and $A_{others,i}[k]$ to produce a behaviour, $B_i[k]$. Moreover, as $A_{conf,i}[k]$ is another cognition in the two-person model, it forms a cognitive pair with $B_i[k]$, which is either consistent or inconsistent. If this (or any) cognitive pair in the two-person system is inconsistent, then dissonance pressure arises and is calculated in the usual manner.

In cases where acceptance occurs, research in the field of conformity suggests that behaviour changes tend to precede attitude change [14]. To model this attitude change, as well as to ensure consistency in the order in which behaviour and attitude change, cognitive dissonance theory can be used. In [9], Festinger discusses dissonance pressure arising from social situations and argues that the behaviour of others can produce dissonance pressure, which can be reduced through (internal) attitude change. Since our model includes the effects of dissonance pressure, it already contains a mechanism through which acceptance can occur.

The final point to consider is the effect of $A_{conf,i}[k]$ on the overjustification pressure. Unlike the social pressures $A_{rew,i}[k]$ and $A_{others,i}[k]$, an individual does not feel overtly controlled through conformity pressure. Since $A_{conf,i}[k]$ is not controlling in nature, the following psychological assumption is made:

Assumption 4.2. *Overjustification pressure cannot arise due to conformity pressure.*

Note that, in particular, if the cognitive pair $(A_{conf,i}[k], B_i[k])$ is consistent and no other inconsistencies arise, then $A_{conf,i}[k]$ does not contribute to overjustification pressure.

To summarize, conformity pressure arises from one individual observing the behaviour of another. The compliance element of conformity describes how an individual's behaviour is directly influenced by that of another person. In the context of the piano situation, the

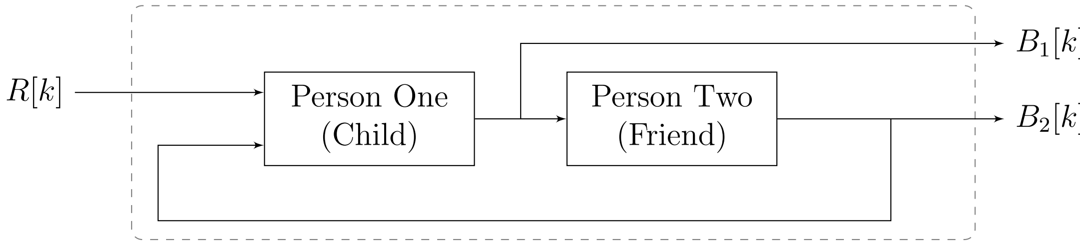

Figure 4.1: The overall psychological system of the two-person model (dotted box), decomposed into attitude-behaviour models of each person, each of which is an augmented one-person system, updated to include the effects of other external influences, $A_{conf}[k]$ and $A_{others}[k]$. The system input is $R[k]$ and the system outputs are $B_1[k]$ and $B_2[k]$.

behaviour of the child's friend and the extent to which this friend can influence the child combine to form a conformity attitude, $A_{conf}[k]$. This conformity attitude has a direct influence on the child's behaviour. Furthermore, the acceptance element of conformity, which says that an individual's attitude may also change, is modelled by the existing dissonance pressure mechanism. If any cognitive pair in the two-person system is inconsistent, then dissonance pressure arises and $A_{conf}[k]$ contributes to the dissonance pressure magnitude and, thus, attitude change.

4.2 The Augmented Dynamic Feedback Model

The framework given in Section 2.3 is now used to develop the two-person discrete-time model. First, Component A, which models the theory of planned behaviour, is updated to include $A_{others}[k]$ and $A_{conf}[k]$ as described in Section 4.1.1 and 4.1.2 respectively. Then, Component B, which models cognitive dissonance and the overjustification effect, is updated to include these two new cognitions. Component A and Component B form an updated version of the one-person system with an additional input: the behaviour of another person. The two-person system, shown in Figure 4.1, is essentially composed of two, one-person models connected via behaviour. Specifically, the behaviour of one person in the system is an input to the other person. To distinguish between the signal and parameter values for each person, additional subscripts are used in the notation given in Table 2.2; Table 4.1 summarizes new signals and provides the parameter values used for simulating the two-person system. Finally, although each one-person system allows for two inputs, for this thesis only one of the two people receives a reward input. An item for future work is to consider two input reward signals, one for each person. In the context of the piano situation, person one is the child and person two is the child's friend.

Symbol	**Description [*value or units*]**
$A_{others,i}[k] \in (-\infty, \infty)$	attitude towards the behaviour arising from social pressure [*attitude units*]
$A_{conf,i}[k] \in (-\infty, \infty)$	attitude towards the behaviour arising from conformity pressure [*attitude units*]
$P_i^{SP}[k] \in (-\infty, \infty)$	social pressure [*social pressure units*]
$P_{raw,i}^{OJ,rew}[k], P_i^{OJ,rew}[k] \in [0, \infty)$	pressure arising from the overjustification effect due to rewards [(*attitude units*)^2]
$P_{raw,i}^{OJ,others}[k], P_i^{OJ,others}[k] \in [0, \infty)$	pressure arising from the overjustification effect due to social pressure [(*attitude units*)^2]
$K_{1i} \in [0, \infty)\dagger$	gain reflecting how the dissonance pressure affects attitude change [$K_{11} = K_{12} = 30$ *attitude units*]
$K_{2i} \in [0, \infty)\dagger$	gain reflecting how the overjustification pressure due to rewards affects attitude change [$K_{21} = K_{22} = 0.1$ *1/(attitude unit)*]
$K_{3i} \in [0, \infty)\dagger$	gain reflecting how the overjustification pressure due to social pressure affects attitude change [$K_{31} = K_{32} = 0.1$ *1/(attitude unit)*]
$\mu_{1i} \in [0, \infty)$	value assigned to one dollar [$\mu_{11} = \mu_{12} = 1$ *attitude unit per dollar*]
$\mu_{2i} \in [0, \infty)$	importance of other people who apply social pressure [$\mu_{21} = \mu_{22} = 50$ *attitude units per social pressure unit*]
$\mu_{3i} \in [0, \infty)$	extent to which person one influences person two and vice versa [$\mu_{31} = \mu_{32} = 0.5$ *unitless*]
$r_{1i} \in [0, 1)\dagger$	mental processing pole location for reward attitude formation in (4.2) [$r_{11} = r_{12} = 0$]
$r_{2i} \in [0, 1)\dagger$	mental processing pole location for social pressure attitude formation in (4.3) [$r_{21} = r_{22} = 0$]
$r_{3i} \in [0, 1)\dagger$	mental processing pole location for conformity attitude formation in (4.4) [$r_{31} = r_{32} = 0$]
$r_{4i} \in [0, 1)\dagger$	mental processing pole location for dissonance pressure in (4.16) [$r_{41} = r_{42} = 0.5$]

Continued on Next Page...

Table 4.1 – Continued

Symbol	**Description** [***value or units***]
$r_{5i} \in [0, 1)\dagger$	mental processing pole location for reward overjustification effect in (4.20) $[r_{51} = r_{52} = 0.5]$
$r_{6i} \in [0, 1)\dagger$	mental processing pole location for others overjustification effect in (4.21) $[r_{61} = r_{62} = 0.5]$

Table 4.1: Key signals and parameters that appear in the two-person system. The parameter values listed are those used for analysis in Chapter 5. Parameters marked † depend on the sampling period, taken here to be one day.

4.2.1 Revising Component A

Component A models the theory of planned behaviour. The equations presented in Section 2.3.1 detail how an individual's internal attitude and reward attitude combine to produce a behavioural intent (see (2.2)–(2.4)). Since our model now contains two people, the attitudes and behavioural intent of each person must be represented by distinct equations; thus, let the subscripts i and j (where $i = 1$ and $j = 2$, or $i = 2$ and $j = 1$) differentiate between the attitudes and behaviours of the two individuals. In Section 4.1, two additional attitudes were considered: the attitude arising from the social pressure exerted by others, $A_{others}[k]$, and the attitude arising from conformity pressure, $A_{conf}[k]$. The former attitude depends on social pressure ($P^{SP}[k]$), the importance of the people applying the social pressure (μ_2), and a first-order mental processing model (with r_2 denoting the pole location). Furthermore, the conformity attitude of person i ($A_{conf,i}[k]$) depends on the behaviour of person j ($B_j[k]$), the amount of influence person j has on person i (μ_{3j}), and again, a first order-mental processing model (with r_3 denoting the pole location). The internal attitude and reward attitude are the same as in the one-person model. As before, the theory of planned behaviour states that the weight of each attitude varies by situation and individual; thus, for our model, this weighting is contained in μ_1, μ_2 and μ_3, scaling the reward, social pressure and conformity pressure into attitude units respectively. Thus, Component A for person i is now modelled by the following five equations:

$$A_{out,i}[k] = A_{out,i}[k-1] + \Delta A_{out,i}[k-1] \tag{4.1}$$

$$A_{rew,i}[k] = r_{1i} A_{rew,i}[k-1] + \mu_{1i}\,(1 - r_{1i})\, R_i[k-1] \tag{4.2}$$

$$A_{others,i}[k] = r_{2i} A_{others,i}[k-1] + \mu_{2i}\,(1 - r_{2i})\, P_i^{SP}[k-1] \tag{4.3}$$

$$A_{conf,i}[k] = r_{3i} A_{conf,i}[k-1] + \mu_{3j}\,(1 - r_{3i})\, B_j[k-1] \tag{4.4}$$

$$B_i[k] = A_{out,i}[k] + A_{rew,i}[k] + A_{others,i}[k] + A_{conf,i}[k]. \tag{4.5}$$

4.2.2 Revising Component B

As before, Component B, shown in detail in Figure 4.2, is based on cognitive dissonance theory and the overjustification effect. These two psychological effects create pressure, producing attitude change. Thus, the output of Component B, $\Delta A_{out}[k]$, remains the same as in the one-person system:

$$\Delta A_{out,i}[k] = \Delta A_{out,i}^{CD}[k] + \Delta A_{out,i}^{OJ}[k]. \tag{4.6}$$

The difference between the one-person system and the two-person system lies in how these attitude change terms are formed. Each of these terms will now be examined.

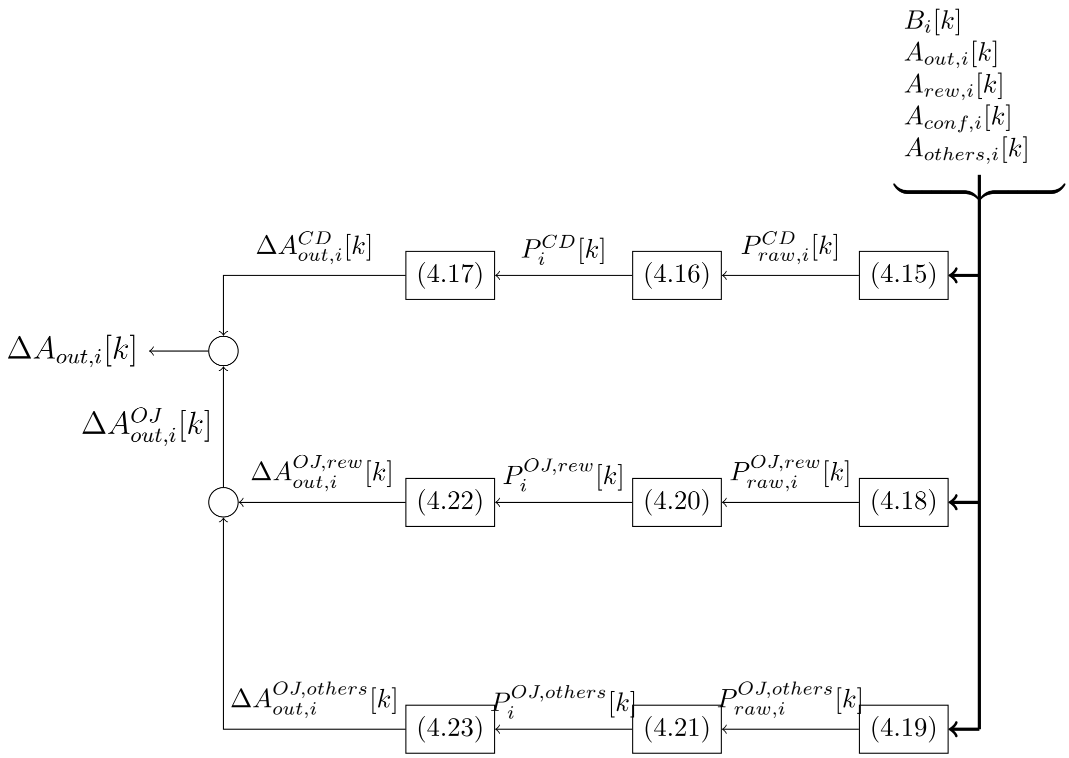

Figure 4.2: Details of the revised Component B. Thick line is used to indicate multiple signals.

Revising the Cognitive Dissonance Theory Model

For the one-person system, the induced compliance paradigm of cognitive dissonance theory is used, resulting in simplified dissonance pressure equations. This paradigm applies when $A_{out}[k] < 0$ and $A_{rew}[k] > 0$ and does not consider how other people affect dissonance pressure. As such, a more general approach is now needed for modelling the dissonance pressure. The assumption that the internal attitude is negative is dropped, while the reward attitude is still assumed to be non-negative. Recall that the dissonance pressure is the ratio between the inconsistency magnitudes and the magnitudes of all cognitive pairs and therefore, the consistency and inconsistency magnitudes for each cognitive pair are first presented.

For the two-person model, person i has four attitudes, $A_{out,i}[k]$, $A_{rew,i}[k]$, $A_{others,i}[k]$ and $A_{conf,i}[k]$, each of which forms a cognitive pair with the generative cognition, $B_i[k]$. Each cognitive pair has an associated $M_{con,i}[k]$ or $M_{incon,i}[k]$. Similar to the one-person model, these magnitudes are proportional to the strength of each element in the cognitive pair; thus, define

$$M_{incon,i}^{1}[k] = \begin{cases} |A_{rew,i}[k]B_i[k]| & \text{if } A_{rew,i}[k] \geq 0, B_i[k] < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.7}$$

$$M_{incon,i}^{2}[k] = \begin{cases} |A_{out,i}[k]B_i[k]| & \text{if } A_{out,i}[k] < 0, B_i[k] \geq 0, \\ & \text{or } A_{out,i}[k] \geq 0, B_i[k] < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.8}$$

$$M_{incon,i}^{3}[k] = \begin{cases} |A_{others,i}[k]B_i[k]| & \text{if } A_{others,i}[k] < 0, B_i[k] \geq 0, \\ & \text{or } A_{others,i}[k] \geq 0, B_i[k] < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.9}$$

$$M_{incon,i}^{4}[k] = \begin{cases} |A_{conf,i}[k]B_i[k]| & \text{if } A_{conf,i}[k] < 0, B_i[k] \geq 0, \\ & \text{or } A_{conf,i}[k] \geq 0, B_i[k] < 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.10}$$

and

$$M_{con,i}^{1}[k] = \begin{cases} |A_{rew,i}[k]B_i[k]| & \text{if } A_{rew,i}[k] \geq 0, B_i[k] \geq 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.11}$$

$$M_{con,i}^{2}[k] = \begin{cases} |A_{out,i}[k]B_i[k]| & \text{if } A_{out,i}[k] < 0, B_i[k] < 0, \\ & \text{or } A_{out,i}[k] \geq 0, B_i[k] \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.12}$$

$$M_{con,i}^{3}[k] = \begin{cases} |A_{others,i}[k]B_i[k]| & \text{if } A_{others,i}[k] < 0, B_i[k] < 0, \\ & \text{or } A_{others,i}[k] \geq 0, B_i[k] \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.13}$$

$$M_{con,i}^{4}[k] = \begin{cases} |A_{conf,i}[k]B_i[k]| & \text{if } A_{conf,i}[k] < 0, B_i[k] < 0, \\ & \text{or } A_{conf,i}[k] \geq 0, B_i[k] \geq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.14}$$

Note that if $A_{conf,i}[k] = 0$, $A_{others,i}[k] = 0$ and $A_{out,i}[k] < 0$, then the above cognitive pair magnitudes reduce to those of the one-person system, given in (2.6)–(2.9). Similar to the one-person model, let $M_{incon,i}[k] = \sum_{m=1}^{4} M_{incon,i}^{m}[k]$ and $M_{con,i}[k] = \sum_{m=1}^{4} M_{con,i}^{m}[k]$. Then, for $B_i[k] \neq 0$, the (raw, unprocessed) dissonance pressure at time k is

$$P_{raw,i}^{CD}[k] = \operatorname{sgn}(B_i[k]) \frac{M_{incon,i}[k]}{M_{incon,i}[k] + M_{con,i}[k]} \tag{4.15}$$

In the case of $B_i[k] = 0$, $P^{CD}_{raw,i}[k]$ has many possible expressions, as there are many possible combinations of consistent/inconsistent cognitive pairs. Instead of listing each of these possible expressions, as was done for the one-person model, the method through which $P^{CD}_{raw,i}[k]$ is determined is presented. The raw, unprocessed dissonance pressure is formed much in the same way as in the one-person case, given in (2.10). Specifically, the denominator equals the sum of all cognition magnitudes (other than $B_i[k]$) and the numerator is formed by summing the magnitude of each cognition that is inconsistent with $B_i[k]$. From (4.7)–(4.14), a cognition is inconsistent with $B_i[k] = 0$ if it is strictly negative. As a result, only $A_{out,i}[k]$, $A_{others,i}[k]$ and $A_{conf,i}[k]$ can be inconsistent with behaviour $B_i[k]$ (since the reward is restricted to non-negative values). Finally, as in the one-person system, the "sgn $(B[k])$" factor ensures that the sign of $\Delta A^{CD}_{out}[k]$ is consistent with cognitive dissonance theory in that it results in a decrease in dissonance pressure at the next time instant (since by Assumption 2.4, the dissonance pressure reduction mechanism is changing the internal attitude).

Having derived an expression for the raw dissonance pressure, the first-order mental processing model is applied in the same way as in the one-person model; thus,

$$P^{CD}_i[k] = r_{4i} P^{CD}_i[k-1] + (1 - r_{4i})\, P^{CD}_{raw,i}[k]. \tag{4.16}$$

Finally, the dissonance pressure results in attitude change, denoted $\Delta A^{CD}_{out,i}[k]$, is given by

$$\Delta A^{CD}_{out,i}[k] = K_{1i} P^{CD}_i[k]. \tag{4.17}$$

Similar to the one-person model, in (4.17) it is assumed that attitude change is proportional to the dissonance pressure; psychologists have yet to identify the exact relationship.

Revising the Overjustification Effect Model

As discussed in Section 2.2.3, the overjustification effect applies only when all cognitions are non-negative. The additional cognition $A_{others,i}[k]$ arises from overt social pressure due to other individuals who are significant to person i. Similar to a reward that is used to overtly control the behaviour of person i, this social pressure tends to generate overjustification pressure (see Assumption 4.1). On the other hand, by Assumption 4.2, the second new cognition, $A_{conf,i}[k]$, does not generate overjustification pressure as person i does not feel overtly controlled through conformity pressure, as is the case when rewards ($R[k]$) and social pressure ($P^{SP}[k]$) are applied. Therefore, two overjustification pressures may arise: $P^{OJ,rew}_i[k]$ and $P^{OJ,others}_i[k]$. The basic pressures arising from the overjustification effect are modelled as

$$P^{OJ,rew}_{raw,i}[k] = \begin{cases} A_{out,i}[k] A_{rew,i}[k] & \text{if } A_{out,i}[k] > 0, A_{rew,i}[k] > 0, \\ & A_{others,i}[k] \geq 0, A_{conf,i}[k] \geq 0 \text{ and } B_i[k] > 0, \\ 0 & \text{otherwise,} \end{cases} \tag{4.18}$$

and

$$P_{raw,i}^{OJ,others}[k] = \begin{cases} A_{out,i}[k]A_{others,i}[k] & \text{if } A_{out,i}[k] > 0, A_{rew,i}[k] \geq 0, \\ & A_{others,i}[k] > 0, A_{conf,i}[k] \geq 0 \text{ and } B_i[k] > 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4.19}$$

Like the one-person system, in (4.18) and (4.19), it is assumed that the pressure depends on the product of the relevant attitudes.

Assuming first-order mental processing of the pressures $P_{raw,i}^{OJ,rew}[k]$ and $P_{raw,i}^{OJ,others}[k]$, the following equations for the processed pressures are obtained:

$$P_i^{OJ,rew}[k] = r_{5i}P_i^{OJ,rew}[k-1] + (1 - r_{5i})\, P_{raw,i}^{OJ,rew}[k], \tag{4.20}$$

$$P_i^{OJ,others}[k] = r_{6i}P_i^{OJ,others}[k-1] + (1 - r_{6i})\, P_{raw,i}^{OJ,others}[k]. \tag{4.21}$$

Finally, to avoid the situation where $P_i^{OJ,rew}[k]$ and $P_i^{OJ,others}[k]$ are large enough to actually change the sign of $A_{out,i}[k]$ (an effect that is inconsistent with overjustification theory), $P_i^{OJ,rew}[k]$ and $P_i^{OJ,others}[k]$ are saturated as follows:

$$\Delta A_{out,i}^{OJ,rew}[k] = \begin{cases} -K_{2i}P_i^{OJ,rew}[k] & \text{if } P_i^{OJ,rew}[k] > 0, K_{2i}P_i^{OJ,rew}[k] \leq A_{out,i}[k], \\ -A_{out,i}[k] & \text{if } P_i^{OJ,rew}[k] > 0, K_{2i}P_i^{OJ,rew}[k] > A_{out,i}[k], \\ 0 & \text{otherwise,} \end{cases} \tag{4.22}$$

and

$$\Delta A_{out,i}^{OJ,others}[k] = \begin{cases} -K_{3i}P_i^{OJ,others}[k] & \text{if } P_i^{OJ,others}[k] > 0, K_{3i}3P_i^{OJ,others}[k] \leq A_{out,i}[k], \\ -A_{out,i}[k] & \text{if } P_i^{OJ,others}[k] > 0, K_{3i}P_i^{OJ,others}[k] > A_{out,i}[k], \\ 0 & \text{otherwise.} \end{cases} \tag{4.23}$$

Combining these two attitude change equations yields

$$\Delta A_{out,i}^{OJ}[k] = \Delta A_{out,i}^{OJ,rew}[k] + \Delta A_{out,i}^{OJ,others}[k].$$

Again, like the one-person model, attitude change in (4.22) and (4.23) is assumed to be proportional to the experienced psychological pressure.

4.2.3 Simplifying Assumptions and Initial Conditions

Similar to the one-person system, the problem statement given in Chapter 1 suggests some preliminary assumptions and initial conditions. The original problem statement considers the child and his friend. The child's initial attitude is negative, whereas his friend's initial attitude is arbitrary. Furthermore, only the child is offered a reward and as a result, the following assumption is made on the reward offered to the friend:

Assumption 4.3. $A_{rew,2}[k] = 0$ *for* $k \geq 0$.

Moreover, the child experiences two external pressures: a reward offered to him by his parents, and conformity pressure exerted by his friend; thus, this thesis is restricted to the effects of rewards and conformity pressure (and not any other social pressure) and the following assumption is made:

Assumption 4.4. $A_{others,i}[k] = 0$ *for* $k \geq 0$, $i = 1, 2$.

Assumption 4.4 simplifies the expression for $B_i[k]$ in (4.5) to

$$B_i[k] = A_{out,i}[k] + A_{rew,i}[k] + A_{conf,i}[k]. \tag{4.24}$$

In addition, note that from Assumption 4.4 and (4.19), $P_{raw,i}^{OJ,others}[k] = 0$ for all $k \geq 0$ and thus from (4.21), $P_i^{OJ,others}[k] = 0$ for all $k \geq 0$. As a result, $\Delta A_{out,i}^{OJ,others}[k] = 0$ and consequently the expression for $\Delta A_{out,i}^{OJ}[k]$ simplifies to

$$\Delta A_{out,i}^{OJ}[k] = \Delta A_{out,i}^{OJ,rew}[k]. \tag{4.25}$$

Furthermore, previous to sample $k = 0$, the child and his friend do not experience any external influences; thus, initially, the behavioural intentions of both the child and his friend match their respective initial internal attitudes. Finally, a last reasonable assumption is that the child and his friend have not previously experienced any pressures due to dissonance of overjustification effects and thus, both experience no initial attitude change. Hence, the initial conditions for $i = 1, 2$ are:

$$\begin{aligned} P_i^{CD}[0] = P_i^{OJ}[0] &= 0, \\ \Delta A_{out,i}^{CD}[0] = \Delta A_{out,i}^{OJ}[0] &= 0, \\ A_{rew,i}[0] = A_{conf,i}[0] &= 0, \\ B_i[0] = A_{out,i}[0] &= A_{oi}, \end{aligned} \tag{4.26}$$

where $A_{o1} < 0$ and A_{o2} is arbitrary. Note that, by Assumption 4.4, the initial conditions on $A_{others,1}[k]$ and $A_{others,2}[k]$ are omitted.

Finally, similar to the one-person system, assumptions are made on our model's parameter values as follows:

Assumption 4.5. *Gains reflecting how the dissonance and overjustification pressures affect attitude change are strictly positive, i.e.,* $K_{1i} > 0$, $K_{2i} > 0$ *and* $K_{3i} > 0$.

Assumption 4.6. *The value assigned to one dollar is strictly positive, i.e.,* $\mu_{1i} > 0$.

Assumption 4.7. *The extent to which each person in the two-person system experiences conformity pressure cannot be more than* 100%, *i.e.,* $0 \leq \mu_{31} < 1$ *and* $0 \leq \mu_{32} < 1$.

Assumption 4.8. *The mental processing pole locations for reward attitude formation and the conformity attitude formation in (4.2) and (4.4) are zero, i.e.,* $r_{1i} = r_{3i} = 0$.

Assumption 4.8 simplifies the expressions for $A_{rew,i}[k]$ and $A_{conf,i}[k]$ in (4.2) and (4.4) respectively to

$$A_{rew,i}[k] = \mu_{1i} R_i[k-1] \tag{4.27}$$

$$A_{conf,i}[k] = \mu_{3j} B_j[k-1]. \tag{4.28}$$

Assumption 4.9. *The mental processing pole locations for dissonance and overjustification pressures in (4.16), (4.20), and (4.21) are contained in the range* $[0, 1)$, *i.e.,* $0 \leq r_{4i}, r_{5i}, r_{6i} < 1$.

These assumptions are used throughout the remainder of this chapter, as well as in Chapter 5. A final assumption is used only for analysis and does not apply to simulations:

Assumption 4.10. *The mental processing pole location for dissonance pressure in (4.16) is zero, i.e.,* $r_{4i} = 0$ *for* $i = 1, 2$.

Assumption 4.10 simplifies the expressions for $P_i^{CD}[k]$ (4.16) to

$$P_i^{CD}[k] = P_{raw,i}^{CD}[k]. \tag{4.29}$$

4.3 Model Consistency Verification

The model of Section 4.2 is now verified to be qualitatively consistent with the psychology presented in Section 4.1. First, the social pressure component is studied by considering how it contributes to the theory of planned behaviour, cognitive dissonance theory, and the overjustification effect. Then, the conformity pressure component of the two-person system is examined.

4.3.1 Social Pressure Consistency Verification

We show that the following four qualitative features are exhibited by the social pressure component of the two-person system:

1. Strong social pressures have a greater influence on behavioural intent then weak social pressures.
2. A person tends to reduce dissonance pressure, should it arise.
3. Overjustification pressure, should it arise, cannot cause attitude to switch signs.
4. The magnitude of the attitude change arising from overjustification pressure increases with an increase in social pressure and/or attitude towards behavioural outcome.

To perform this qualitative analysis, it is assumed for simplicity that no conformity pressure arises and no reward is offered and thus, the subscript i is omitted from the model equations.

First, a strong social pressure is given by higher magnitudes of $P^{SP}[k]$ in (4.3), leading to a large $A_{others}[k]$ magnitude. From (4.5), large $A_{others}[k]$ magnitudes have a greater influence on $B[k]$ than small magnitudes; hence, the model displays the first qualitative characteristic listed above.

The second trend, related to dissonance pressure, is the same as in the one-person system. Thus, similar to the one-person system, to show dissonance pressure is reduced through attitude change, it is sufficient (recalling Assumption 4.10) to show that

$$|P_{raw}^{CD}[k+1]| < |P_{raw}^{CD}[k]|$$

for any combination of dissonance pressures. Given that $A_{conf}[k]$ and $A_{rew}[k]$ are assumed to be zero for this analysis, there are two cognitive pairs, $(A_{others}[k], B[k])$ and $(A_{out}[k], B[k])$; thus, from the two-person model equations, there are four possible situations in which dissonance may arise. These situations are given as follows:

1. $A_{others}[k] \geq 0$, $A_{out}[k] < 0$, $B[k] \geq 0$,
2. $A_{others}[k] \geq 0$, $A_{out}[k] < 0$, $B[k] < 0$,
3. $A_{others}[k] < 0$, $A_{out}[k] \geq 0$, $B[k] \geq 0$, and
4. $A_{others}[k] < 0$, $A_{out}[k] \geq 0$, $B[k] < 0$.

The first two situations are the same as the induced compliance paradigm discussed in the one-person case, with $A_{others}[k]$ replacing $A_{rew}[k]$. As such, the first two situations have already been shown to satisfy the requirement that $|P_{raw}^{CD}[k+1]| < |P_{raw}^{CD}[k]|$. Now, it only remains to show that the last two situations satisfy this relationship.

For the situation in which $B[k] \geq 0$, from (4.12), $M_{con}[k] = |A_{out}[k]B[k]|$ and from (4.9), $M_{incon}[k] = |A_{others}[k]B[k]|$. Assuming $B[k] \neq 0$, from (4.15),

$$P_{raw}^{CD}[k] = \frac{|A_{others}[k]B[k]|}{|A_{others}[k]B[k]| + |A_{out}[k]B[k]|} = \frac{|A_{others}[k]|}{|A_{others}[k]| + |A_{out}[k]|}. \tag{4.30}$$

Note that if $B[k] = 0$, then $P^{CD}_{raw}[k] = \frac{|A_{others}[k]|}{|A_{others}[k]|+|A_{out}[k]|}$, which is the same as the above expression. From (4.16), this raw, unprocessed dissonance pressure produces a positive $P^{CD}[k]$ and thus, from (4.17) and Assumption 4.5, a positive attitude change, i.e.,

$$A_{out}[k+1] = A_{out}[k] + K_1 P^{CD}[k] > A_{out}[k].$$

Note that $\Delta A^{OJ}_{out}[k] = 0$ since the conditions required for overjustification effect to hold are not met. To see how this increased attitude affects the raw, unprocessed dissonance pressure at sample $k+1$, assume that the social pressure is constant, ensuring only the effect of changes to the internal attitude is studied. Since the internal attitude increases and the social pressure remains constant, $B[k]$ necessarily increases and is, therefore, still positive. Hence, from (4.15),

$$\begin{aligned} P^{CD}_{raw}[k+1] &= \frac{|A_{others}[k+1]B[k+1]|}{|A_{others}[k+1]B[k+1]| + |A_{out}[k+1]B[k+1]|} \\ &= \frac{|A_{others}[k+1]|}{|A_{others}[k+1]| + |A_{out}[k+1]|} \end{aligned}$$

Since $A_{others}[k+1] = A_{others}[k]$ and $|A_{out}[k+1]| > |A_{out}[k]|$, it follows that $P^{CD}_{raw}[k+1] < P^{CD}_{raw}[k]$, as required. Note that since $P^{CD}_{raw}[k] > 0$, the magnitude signs are omitted.

For the situation in which $B[k] < 0$, from (4.13), $M_{con}[k] = |A_{others}[k]B[k]|$ and from (4.8), $M_{incon}[k] = |A_{out}[k]B[k]|$. From (4.15),

$$P^{CD}_{raw}[k] = -\frac{|A_{out}[k]B[k]|}{|A_{out}[k]B[k]| + |A_{others}[k]B[k]|} = -\frac{|A_{out}[k]|}{|A_{out}[k]| + |A_{others}[k]|}. \quad (4.31)$$

From (4.16), this raw, unprocessed dissonance pressure produces a negative $P^{CD}[k]$ and thus from (4.17) and Assumption 4.5, a negative attitude change, i.e.,

$$A_{out}[k+1] = A_{out}[k] + K_1 P^{CD}[k] < A_{out}[k].$$

Similar to before, $\Delta A^{OJ}_{out}[k] = 0$. The above attitude effectively decreases $B[k]$ (again, assuming a constant reward attitude); thus $B[k+1] < 0$. Moreover, if this decrease causes $A_{out}[k+1] < 0$, then $P^{CD}_{raw}[k+1] = 0$ and thus, dissonance pressure magnitude is reduced via attitude change. On the other hand, if $A_{out}[k+1] \geq 0$, then from (4.15),

$$\begin{aligned} P^{CD}_{raw}[k+1] &= -\frac{|A_{out}[k+1]B[k+1]|}{|A_{out}[k+1]B[k+1]| + |A_{others}[k+1]B[k+1]|} \\ &= -\frac{|A_{out}[k+1]|}{|A_{out}[k+1]| + |A_{others}[k+1]|}. \end{aligned}$$

Since $A_{others}[k+1] = A_{others}[k]$ and $A_{out}[k+1] < A_{out}[k]$, it follows that $|P^{CD}_{raw}[k+1]| < |P^{CD}_{raw}[k]|$, as required. Therefore, in each of the four possible situations in which dissonance pressure arises exclusively from social pressure, attitude change results in a reduction of the amount of dissonance pressure experienced.

For illustrative purposes, a simulation is performed for two of the situations in which dissonance may arise. In both simulations, shown in Figure 4.3, a constant, negative social pressure is applied to a person in the system, whose initial attitude is positive. Furthermore, this person has no reward attitude and no conformity attitude. The simulation results show the response of the person in question (since no conformity is present). The first simulation, shown in the left plots, is related to the situation in which $A_{others}[k] < 0$, $A_{out}[k] \geq 0$, and $B[k] < 0$, whereas the second simulation is shown the right plots, and is related to the situation in which $A_{others}[k] < 0$, $A_{out}[k] \geq 0$, and $B[k] \geq 0$. As both simulations in Figure 4.3 show, when dissonance pressure arises, the raw, unprocessed dissonance pressure strength decreases through attitude change. Furthermore, the processed dissonance pressure magnitude, $|P^{CD}[k]|$, also decreases, as expected. (Since the situations in which $A_{others}[k] \geq 0$ and $A_{out}[k] < 0$ are essentially the same as the induced compliance paradigm of cognitive dissonance theory, no simulations are given because they would be the same as those shown in Figure 2.5.) Therefore, the two-person model is consistent with the second qualitative characteristic on page 90.

The remaining two characteristics, related to the overjustification pressure, are the same as the overjustification pressure trends in the one-person system. As the equations for the overjustification pressure arising from social pressure, i.e., $P^{OJ}_{others}[k]$, are the same as those for the overjustification pressure arising from rewards, i.e., $P^{OJ}_{rew}[k]$, which has been shown to demonstrate these two trends (see Section 2.4), the two remaining characteristics are indeed exhibited by the two-person system.

4.3.2 Conformity Pressure Consistency Verification

The conformity pressure component of the two-person model is now considered. For this verification, assume no reward or social pressure is applied. In this case, the two-person model is called the *zero-input system* because no external input is applied to the system. Recall from Section 4.1.2 that there are two types of conformity, compliance and acceptance. Compliance occurs when a person's behaviour switches sign due to conformity pressure, while his attitude does not switch sign, whereas acceptance occurs when a person switches the sign of both his behaviour and attitude; typically behaviour switches first. For our model, showing that it is possible for a person's behaviour to switch sign due to strong *counter-attitudinal conformity pressure* (which refers to the conformity pressure a person experiences that is contrary to his internal attitude) is evidence that our model is consistent with at least compliance. Additionally, showing that the sign of the internal attitude can also change is evidence that our model exhibits acceptance. As the acceptance type of conformity was not explicitly taken into account when modelling the two-person

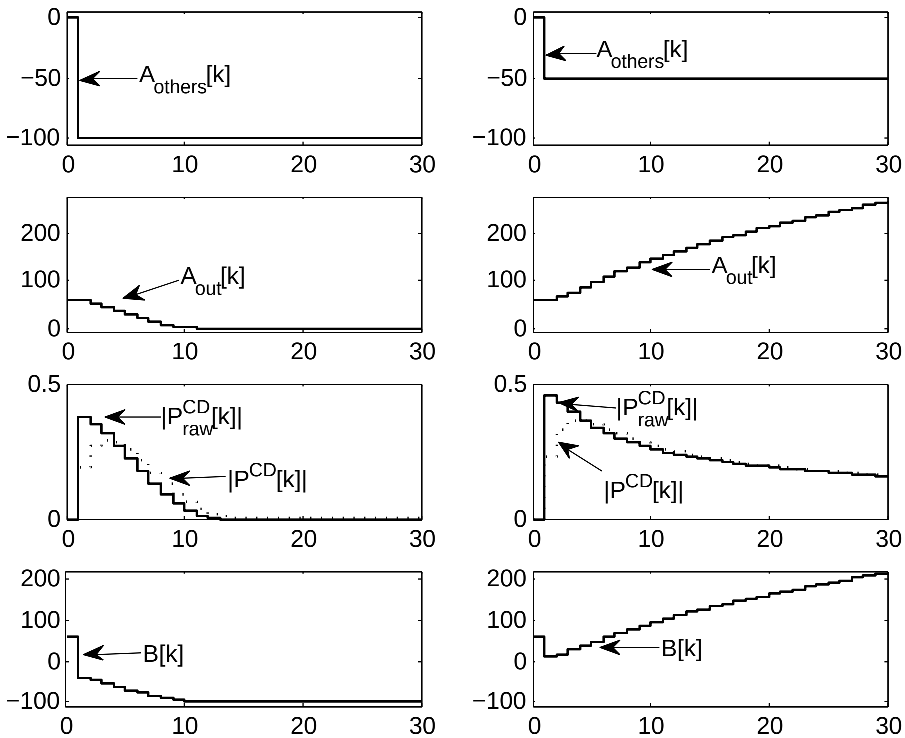

Figure 4.3: Simulation demonstrating how the two-person system responds to constant social pressure. Set $A_o = 60$ attitude units and consider two social pressure magnitudes: negative social pressure that is strong enough to drive behaviour negative (say, $P^{SP}[k] = -2$ *social pressure units*, left plots) and negative social pressure that does not cause behaviour to become negative (say, $P^{SP}[k] = -1$ *social pressure units*, right plots). The simulations results show how $A_{out}[k]$, $A_{others}[k]$, $P^{CD}[k]$ and $B[k]$ respond to the applied social pressure. For the stronger social pressure, $A_{out}[k]$ decreases. This has the effect of reducing $|P^{CD}_{raw}[k]|$ and $|P^{CD}[k]|$. For the weaker social pressure, $A_{out}[k]$ increases, which has the effect of reducing $P^{CD}_{raw}[k]$ and $P^{CD}[k]$. However, since $A_{out}[k] \geq 0$, $B[k] \geq 0$ and $A_{others}[k] < 0$ for $k \geq 0$, there is always dissonance pressure. It can be shown that this dissonance pressure cannot converge to zero as k tends to infinity; thus $A_{out}[k]$ (and hence, $B[k]$) is an unbounded increasing function of k as k tends to infinity.

system, showing the model exhibits this effect is of potential interest to psychologists as conformity theory does not traditionally link acceptance with dissonance pressure effects.

To begin, our model is shown to exhibit compliance. From the psychology literature, strong conformity pressure influences an individual's behaviour more than weak conformity pressure [11]. This trend is exhibited by our model's theory of planned behaviour equations, specifically, for the zero-input system,

$$B_i[k] = A_{out,i}[k] + A_{conf,i}[k].$$

Thus, when $|A_{conf,i}[k]|$ is greater, its effect on $B_i[k]$ is also greater. Moreover, if $A_{out,i}[k] < 0$, then, provided $A_{conf,i}[k] \geq 0$ and is sufficiently large, it is possible for $B_i[k] \geq 0$. In other words, a person's behaviour can become positive due to positive conformity pressure, even though his internal attitude is negative. A similar argument holds when $A_{out,i}[k] \geq 0$ and $A_{conf,i}[k] < 0$. Hence, our model can demonstrate compliance effects.

To show our model can exhibit acceptance, the internal attitude is considered. In particular, if conformity pressure forces behaviour to become positive while the internal attitude is negative, then acceptance is demonstrated if the internal attitude also becomes, at some point in time, positive. Similarly, if conformity pressure forces behaviour to become negative while the internal attitude is positive, then acceptance is demonstrated if the internal attitude also becomes, at some point in time, negative. To determine whether or not these two trends are exhibited by the two-person system, the effect of a counter-attitudinal conformity pressure on one of the two people in our model is studied. This conformity pressure is chosen to be high enough to force the initially consistent cognitive pair, $(A_{out,i}[k], B_i[k])$, to become inconsistent, by changing the sign of $B_i[k]$. If the internal attitude changes according to the above two trends, then we will have shown that acceptance can be exhibited by our model.

Suppose person one has a negative internal attitude and experiences zero conformity pressure before sample $\bar{k} > 0$, i.e., $A_{conf,1}[k] = 0$ for $0 \leq k < \bar{k}$. Then, since $A_{rew,1}[k] = A_{others,1}[k] = 0$ for $k \geq 0$, the behaviour of person one, $B_1[k]$, equals the internal attitude, $A_{out,1}[k]$, for $0 \leq k < \bar{k}$. Both $A_{out,1}[k]$ and $B_1[k]$ remain constant for $0 \leq k < \bar{k}$ since $P_1^{CD}[k] = 0$ and $P_1^{OJ}[k] = 0$ for $0 \leq k < \bar{k}$, implying $\Delta A_{out,1}[k] = 0$ over this sample range. Let $A_{conf,1}[k] > 0$ at $k = \bar{k}$ and suppose this conformity pressure is high enough to cause $B_1[\bar{k}] \geq 0$. Then $(A_{out,1}[\bar{k}], B_1[\bar{k}])$ form an inconsistent cognitive pair and, therefore, from our model's dissonance pressure equations,

$$P_{raw,1}^{CD}[\bar{k}] = \frac{|A_{out,1}[\bar{k}]|}{|A_{out,1}[\bar{k}]| + |A_{conf,1}[\bar{k}]|} > 0.$$

Since the raw unprocessed dissonance pressure is positive at $k = \bar{k}$ and zero for $0 \leq k < \bar{k}$, it follows from the model equations that $\Delta A_{out,1}^{CD}[\bar{k}] > 0$. Moreover, since no overjustification pressure arises, $\Delta A_{out,1}[\bar{k}] = \Delta A^{CD}_{out,1}[\bar{k}] > 0$. Therefore,

$$A_{out,1}[\bar{k}+1] = A_{out,1}[\bar{k}] + \Delta A_{out,1}[\bar{k}] > A_{out,1}[\bar{k}].$$

In other words, person one's internal attitude increases. If $A_{out,1}[\bar{k}+1] \geq 0$, then acceptance occurs. Otherwise, even though attitude increases, compliance occurs. A similar argument can be used to show that if person one has a positive internal attitude and experiences a sufficiently large, negative conformity pressure, then $A_{out,1}[k]$ decreases. If the internal attitude decreases to a negative value, then acceptance occurs; otherwise, compliance occurs.

To summarize, our model is consistent with the psychological trend that conformity pressure influences behaviour. Moreover, a sufficiently large counter-attitudinal conformity pressure changes a behaviour's sign and thus, our model can exhibit compliance. Finally, since this sufficiently large counter-attitudinal conformity pressure produces attitude change (through our model's dissonance equations), acceptance may also arise. Therefore, our model is qualitatively consistent with the psychology related to conformity.

Combining the results of this section show that the two-person system is qualitatively consistent with the psychology used to develop our model. There are, however, other psychological phenomena that arise in a group setting, which were not taken into account when modelling the two-person system. In the next section, these trends are introduced and discussed within the context of our model.

4.4 Other Psychological Phenomena

Psychology researchers who study group dynamics have discovered various phenomena related to how people's attitudes and behaviours change when in a group setting. Three of these possible trends are relevant to the two-person system and are given below:

1. When all people in a group share a similar attitude, i.e., the sign of each person's attitude is the same, the average attitude and/or behaviour of the group tends to increase. This is called *group polarization* [14].

2. When a person in a group experiences pressure to carry out a behaviour that is contrary (in sign) to his internal attitude, this person conforms (by complying to this pressure) if his attitude is weak, relative to the average pressure exerted by the group [18].

3. When a person in a group experiences pressure to perform a behaviour that is contrary (in sign) to his internal attitude, this person *reacts* (by strengthening the magnitude of his pre-existing attitude and/or behaviour) if his attitude is strong, relative to the average pressure exerted by the group [18].

Initial Region	A_{o1}	A_{o2}	**Relationship**
I	positive	positive	any
II	positive	negative	$\lvert A_{o2}\rvert \leq \mu_{31} A_{o1}$
III	positive	negative	$\mu_{31} A_{o1} < \lvert A_{o2}\rvert \leq \frac{A_{o1}}{\mu_{32}}$
IV	positive	negative	$\frac{A_{o1}}{\mu_{32}} < \lvert A_{o2}\rvert$
V	negative	negative	any
VI	negative	positive	$\frac{A_{o2}}{\mu_{31}} < \lvert A_{o1}\rvert$
VII	negative	positive	$\mu_{32} A_{o2} < \lvert A_{o1}\rvert \leq \frac{A_{o2}}{\mu_{31}}$
VIII	negative	positive	$\lvert A_{o1}\rvert \leq \mu_{32} A_{o2}$

Table 4.2: Initial operating regions for the two-person system.

The development of the two-person system modelled the immediate impact conformity pressure has on the behaviour of an individual and did not attempt to model the three trends described above; thus, showing the two-person system is qualitatively consistent with these trends is of potential interest to psychologists and to us, since it provides further evidence of the validity of our model. Most interesting, perhaps, is the consistency of the model with the reactance trend above, as in this situation, dissonance pressure arises and is the mechanism through which this trend is exhibited. To the best of our knowledge, psychologists have not explicitly linked the reactance trend with cognitive dissonance theory, but our model suggests that dissonance theory may be a possible explanation for this characteristic.

The three psychological trends above depend on a person's internal attitude, and its strength, relative to the average conformity pressure exerted by the group. For our model, there are two people, and therefore, there is no need to average the conformity pressure experienced by each person. To determine which of the three psychological trends a person in our model will exhibit, the relative strength of each initial attitude is examined (because each initial attitude produces a conformity pressure at $k = 1$). Table 4.2 defines eight possible relationships between the two initial attitudes. These eight relationships form eight initial operating regions, shown graphically in Figure 4.4. If the zero-input system begins in regions I or V, then both people share a similar (i.e., same sign) attitude and, therefore, psychology predicts that group polarization should occur. If the system begins in region II, IV, VI or VIII, then the two people do not share a similar attitude and one attitude is relatively weak, while the other is relatively strong; thus, psychology predicts that the person with the weak attitude should conform and the person with the strong attitude should react. Finally, if the system begins in region III or VII, then the two people do not share a similar attitude, and neither attitude is weak (or strong) relative to the other. In this case, psychology does not predict whether reactance or conformity

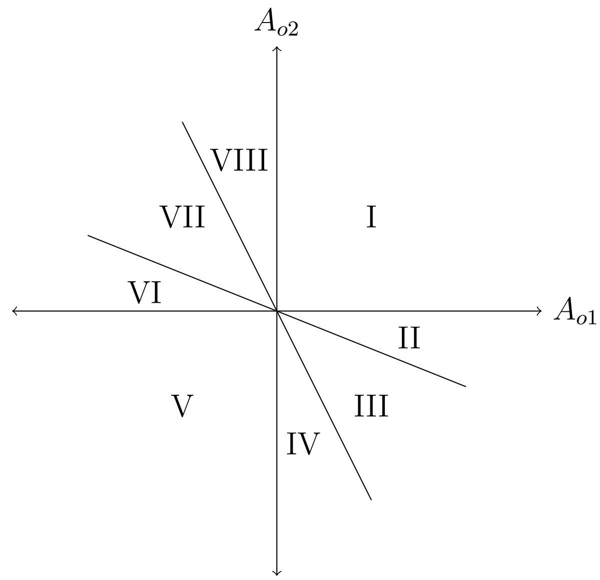

Figure 4.4: Graphical representation of initial operating regions of the two-person system. Regions are denoted on the graph and lines are given by the relationships in Table 4.2.

should occur (only that each person should exhibit one of these two trends). Table 4.3 summarizes these expected psychological trends for each initial region given in Table 4.2. Note that because each person's attitude may change over time, the relative strength of each person's attitude may also change, in which case, the expected psychological trend may change. As a result, Table 4.3 summarizes the *short-term* expected psychological trend for the zero-input system. Showing that our model demonstrates the three group dynamic trends above (group polarization, reactance, conformity) amounts to showing the zero-input system response of each initial region is consistent with the relevant psychological trend given in Table 4.3. This analysis is carried out in Chapter 5 due to its relevance to controller design.

To summarize, this chapter relaxes the assumption that an individual experiences only one external influence (a reward) and considers the effect a second person has on the attitude and behaviour of the first individual. Two general types of external influences were explained, social pressure and conformity pressure. Then, a two-person discrete-time model was developed using the framework of the one-person model and the new psychology presented in this chapter. The two-person system was shown to be qualitative consistency with the psychology related to social pressure and conformity pressure. Finally, other psychological phenomena arising from group dynamics were introduced. Verification of the consistency between the zero-input system and these phenomena is given in Chapter 5, along with open-loop and closed-loop controller design.

Initial Region	**Expected Psychological Trend for Person One**	**Expected Psychological Trend for Person Two**
I	Group Polarization	Group Polarization
II	Reactance	Conformity
III	Conformity or Reactance	Reactance or Conformity
IV	Conformity	Reactance
V	Group Polarization	Group Polarization
VI	Reactance	Conformity
VII	Conformity or Reactance	Reactance or Conformity
VIII	Conformity	Reactance

Table 4.3: Short-term psychological trends expected for each initial region of the two-person system.

Chapter 5

Two-Person System: Simulation and Control Strategies

This chapter explores open-loop and closed-loop control strategies to achieve a control objective for the two-person system with the initial conditions given in (4.26): the control objective is to determine whether or not, for any $B_d > 0$, there exists a reward sequence, $R_1[k]$, such that $B_1[k] \geq B_d$ as k tends to infinity. For this analysis, the assumptions presented in Section 4.2.3 are used.

In the last section of Chapter 4, three psychological phenomena related to group dynamics were introduced: group polarization, conformity and reactance. To determine whether or not these trends are exhibited by the two-person system, the zero-input response of each of the eight initial operating regions (see Figure 4.4) is determined. This chapter begins with this analysis. Key qualitative characteristics uncovered in this analysis are then exploited to design controllers that meet the control objective. First, open-loop controllers are considered; then, due to drawbacks of open-loop control, closed-loop controllers are designed to meet the required control objective.

5.1 Zero-Input Response Analysis

Table 5.1 collects together various responses of the zero-input system, labelled A1–D1. Each of these responses characterizes at least one of the three psychological phenomena arising from group dynamics, as indicated in Table 5.1. By comparing the zero-input response for each initial region with the response descriptions in Table 5.1, the psychological trend(s) arising for each initial region can be determined. In some initial regions, the system exhibits different psychological trends over time, i.e., the short-term psychological trends exhibited by some initial regions are different than the long-term trends. For each initial region, if the short-term psychological trends exhibited by the model's zero-input response are consistent with the expected psychological trends, as given in Table 4.3, then the model is qualitatively consistent with the trends predicted by psychology related to group dynamics.

For this analysis, regions V, VI, VII and VIII are considered since, from the initial conditions, A_{o1} is restricted to negative values. From Table 4.2, initial regions I, II, III and IV correspond with non-negative values for A_{o1}. Even though these four regions are not included in the analysis, the initial regions are symmetric about the $A_{o1} = -A_{o2}$ axis (see Figure 4.4). Furthermore, each person in the two-person system is modelled by the same set of equations. As such, the zero-input system response is symmetric about the $A_{o1} = -A_{o2}$ axis, meaning the responses of initial regions I, II, III and IV can be inferred from those of initial regions V, VI, VII and VIII. Finally, in some cases, an approximation is made for $|\mu_{31}^n \mu_{32}^m|$ for any whole numbers m and n with $m + n \geq 2$. Specifically, since $|\mu_{31}| < 1$ and $|\mu_{32}| < 1$ (by Assumption 4.7), their product is also less than one. For some of this analysis, this product is approximated as $|\mu_{31}^n \mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$.

5.1.1 Zero-Input Response Analysis: Region V

First, suppose the system begins in region V, which is defined by the following initial attitude relationship:

$$A_{o1}, A_{o2} < 0. \tag{5.1}$$

From the system's initial conditions given in (4.26), $B_1[0] = A_{o1}$ and $B_2[0] = A_{o2}$; thus, both behaviours are also initially negative. As a result, each person experiences negative conformity pressure, further decreasing their behaviour at $k = 1$. Each person's behaviour at $k = 1$ is still consistent with their attitude at $k = 1$ because for $i = 1, 2$, $A_{out,i}[1] = A_{out,i}[0] + \Delta A_{out,i}[0]$ and, from (4.26), $\Delta A_{out,i}[0] = 0$ for $i = 1, 2$, meaning $A_{out,i}[1] = A_{oi} < 0$. Moreover, since each person experiences negative conformity pressure at $k = 1$, the cognitive pair $(A_{conf,i}[1], B_i[1])$ is also consistent, implying dissonance pressure does not arise at $k = 1$. Finally, by Assumption 4.2 and the fact that the zero-input system assumes $A_{rew,i}[k] = 0$ for $k \geq 0$, $i = 1, 2$, overjustification pressure cannot arise, i.e., $P_i^{OJ,rew}[k] = 0$ for $k \geq 0$, $i = 1, 2$; consequently, $\Delta A_{out,i}[1] = 0$ for $i = 1, 2$. This argument can be repeated for each $k \geq 0$, leading to the following conclusion:

Lemma 5.1. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.1), then, for $k \geq 0$, and $i = 1, 2$,*

(a) $\Delta A_{out,i}^{CD}[k] = 0$ *and*

	Response Description	Psychological Trend for Person One	Psychological Trend for Person Two
A1	For $k \geq \hat{k}$, $A_{out,1}[k], A_{out,2}[k], B_1[k], B_2[k] \geq 0$. $B_1[k], B_2[k]$ are increasing functions of k and settle to steady-state values.	Group Polarization	Group Polarization
A2	For $k \geq \hat{k}$, $A_{out,1}[k], B_1[k], B_2[k] \geq 0$, $A_{out,2}[k] < 0$. $B_1[k], B_2[k]$ are increasing functions of k and settle to steady-state values.	Reactance	Conformity
A3	For $k \geq \hat{k}$, $A_{out,2}[k], B_1[k], B_2[k] \geq 0$, $A_{out,1}[k] < 0$. $B_1[k], B_2[k]$ are increasing functions of k and settle to steady-state values.	Conformity	Reactance
B1	For $k \geq \hat{k}$, $A_{out,1}[k], B_1[k] \geq 0$, $A_{out,2}[k], B_2[k] < 0$. $B_1[k]$ increases and is unbounded and $B_2[k]$ decreases and is unbounded.	Reactance	Reactance
C1	For $k \geq \hat{k}$, $A_{out,1}[k], A_{out,2}[k], B_1[k], B_2[k] < 0$. $B_1[k], B_2[k]$ are decreasing functions of k and settle to steady-state values.	Group Polarization	Group Polarization
C2	For $k \geq \hat{k}$, $A_{out,1}[k], B_1[k], B_2[k] < 0$, $A_{out,2}[k] \geq 0$. $B_1[k], B_2[k]$ are decreasing functions of k and settle to steady-state values.	Reactance	Conformity
C3	For $k \geq \hat{k}$, $A_{out,2}[k], B_1[k], B_2[k] < 0$, $A_{out,1}[k] \geq 0$. $B_1[k], B_2[k]$ are decreasing functions of k and settle to steady-state values.	Conformity	Reactance
D1	For $k \geq \hat{k}$, $A_{out,2}[k], B_2[k] \geq 0$, $A_{out,1}[k], B_1[k] < 0$. $B_2[k]$ increases and is unbounded and $B_1[k]$ decreases and is unbounded.	Reactance	Reactance

Table 5.1: Description of two-person system responses and the related psychological trends.

(b) $\Delta A^{OJ}_{out,i}[k] = 0$.

Proof. See Appendix E.1. ☐

Lemma 5.1 dictates that if the zero-input, two-person system begins in region V, then both attitudes remain constant for $k \geq 0$. (Moreover, a similar argument holds if the system begins in region I.) In the case when both people in the zero-input system experience no dissonance or overjustification pressures, the expressions for $B_1[k]$ and $B_2[k]$ can be simplified:

Lemma 5.2. *For the zero-input two-person system with the initial conditions given by (4.26), if for $k \geq 0$ and $i = 1, 2$*

(i) $\Delta A^{CD}_{out,i}[k] = 0$, *and*

(ii) $\Delta A^{OJ}_{out,i}[k] = 0$,

then, for $k \geq 1$,

$$B_1[k] = A_{o1}\left(\frac{1-(\mu_{31}\mu_{32})^{1+\lfloor\frac{k}{2}\rfloor}}{1-(\mu_{31}\mu_{32})}\right) + \mu_{32}A_{o2}\left(\frac{1-(\mu_{31}\mu_{32})^{1+\lfloor\frac{k-1}{2}\rfloor}}{1-(\mu_{31}\mu_{32})}\right) \tag{5.2}$$

$$B_2[k] = \mu_{31}A_{o1}\left(\frac{1-(\mu_{31}\mu_{32})^{1+\lfloor\frac{k-1}{2}\rfloor}}{1-(\mu_{31}\mu_{32})}\right) + A_{o2}\left(\frac{1-(\mu_{31}\mu_{32})^{1+\lfloor\frac{k}{2}\rfloor}}{1-(\mu_{31}\mu_{32})}\right). \tag{5.3}$$

Proof. See Appendix E.2. ☐

By Lemma 5.1, if the system begins in region V, then the conditions of Lemma 5.2 apply; therefore, the above lemma provides simplified expressions for $B_1[k]$ and $B_2[k]$ when the system begins in region V. From these expressions, the system response is determined.

Along with Assumption 4.7, (5.2) and (5.3) imply that the zero-input response of initial region V is stable in the sense that $B_1[k]$ and $B_2[k]$ converge and all other signals remain bounded as k tends to infinity. Moreover, Assumption 4.7, (5.2) and (5.3) imply that $B_1[k] < 0$ and $B_2[k] < 0$ for $k \geq 0$. Not only are both behaviours negative for all $k \geq 0$, it can be shown that they are decreasing functions of k for $k \geq 0$:

Lemma 5.3. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.1), then*

(a) $A_{out,1}[k] < 0$, $A_{out,2}[k] < 0$, $B_1[k] < 0$, $B_2[k] < 0$, *for all $k \geq 0$, and*

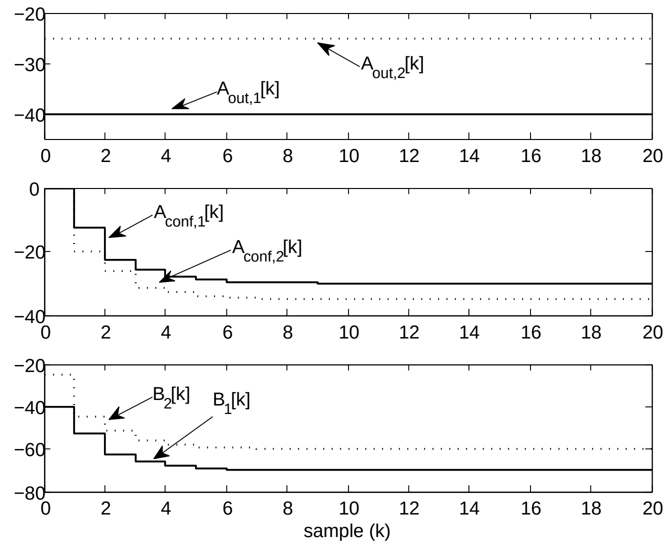

Figure 5.1: Simulation results of the two-person system, with the initial conditions given in (4.26), parameter values in Table 4.1, beginning in region V with no reward applied to the system. Person one's initial attitude is $A_{o1} = -40$ *attitude units*, whereas person two's initial attitude is $A_{o2} = -25$ *attitude units*. The simulation results confirm Lemma 5.3 since all attitudes and behaviours remain negative for $k \geq 0$, while both behaviours are decreasing functions of k for $k \geq 0$, eventually converging to steady-state values.

(b) $B_1[k]$ and $B_2[k]$ are decreasing functions of k, converging as k tends to infinity.

Proof. See Appendix E.3. □

The system response given by Lemma 5.3 is the same as that described by response type C1 in Table 5.1, where $\hat{k} = 0$; consequently, if the system begins in initial region V, then its response is of type C1 (both in the short-term and in the long-term). From Table 5.1, response type C1 characterizes group polarization for person one and person two. This is consistent with the prediction given by Table 4.3, which indicates that if the system begins in region V, then the short-term psychological trend it should exhibit is the group polarization phenomenon. To illustrate the trends of Lemma 5.3, simulation results are shown in Figure 5.1.

5.1.2 Zero-Input Response Analysis: Region VI

Now, suppose the system begins in region VI, which is defined by the following conditions on the two initial attitudes:

$$\begin{aligned} &A_{o1} < 0, A_{o2} \geq 0, \\ &A_{o2} < \mu_{31}|A_{o1}|. \end{aligned} \tag{5.4}$$

Unlike in the case of initial region V, for initial region VI dissonance pressure may arise and, thus, the expressions for $B_1[k]$ and $B_2[k]$ in Lemma 5.2 are not applicable. Instead, more general equations are needed. Iterating through the system equations yields

$$B_1[k] = \begin{cases} A_{o1} & \text{if } k = 0 \\ B_1[0] + \mu_{32}A_{o2} & \text{if } k = 1, \\ B_1[1] + K_{11}P_1^{CD}[1] + \mu_{31}\mu_{32}A_{o1} & \text{if } k = 2, \\ B_1[k-1] + K_{11}P_1^{CD}[k-1] + & \\ \quad \mu_{32}\left[\sum_{i=1}^{\lfloor\frac{k-1}{2}\rfloor} K_{12}P_2^{CD}[k-2i]\left(\mu_{31}\mu_{32}\right)^{i-1}\right. & \text{if } k \geq 3, \\ \quad \left. + \sum_{i=1}^{\lfloor\frac{k-2}{2}\rfloor} K_{11}P_1^{CD}[k-2i-1]\mu_{31}^{i}\mu_{32}^{i-1} + \mu_{31}^{\lfloor\frac{k-2}{2}\rfloor+1}\mu_{32}^{\lfloor\frac{k-1}{2}\rfloor}A_{oi}\right] & \end{cases} \tag{5.5}$$

$$B_2[k] = \begin{cases} A_{o2} & \text{if } k = 0 \\ B_2[0] + \mu_{31}A_{o1} & \text{if } k = 1, \\ B_2[1] + K_{12}P_2^{CD}[1] + \mu_{31}\mu_{32}A_{o2} & \text{if } k = 2, \\ B_2[k-1] + K_{12}P_2^{CD}[k-1] + & \\ \quad \mu_{31}\left[\sum_{i=1}^{\lfloor\frac{k-1}{2}\rfloor} K_{11}P_1^{CD}[k-2i]\left(\mu_{31}\mu_{32}\right)^{i-1}\right. & \text{if } k \geq 3, \\ \quad \left. + \sum_{i=1}^{\lfloor\frac{k-2}{2}\rfloor} K_{12}P_2^{CD}[k-2i-1]\mu_{32}^{i}\mu_{31}^{i-1} + \mu_{32}^{\lfloor\frac{k-2}{2}\rfloor+1}\mu_{31}^{\lfloor\frac{k-1}{2}\rfloor}A_{oj}\right] & \end{cases} \tag{5.6}$$

where

$$\begin{cases} j = 1, i = 2, & \text{if } k \text{ is odd} \\ j = 2, i = 1 & \text{if } k \text{ is even.} \end{cases}$$

Along with Assumption 4.10 and the approximation that $|\mu_{31}^n\mu_{32}^m| \approx 0$ for any m and n with $m+n \geq 2$, these two equations can be used to deduce the signs of $B_1[k]$ and $B_2[k]$ for $k \geq 0$. Once the signs of $B_1[k]$ and $B_2[k]$ are known for $k \geq 0$, the signs of $P_1^{CD}[k]$ and $P_2^{CD}[k]$ are easily obtained from Assumption 4.10. Specifically, if $B_i[k] < 0$ then $P_i^{CD}[k] \leq 0$;

otherwise $P_i^{CD}[k] \geq 0$. (It is possible for there to be zero dissonance pressure at some sample and, thus, this possibility is included.) By Assumption 4.2, dissonance pressure is the only pressure driving attitude change and, therefore, the sign of $B_i[k]$ indicates the attitude change direction of $A_{out,i}[k]$.

When the system begins in region VI, the initial attitude of person one is strong relative to person two and, therefore, the conformity pressure of person one should dominate. In other words, the behaviour of person one is strong enough such that it should drive person two's behaviour negative. From (5.5) and (5.6), it can be shown that this is the case. In fact, not only does person two's behaviour eventually become negative due to the negative conformity pressure exerted by person one, it becomes negative at $k = 1$ and remains negative for $k \geq 1$. Additionally, person one's behaviour remains negative for $k \geq 0$. The lemma below formalizes this conclusion.

Lemma 5.4. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.4), and the approximation* $|\mu_{31}^n \mu_{32}^m| \approx 0$ *for any* m *and* n *with* $m + n \geq 2$ *is used, then for* $k \geq 1$

$$\begin{aligned} B_1[k] &< 0 \\ B_2[k] &< 0. \end{aligned}$$

Proof. See Appendix E.4. □

The above lemma states that if the system begins in region VI, then both behaviours are negative for $k \geq 1$, implying (from the previous discussion) that $A_{out,i}[k]$ is a non-increasing function of k for $k \geq 1$ and $i = 1, 2$. For $A_{out,1}[k]$, this conclusion guarantees the attitude of person one remains negative for $k \geq 0$. For $A_{out,2}[k]$, this conclusion means that the attitude of person two may become negative at some sample, k^*, or by Lemma A.1, converge to some non-negative constant as k tends to infinity.

Although Lemma 5.4 says that both behaviours remain negative for $k \geq 1$, more information about the response of these two signals can be obtained by considering (5.5) and (5.6). In particular, from (5.5) and the approximation on $|\mu_{31}^n \mu_{32}^m|$, $B_1[k]$ can be approximated as

$$B_1[k] \approx B_1[k-1] + K_{11} P_1^{CD}[k-1] + \mu_{32} K_{12} P_2^{CD}[k-2].$$

for $k \geq 3$. From Lemma 5.4, for $k \geq 3$, $B_1[k-1] < 0$ and $B_2[k-2] < 0$, thus from (4.29) $P_1^{CD}[k-1] < 0$ and $P_2^{CD}[k-2] < 0$. As a result, $B_1[k] < B_1[k-1]$ for $k \geq 3$. Similarly, $B_2[k] < B_2[k-1]$ for $k \geq 3$. Combining these two conclusions implies that for $k \geq 3$, $B_1[k]$ and $B_2[k]$ are decreasing functions of k. Not only are these two signals decreasing functions of k, it can also be shown that each converges as k tends to infinity. The following lemma summarizes the response of the two-person system beginning in region VI.

Lemma 5.5. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.4), and the approximation* $|\mu_{31}^n \mu_{32}^m| \approx 0$ *for any* m *and* n *with* $m + n \geq 2$ *is used, then*

(a) $A_{out,1}[k] < 0$, $B_1[k] < 0$ *and* $B_2[k] < 0$, *for* $k \geq 1$,

(b) $A_{out,2}[k] \geq 0$ *for* $k \geq 0$, *or there exists a* k^* *such that* $A_{out,2}[k] \geq 0$ *for* $0 \leq k < k^*$ *and* $A_{out,2}[k] < 0$ *for all* $k \geq k^*$, *and*

(c) $B_1[k]$ *and* $B_2[k]$ *are non-increasing functions of* k *for* $k \geq 3$, *converging as* k *tends to infinity.*

Proof. See Appendix E.5. □

The system response given by Lemma 5.5 is the same as that described by response types C1 and C2 in Table 5.1. The case when $A_{out,2}[k] \geq 0$ for $k \geq 0$ corresponds with response type C2, where $\hat{k} = 3$, whereas the case when $A_{out,2}[k] < 0$ for $k \geq k^*$ corresponds with response type C1 where $\hat{k} = k^*$. From Table 5.1, response type C2 characterizes reactance for person one and conformity for person two. This is consistent with the prediction given in Table 4.3, which indicates that if the system begins in region VI, then it should exhibit reactance for person one and conformity for person two. Additionally, although response type C1, which corresponds to group polarization, does not match the expected psychological trend give in Table 4.3, recall that Table 4.3 gives the *short-term* expected psychological trends for each initial region. In the case when $A_{out,2}[k] < 0$ for $k \geq k^*$, it's also true that $A_{out,2}[k] \geq 0$ for $0 \leq k < k^*$, and therefore, over this range, person one exhibits reactance and person two exhibits conformity, which is consistent with the prediction given in Table 4.3. To illustrate the trends of Lemma 5.5, simulation results are shown in Figure 5.2.

5.1.3 Zero-Input Response Analysis: Region VIII

Before continuing to initial region VII, initial region VIII is considered. This region is defined by the following initial attitudes:

$$A_{o1} < 0, A_{o2} \geq 0,$$
$$|A_{o1}| \leq \mu_{32} A_{o2}. \tag{5.7}$$

The initial signs of A_{o1} and A_{o2} are the same as those of region VI. Furthermore, the above relationship between the two attitudes is symmetric (across the $A_{o1} = -A_{o2}$ axis) with their relationship in region VI, given by (5.4). This similarity between the initial attitude

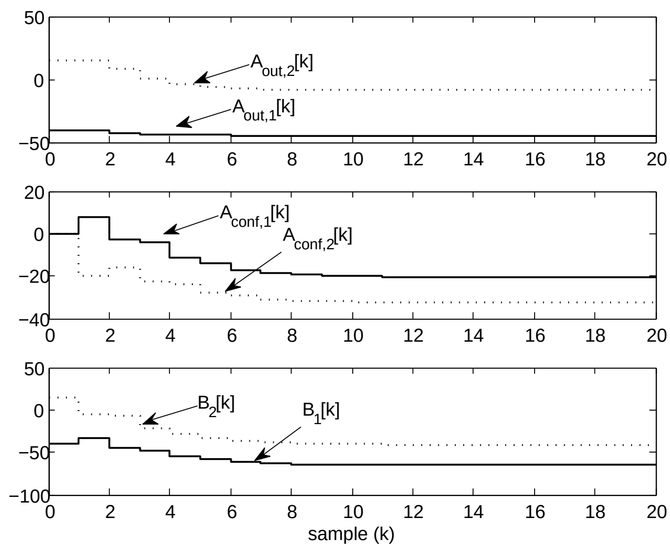

Figure 5.2: Simulation results of the two-person system, with the initial conditions given in (4.26), parameter values in Table 4.1, beginning in region VI with no reward applied to the system. Person one's initial attitude is $A_{o1} = -40$ *attitude units*, whereas person two's initial attitude is $A_{o2} = 15$ *attitude units*, thus satisfying the relationship given in (5.4). This simulation confirms Lemmas 5.4 and 5.5 since both behaviours are negative, decreasing functions of k for $k \geq 1$, converging as $k \to \infty$, while $A_{out,1}[k] < 0$ for $k \geq 0$. In the above simulation, $A_{out,2}[k]$ becomes negative at $k = 4$, resulting in response type C1 where $\hat{k} = 4$.

relationships of regions VI and VIII suggests a symmetry in their responses. Since, in initial region VI, the attitude of person one is strong relative to person two, person two conforms to the behaviour of person one. On the other hand, in initial region VIII, the attitude of person two is strong relative to person one and thus by symmetry, person one should conform to the behaviour of person two and therefore their behaviour should become positive, suggesting a type A response. From Table 5.1, apart from the attitude and behaviour signs, response type A1 is identical to response type C1. Similarly, response type A3 is symmetric (across the $A_{o1} = -A_{o2}$ axis) with type C2. Consequently, the following lemma describes the system response when the initial attitudes are in region VIII, which is symmetric with the zero-input system response of initial region VI, given in Lemma 5.5.

Lemma 5.6. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.7), and the approximation* $|\mu_{31}^n \mu_{32}^m| \approx 0$ *for any* m *and* n *with* $m + n \geq 2$ *is used, then*

(a) $A_{out,2}[k] \geq 0$*,* $B_1[k] \geq 0$ *and* $B_2[k] \geq 0$*, for* $k \geq 1$*,*

(b) $A_{out,1}[k] < 0$ *for* $k \geq 0$*, or there exists a* k^* *such that* $A_{out,1}[k] < 0$ *for* $0 \leq k < k^*$ *and* $A_{out,1}[k] \geq 0$ *for all* $k \geq k^*$*, and*

(c) $B_1[k]$ *and* $B_2[k]$ *are increasing functions of* k *for* $k \geq 3$*, converging as* k *tends to infinity.*

Proof. The lemma follows from Lemma 5.5 and symmetry with region VI. □

Similar to Lemma 5.5, Lemma 5.6 is also the same as two responses described in Table 5.1. In particular, the case when $A_{out,1}[k] < 0$ for $k \geq 0$ corresponds with response type A3, where $\hat{k} = 3$, while the case when $A_{out,1}[k] \geq 0$ for $k \geq k^*$ corresponds with response type A1, where $\hat{k} = k^*$. As before, Table 5.1 indicates the psychological phenomenon exhibited by each response type. Using the same arguments as initial region VI, the response of the zero-input system beginning in region VIII can be shown to be consistent with the short-term expected psychological trends given in Table 4.3. To illustrate the trends of Lemma 5.6, simulation results are shown in Figure 5.3. These simulation results also illustrate further the point that the system responses of initial regions VI and VIII are symmetric.

5.1.4 Zero-Input Response Analysis: Region VII

Finally, the response of the system beginning in region VII is studied. This region is the most complicated and, perhaps, the most interesting region to study, as very different types

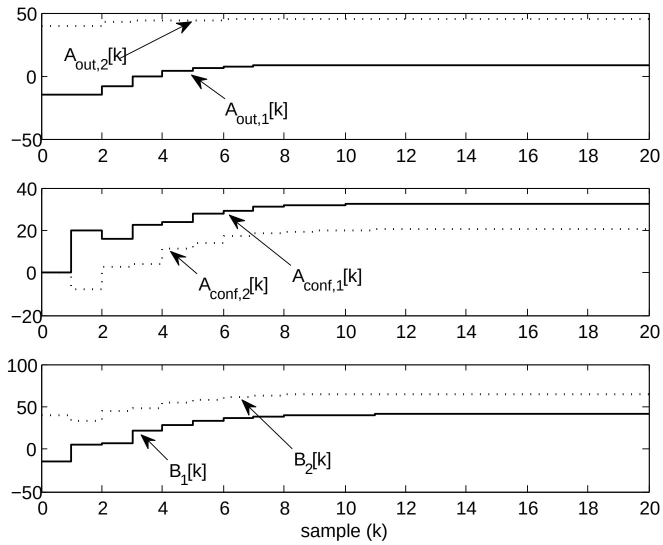

Figure 5.3: Simulation results of the two-person system, with the initial conditions given in (4.26), parameter values in Table 4.1, beginning in region VIII with no reward applied to the system. Person one's initial attitude is $A_{o1} = -15$ *attitude units*, whereas person two's initial attitude is $A_{o2} = 40$ *attitude units*, thus satisfying the relationship given in (5.7). The initial conditions used for this simulation are symmetric with those of the simulation for region VI. The magnitudes of person one's signals in this simulation are the same as the magnitudes of the corresponding signals of person two in the simulation results given in Figure 5.2, while their sign is different (and vice versa for person two). Therefore, this simulation also provide evidence supporting the argument that the zero-input system response is symmetric across the $A_{o1} = -A_{o2}$ axis. Moreover, this simulation confirms Lemma 5.6 since both behaviours and $A_{out,2}[k]$ are positive for $k \geq 1$. At $k = 4$, $A_{out,1}[k]$ becomes positive, resulting in response type A1 where $\hat{k} = 4$.

of system response can occur. In this region, the two attitudes have approximately the same strength, but differ in sign. Psychology literature suggests that in this case, people can either react or conform. If one person conforms, then the other person likely reacts. However, as will be demonstrated in this subsection, our model predicts that it is also possible for both people to react, with both people's attitudes and behaviours becoming stronger.

For initial region VII, defined by the following initial attitudes,

$$A_{o1} < 0, A_{o2} \geq 0,$$
$$\mu_{32} A_{o2} < |A_{o1}| \leq \frac{A_{o2}}{\mu_{31}}, \tag{5.8}$$

the behaviour can be expressed by (5.5) and (5.6). From the above relationship between the initial attitudes, it can be shown (from (5.5) and (5.6)) that for $0 \leq k \leq 2$, $B_1[k] < 0$ and $B_2[k] \geq 0$. From (4.29), negative dissonance pressure arises from a negative behaviour, and positive dissonance pressure arises from a positive behaviour. By Assumption 4.2, attitude change is driven only by dissonance pressure and, therefore, for $0 \leq k \leq 2$, $A_{out,1}[k]$ is a decreasing function of k, whereas $A_{out,2}[k]$ is an increasing function of k. At this point, the system response depends on the relative size of K_{11} and K_{12}; thus three possible situations arise for $k = 3$:

1. K_{11} and K_{12} are similar enough to ensure $B_1[3] < 0$ and $B_2[3] \geq 0$;
2. K_{11} is small relative to K_{12} and therefore, $B_1[3] \geq 0$ and $B_2[3] \geq 0$; and
3. K_{12} is small relative to K_{11} and therefore, $B_1[3] < 0$ and $B_2[3] < 0$.

For the last two of these three situations, it is possible to show that after $k = 3$, each behaviour signal maintains its sign for $k \geq 3$, as summarized in the lemma below.

Lemma 5.7. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.8), and the approximation $|\mu_{31}^n \mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$ is used, then*

(a) if $B_1[3] \geq 0$ and $B_2[3] \geq 0$, then for $k \geq 3$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$; whereas

(b) if $B_1[3] < 0$ and $B_2[3] < 0$, then for $k \geq 3$, $B_1[k] < 0$ and $B_2[k] < 0$.

Proof. See Appendix E.6. □

Lemma 5.7 gives two possible responses for the behaviour signals. The analysis for the situation in which $B_1[3] < 0$ and $B_2[3] \geq 0$ is more complicated, yet simulations suggest that in this case, $B_1[k] < 0$ and $B_2[k] \geq 0$ for $k \geq 3$. This detail is presented after analysing the two above known system responses.

First, consider the case when $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k \geq 3$. Since $A_{o2} \geq 0$ and $B_2[k] \geq 0$ for $k \geq 0$, from (4.29), $P_2^{CD}[k] \geq 0$ for $k \geq 0$; thus, for $k \geq 0$, $A_{out,2}[k] \geq 0$ and is a non-decreasing function of k. Moreover, since $B_1[k] \geq 0$ for $k \geq 3$, $P_1^{CD}[k] \geq 0$ for $k \geq 3$. If, for all $k \geq 3$, $A_{out,1}[k] < 0$, then $P_1^{CD}[k] > 0$, otherwise, all cognitions are positive and no dissonance pressure arises; hence, either $A_{out,1}[k]$ is an increasing function of k for $k \geq 3$, remaining negative (thus converging) as k tends to infinity, or $A_{out,1}[k]$ becomes positive at some sample $k = k^*$ and remains positive for $k \geq k^*$. This response is the same as that described by Lemma 5.6, which corresponds with initial region VIII. Therefore, if $B_1[3] \geq 0$ and $B_2[3] \geq 0$, the system response is similar to that of initial region VIII, as given in the lemma below.

Lemma 5.8. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.8), $B_1[3] \geq 0$, $B_2[3] \geq 0$, and the approximation $|\mu_{31}^n \mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$ is used, then*

(a) $B_1[k] \geq 0$, $B_2[k] \geq 0$ *and* $A_{out,2}[k] \geq 0$ *for* $k \geq 3$;

(b) $A_{out,1}[k] < 0$ *for* $k \geq 0$, *or there exists some* k^* *such that* $A_{out,1}[k] < 0$ *for* $0 \leq k < k^*$ *and* $A_{out,1}[k] \geq 0$ *for* $k \geq k^*$; *and*

(c) $B_1[k]$ *and* $B_2[k]$ *converge as* $k \to \infty$.

Proof. See Appendix E.7. □

Similar to Lemma 5.6, the system response given by Lemma 5.8 is the same as that described by response types A1 and A3 in Table 5.1. It can be argued (as before) that response given by Lemma 5.8 is consistent with the short-term expected psychological trends given in Table 4.3. To illustrate the trends of Lemma 5.8, simulation results are shown in Figure 5.4.

Now consider the second result of Lemma 5.7, where $B_1[k] < 0$ and $B_2[k] < 0$ for $k \geq 3$. By symmetry across the $A_{o1} = -A_{o2}$ axis, the previous argument can be used to show that the system response is given by the following lemma.

Lemma 5.9. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.8), $B_1[3] < 0$, $B_2[3] < 0$, and the approximation $|\mu_{31}^n \mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$ is used, then*

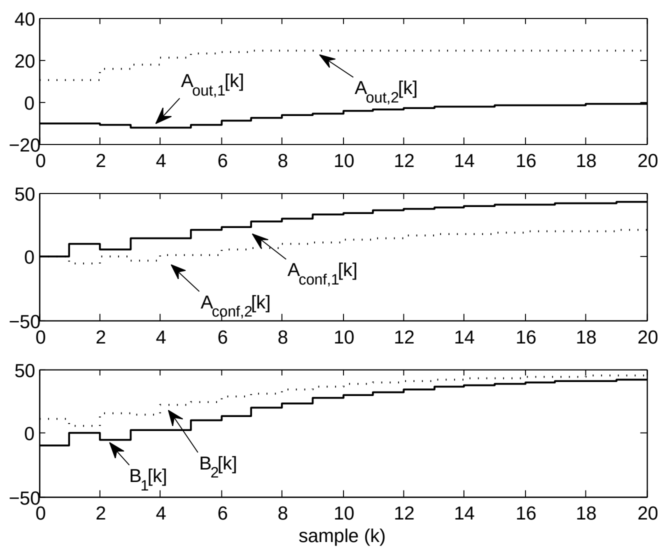

Figure 5.4: Simulation results of the two-person system, with the initial conditions given in (4.26) beginning in region VII with no reward applied to the system. With the exception of μ_{32} and K_{11}, parameter values used are those given in Table 4.1. Person one's initial attitude is $A_{o1} = -10$ *attitude units*, whereas person two's initial attitude is $A_{o2} = 10.5$ *attitude units*. Moreover, $\mu_{32} = 0.95$, thus the relationship given in (5.8) is satisfied. Finally, $K_{11} = 5$ (thus K_{11} is small relative to $K_{12} = 30$). At $k = 3$, both behaviours are positive and remain positive for $k \geq 3$, thus confirming Lemma 5.7. Finally, the simulation results show both attitudes are non-decreasing functions of k for $k \geq 3$ and both behaviours converge as k tends to infinity, which supports the results of Lemma 5.8. Due to $A_{out,1}[k]$ remaining negative for all $k \geq 0$, the system response is of type A3 where $\hat{k} = 3$.

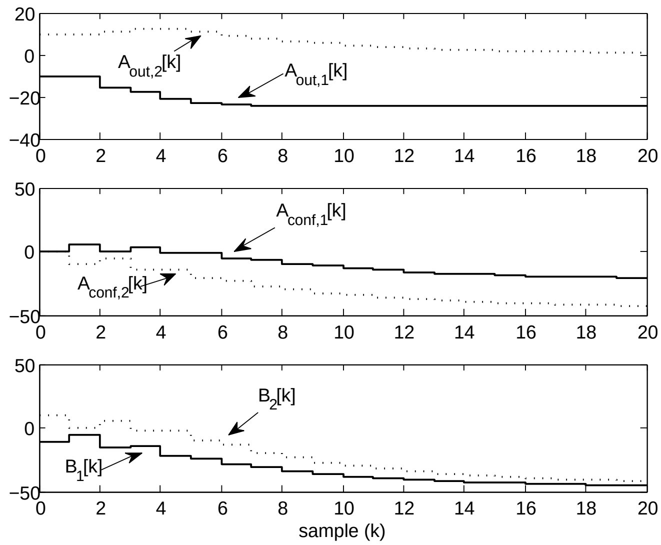

Figure 5.5: Simulation results of the two-person system, with the initial conditions given in (4.26) beginning in region VII with no reward or social pressure applied to the system. With the exception of μ_{31} and K_{12}, parameter values used are those given in Table 4.1. Person one's initial attitude is $A_{o1} = -10.5$ *attitude units*, whereas person two's initial attitude is $A_{o2} = 10$ *attitude units*. Moreover, $\mu_{31} = 0.95$, thus the relationship given in (5.8) is satisfied. Finally, $K_{12} = 5$. At $k = 3$, both behaviours are negative and therefore, the simulation results confirm Lemma 5.7, as both behaviours remain negative for $k \geq 3$. Finally, the simulation results support Lemma 5.9 because they show that both attitudes are decreasing functions of k and both behaviours converge as k tends to infinity. Due to $A_{out,2}[k]$ remaining positive for all $k \geq 0$, the system response is of type C2 where $\hat{k} = 3$.

(a) $B_1[k] < 0$, $B_2[k] < 0$ *and* $A_{out,1}[k] < 0$ *for* $k \geq 3$;

(b) $A_{out,2}[k] \geq 0$ *for* $k \geq 0$, *or there exists some* k^* *such that* $A_{out,2}[k] \geq 0$ *for* $0 \leq k < k^*$ *and* $A_{out,2}[k] < 0$ *for* $k \geq k^*$; *and*

(c) $B_1[k]$ *and* $B_2[k]$ *converge as* $k \to \infty$.

Proof. The lemma follows from symmetry with Lemma 5.8. □

The system response given by Lemma 5.9 is the same as that described by response types C1 and C2 in Table 5.1. It can be argued (as before) that response given by Lemma 5.9 is consistent with the short-term expected psychological trends given in Table 4.3. The simulation results given in Figure 5.5 illustrate the results of Lemma 5.9 and, again, highlight the symmetric trend exhibited by the system.

Finally, the situation in which $B_1[3] < 0$ and $B_2[3] \geq 0$ is discussed. Simulations suggest that when this situation arises, $B_1[k] < 0$ and $B_2[k] \geq 0$ for $k \geq 3$. Suppose it is true that $B_1[3] < 0$ and $B_2[3] \geq 0$ implies $B_1[k] < 0$ and $B_2[k] \geq 0$ for $k \geq 3$; then, both people experience dissonance pressure due to the conformity pressure exerted by the other individual. From (4.29), these dissonance pressures cause $A_{out,1}[k]$ to decrease and $A_{out,2}[k]$ to increase. Since $A_{out,1}[k]$ is initially negative and is a decreasing function of k, it follows that it remains negative. Likewise, $A_{out,2}[k] \geq 0$ for $k \geq 0$. Since dissonance pressure always exists for each person, neither attitude converges as k tends to infinity and as a result $A_{out,1}[k] \to -\infty$ and $A_{out,2}[k] \to \infty$ as $k \to \infty$ and, hence, neither behaviour converges. Therefore, the conjecture below is presented to summarize the observed system behaviour in this situation.

Conjecture 5.1. *For the zero-input two-person system with the initial conditions given by (4.26), if the initial attitudes satisfy (5.8),* $B_1[3] < 0$, $B_2[3] \geq 0$, *and the approximation* $|\mu_{31}^n \mu_{32}^m| \approx 0$ *for any* m *and* n *with* $m + n \geq 2$ *is used, then*

(a) $B_1[k] < 0$ *and* $B_2[k] \geq 0$ *for* $k \geq 0$*;*

(b) $A_{out,1}[k] < 0$ *and* $A_{out,2}[k] \geq 0$ *for* $k \geq 0$*; and*

(c) $B_1[k] \to -\infty$ *and* $B_2[k] \to \infty$ *as* $k \to \infty$*.*

The system response given in the above conjecture is the same as that described by response type D1 in Table 5.1. Interesting to note is that both people exhibit reactance, since the attitude and behaviour of each person maintains its original sign and becomes stronger as k tends to infinity. The simulation shown in Figure 5.6 provides evidence supporting the results of Conjecture 5.1.

The results of the initial region VII analysis are interesting because they confirm that both reactance and compliance may arise when conformity pressure is experienced. These results are consistent with psychology literature related to conformity, as are the possible responses of initial regions V, VI and VIII. Although, when developing our model, the dynamic effects of conformity were not considered, it is interesting that our system exhibits these trends through dissonance pressure. This result suggests that cognitive dissonance may be a key underlying mechanism that is responsible for the trends discovered by psychologists.

5.2 Open-Loop Investigation

To begin investigating whether or not there exists a reward sequence, $R_1[k]$, $0 \leq k < \infty$, such that $B_1[k] \geq B_d$ as $k \to \infty$, an open-loop control strategy is first studied. The

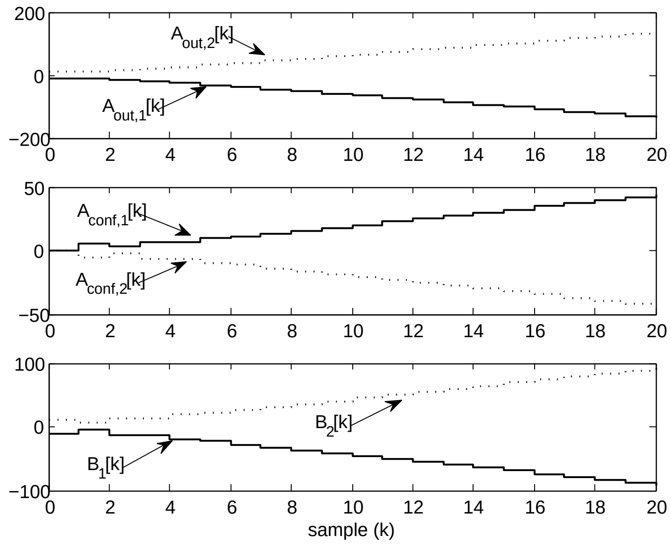

Figure 5.6: Simulation results of the two-person system, with the initial conditions given in (4.26), parameter values in Table 4.1, beginning in region VII with no reward applied to the system. Person one's initial attitude is $A_{o1} = -10$ *attitude units*, whereas person two's initial attitude is $A_{o2} = 10.5$ *attitude units*, thus satisfying the relationship given in (5.8). This simulation provides evidence to support Conjecture 5.1, because $B_1[3] < 0$, $B_2[3] \geq 0$, and all attitudes and behaviours maintain their initial signs for $k \to \infty$. Moreover, the behaviours tend to positive/negative infinity as k tends to infinity, resulting in response type D1 for $k \geq 0$.

dynamics of the two-person system are more complicated than the one-person system as demonstrated by the analysis in the last section. Nevertheless, the one-person open-loop investigation in Section 3.1 provides some insight into where the two-person open-loop investigation could begin. The first open-loop control strategy considered for the one-person system was an impulse reward, which was shown to not necessarily be able to meet the control objective. As the two-person system contains the dynamics of the one-person system, and additional dynamics due to the conformity feedback loop, the impulse-reward control strategy is not even considered in this analysis. Instead, motivated by the success of the step-reward control strategy presented in Section 3.1.2, a step reward is examined here. Since person one is the only person being offered a reward, the reward sequence, $R_1[k]$, is the control signal, and is simply denoted $R[k]$ (see Figure 4.1).

Applying a step reward to the two-person system means that the child's parents offer him a fixed amount of money each day, with the hope that the child *eventually* produces the desired behavioural intent, denoted here by some $B_d > 0$. Moreover, this child and his friend influence each other's behaviour through conformity pressure. The problem is to determine whether or not there exists some $R_o > 0$ and control signal

$$R[k] = R_o,\ k \geq 0 \tag{5.9}$$

such that $B_1[k] \geq B_d$ as k tends to infinity. From (4.27) this control signal generates a reward attitude signal for $k \geq 1$ given by

$$A_{rew,1}[k] = \mu_1 R_o. \tag{5.10}$$

Before proceeding with the analysis to find an R_o that meets the control objective, an easier problem is studied. Since the control objective is to drive $B_1[k] \geq B_d$ as k tends to infinity, it is useful to know whether or not it is possible to at least drive person one's behaviour positive. Instead of finding an R_o that drives $B_1[k] \geq 0$ as $k \to \infty$, the results of the zero-input analysis motivate finding an R_o that drives $B_1[k] \geq 0$ for $k \geq 1$. In particular, the zero-input analysis from the previous section demonstrates that the most complicated dynamics occur within the first few samples, when the signs of $B_1[k]$ and $B_2[k]$ are switching. Knowing the signs of $B_1[k]$ and $B_2[k]$ is one of the more useful pieces of information for understanding the system's response. By driving $B_1[k] \geq 0$ immediately, the analysis is simplified, as the reward's effect on the first few samples of each behaviour can be considered explicitly, and then, through induction, its effect on the remaining samples of each behaviour can be determined. Thus, the easier problem of finding an R_o such that $B_1[k] \geq 0$ for $k \geq 1$ is studied, providing useful tools for solving the main control objective.

5.2.1 Simplified Step-Reward Controller Design

This subsection investigates whether or not there exists an R_o such that a controller of the form given in (5.9) can drive $B_1[k] \geq 0$ for $k \geq 1$. To find such an R_o, a general strategy is developed through consideration of the two-person system's zero-input response, studied in Sections 4.4 and 5.1, and the one-person system's step response, studied in Section 3.1.2.

Recall from Section 4.4 that the relationship between A_{o1} and A_{o2} divides the two-person system into eight different initial operating regions, summarized in Table 4.2. From the initial conditions given in (4.26), $A_{o1} < 0$ and therefore, similar to the previous section, only initial regions V, VI, VII and VIII are considered for this analysis.

The step response of the *one-person* system is now reviewed. In the one-person system analysis, the system's step response is divided into two cases: the first case considers a reward that is insufficient at producing a positive behaviour at $k = 1$, whereas the second case considers a reward that is sufficiently large to accomplish this result. From Theorem 3.3, an insufficient reward has the unintended consequence of driving behaviour to negative infinity as k tends to infinity. This consequence is due to $A_{out}[k]$ decreasing to reduce the dissonance pressure that arises in the system. If, on the other hand, the reward is sufficiently large, then $A_{out}[k]$ is an increasing function of k until $A_{out}[k]$ becomes positive, at which point, it remains positive. Therefore, a sufficiently large reward ensures $B[k] \geq 0$ for $k \geq 1$ in the one-person system.

In the case of the two-person system, a reward that is insufficient at driving $B_1[1] \geq 0$ also produces a decreasing $A_{out,1}[k]$ through the exact same mechanism as the one-person system, namely dissonance pressure reduction. This decreasing $A_{out,1}[k]$ may drive $B_1[k] < 0$ as k tends to infinity. To see why, consider the expression for $B_1[k]$ when a controller of the form (5.9) is applied to the two-person system. From (4.24), for $k \geq 1$, $B_1[k]$ can be expressed as

$$B_1[k] = A_{out,1}[k] + \mu_{32} B_2[k-1] + \mu_{11} R_o. \tag{5.11}$$

Since an insufficient reward causes $A_{out,1}[k]$ to decrease, and since the reward is constant, the only remaining term that could drive $B_1[k]$ positive is $B_2[k-1]$. However, in most cases, the conformity pressure is insufficient for this purpose. The simulation results in Figure 5.7 support this conclusion. In this simulation, a reward that is unable to drive $B_1[1] \geq 0$ is applied to the one-person system, causing $A_{out,1}[k]$ to decrease. Also, a positive conformity pressure is applied by person two, but it is not strong enough to drive $B_1[k]$ positive; thus, $B_1[k] < 0$ for $k \geq 0$. The simulation results support the conclusion that if a step reward is applied to the two-person system and $B_1[1] < 0$, then it is unlikely that, eventually, $B_1[k]$ will become positive. This conclusion provides further support for the decision to consider the problem of *immediately* producing a positive $B_1[k]$ rather than the problem of *eventually* producing a positive $B_1[k]$.

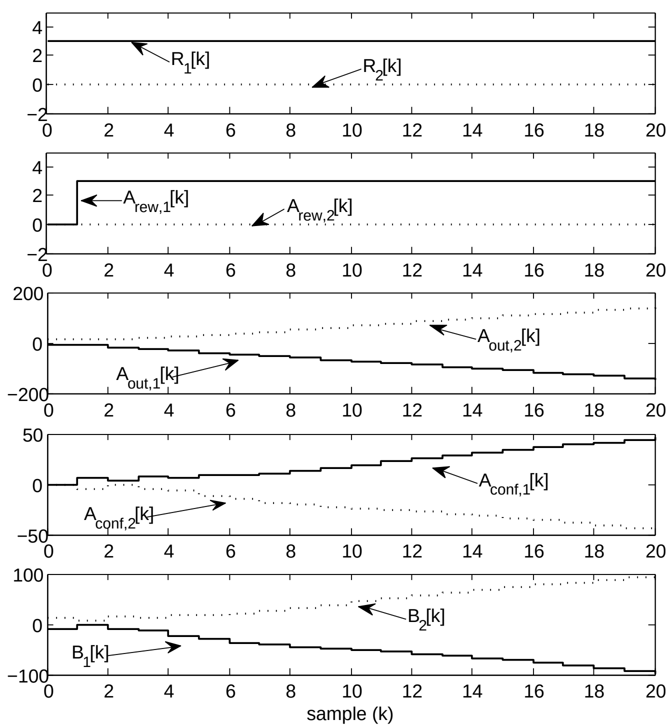

Figure 5.7: A controller of the form (5.9) is applied to the two-person system with the initial conditions given in (4.26), parameter values in Table 4.1, where $A_{o1} = -10$ *attitude units*, $A_{o2} = 12$ *attitude units* and $R_o = \$3$. This reward is insufficient at driving $B_1[1] \geq 0$ and, therefore, $A_{out,1}[k]$ decreases. Even though the conformity pressure exerted on person one is positive, i.e., $A_{conf,1}[k] \geq 0$, it is not sufficient to overcome the decreasing $A_{out,1}[k]$, and, as a result, $B_1[k] < 0$ for $k \geq 0$.

On the other hand, if the reward causes $B_1[1] \geq 0$, then $A_{out,1}[k]$ increases. Should this increase cause $A_{out,1}[k] \geq 0$, overjustification pressure may arise (due to the reward), causing $A_{out,1}[k]$ (and thus $B_1[k]$) to decrease. However, similar to the one-person system, overjustification pressure is able to decrease $A_{out,1}[k]$ only to a certain point: it is not able to drive $A_{out,1}[k] < 0$, as summarized in the lemma below.

Lemma 5.10. *If, for the two-person system, there is a $\bar{k}$ such that*

(i) $A_{rew,1}[\bar{k}] \geq 0$,

(ii) $A_{out,1}[\bar{k}] \geq 0$, *and*

(iii) $A_{conf,1}[k] \geq 0$ *for* $k \geq \bar{k}$,

then for $k \geq \bar{k}$, $A_{out,1}[k]$ is a non-increasing function of k, lower bounded by zero.

Proof. See Appendix E.8. □

The above lemma implies that if $A_{out,1}[k]$ decreases due to overjustification pressure, then a sufficiently large reward can be applied to counteract any negative effect this decreasing internal attitude has on $B_1[k]$.

Although a reward that drives $B_1[1] \geq 0$ causes $A_{out,1}[k]$ to initially increase, the additional dynamics of $B_2[k]$ may be such that $B_2[k]$ (and thus, $A_{conf,1}[k]$) is a decreasing function that causes $B_1[k]$ to become negative. To see why, consider the expression for $B_2[k]$, which, from (4.24), is given by

$$B_2[k] = A_{out,2}[k] + \mu_{31} B_1[k-1]. \tag{5.12}$$

If $B_2[k] < 0$, then any dissonance pressure that arises causes $A_{out,2}[k]$ to decrease. Should $B_2[k]$ remain negative, $A_{out,2}[k]$ remains a decreasing function of k and thus, it is possible that $A_{conf,1}[k]$, through $B_2[k]$, becomes strong enough to drive $B_1[k] < 0$ (see (5.11)). On the other hand, if $B_2[k]$ is positive, then any dissonance pressure that arises causes $A_{out,2}[k]$ to increase and thus, increases $B_2[k]$ (and, through $A_{conf,1}[k]$, $B_1[k]$). Therefore, it is desirable to also drive $B_2[k]$ positive. The reward first affects $B_2[k]$ at $k = 2$. As before, if $B_2[2] < 0$, then $B_2[k]$ is unlikely to become positive due to a decreasing $A_{out,2}[k]$. Thus, the control strategy also attempts to drive $B_2[2] \geq 0$.

To summarize, the approach for achieving the easier control objective is to find an R_o such that when a controller of the form (5.9) is applied to the two-person system, $B_1[1] \geq 0$ and $B_2[2] \geq 0$. Since $B_1[1]$ and $B_2[2]$ are given by

$$B_1[1] = A_{o1} + \mu_{32} A_{o2} + \mu_{11} R_o \tag{5.13}$$

$$B_2[2] = A_{o2} + K_{12} P_2^{CD}[1] + \mu_{31} \left(A_{o1} + \mu_{11} R_o + \mu_{32} A_{o2} \right), \tag{5.14}$$

such an R_o must satisfy the following two inequalities:

$$R_o \geq -\frac{A_{o1} + \mu_{32} A_{o2}}{\mu_{11}} \text{ and} \tag{5.15}$$

$$R_o \geq -\frac{A_{o2}\left(1 + \mu_{31}\mu_{32}\right) + K_{12} P_2^{CD}[1] + \mu_{31} A_{o1}}{\mu_{31}\mu_{11}}. \tag{5.16}$$

Although these two inequalities drive $B_1[1] \geq 0$ and $B_2[2] \geq 0$, they do not necessarily imply that $B_1[k] \geq 0$ for $k \geq 1$, due to the complicated dynamics arising from the conformity feedback loop. However, by slightly modifying the conditions given in (5.15) and (5.16), not only can we guarantee $B_1[k] \geq 0$ for $k \geq 1$, but also $B_2[k] \geq 0$ for $k \geq 2$. Modifications to (5.15) and (5.16) are specific to the sign of $P_2^{CD}[1]$, which depends on the initial operating region. If the system begins in region V, VII or VIII, then $P_2^{CD}[1] \geq 0$, whereas $P_2^{CD}[1] < 0$ if the system begins in region VI. These two cases are considered separately, beginning with initial regions V, VII and VIII .

Suppose the system begins in region V. Then, since $B_2[1]$ is given by $B_2[1] = A_{o2} + \mu_{31} A_{o1}$, both cognitions at $k = 1$ are negative and thus, $P_2^{CD}[1] = 0$. On the other hand, if the system begins in region VII, then $A_{o2} > 0$ and $A_{o2} \geq \mu_{31}|A_{o1}|$. As a result, $B_2[1] \geq 0$ and thus from (4.29), $P_2^{CD}[1] > 0$. Finally, if the system begins in region VIII, then $A_{o2} > 0$ and $A_{o2} \geq \frac{|A_{o1}|}{\mu_{32}}$. As a result, $B_2[1] \geq 0$ and thus from (4.29), $P_2^{CD}[1] > 0$. From (5.14), if $P_2^{CD}[1] \geq 0$, then ensuring

$$A_{o2} + \mu_{31}\left(A_{o1} + \mu_{11} R_o + \mu_{32} A_{o2}\right) \geq 0$$

is sufficient to guarantee $B_2[2] \geq 0$. By rearranging the above inequality, the range of sufficient R_o values able to drive $B_2[2] \geq 0$ is given by

$$R_o \geq -\frac{A_{o2}\left(1 + \mu_{31}\mu_{32}\right) + \mu_{31} A_{o1}}{\mu_{11}\mu_{31}}. \tag{5.17}$$

Thus, for initial regions V, VII and VIII, driving $B_1[1] \geq 0$ and $B_2[2] \geq 0$ requires that both (5.15) and (5.17) hold. The reward magnitude that satisfies these two inequalities can be refined to guarantee $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. The lemma below presents the refined R_o that satisfies the given strategy of driving both behaviours positive.

Lemma 5.11. *If a control signal of the form (5.9) is applied to the two-person system with the initial conditions given by (4.26), and the initial attitudes satisfy (5.1), (5.7) or (5.8), then* $B_1[k] \geq 0$ *for* $k \geq 1$ *and* $B_2[k] \geq 0$ *for* $k \geq 2$ *when*

$$R_o \geq \max\left\{-\frac{A_{o1}\left(1 + \mu_{31}\mu_{32}\right) + \mu_{32} A_{o2}}{\mu_{11}}, -\frac{\left(A_{o2} + \mu_{31} A_{o1}\right)\left(1 + \mu_{31}\mu_{32}\right)}{\mu_{11}\mu_{31}}\right\}. \tag{5.18}$$

Proof. See Appendix E.9. □

The second expression in (5.18) is the same as the right-hand side of (5.17) with an additional term, $\mu_{31}^2\mu_{32}A_{o1}$, which is used to guarantee $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. To demonstrate the effectiveness of the above controller, a simulation is performed, shown in the left-hand side plots of Figure 5.8. The plots on the right-hand side of Figure 5.8 show that the control strategy given in Lemma 5.11 is conservative.

Now, suppose the system begins in region VI. Then, since $A_{o2} \geq 0$ and $A_{o2} < \mu_{31}|A_{o1}|$, it follows that $B_2[1] < 0$ and, thus, dissonance pressure arises due to the inconsistency between $B_2[1]$ and $A_{out,2}[1] = A_{o2}$. From (4.29), this dissonance pressure is given by

$$P_2^{CD}[1] = -\frac{|A_{o2}|}{|A_{o2}| + \mu_{31}|A_{o1}|}.$$

As a result, (5.16) can be expressed as

$$R_o \geq \frac{1}{\mu_{11}\mu_{31}}\left(\frac{K_{12}|A_{o2}|}{|A_{o2}| + \mu_{31}|A_{o1}|} - \left(A_{o2}\left(1 + \mu_{31}\mu_{32}\right) + \mu_{31}A_{o1}\right)\right). \tag{5.19}$$

For initial region VI, ensuring $B_1[1] \geq 0$ and $B_2[2] \geq 0$ requires that both (5.15) and (5.19) hold. The reward magnitude that satisfies these two inequalities can be refined to guarantee $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. The lemma below presents the refined R_o that satisfies the given strategy of driving both behaviours positive.

Lemma 5.12. *If a control signal of the form (5.9) is applied to the two-person system with the initial conditions given by (4.26), and the initial attitudes satisfy (5.4), then* $B_1[k] \geq 0$ *for* $k \geq 1$ *and* $B_2[k] \geq 0$ *for* $k \geq 2$ *when*

$$\begin{aligned} R_o \geq \max\Bigg\{ &-\frac{A_{o1}\left(1 + \mu_{31}\mu_{32}\right) + \mu_{32}A_{o2}}{\mu_{11}}, \\ &\frac{1}{\mu_{11}\mu_{31}}\left(\frac{K_{12}A_{o2}}{A_{o2} + \mu_{31}|A_{o1}|} - \left(A_{o2} + \mu_{31}A_{o1}\right)\left(1 + \mu_{31}\mu_{32}\right)\right)\Bigg\}. \end{aligned} \tag{5.20}$$

Proof. See Appendix E.10. □

The second expression in (5.20) is the same as (5.19) with an additional term, $\mu_{31}^2\mu_{32}A_{o1}$, which is used to guarantee $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. To demonstrate the effectiveness of the above controller, a simulation is performed, shown in the left plots of Figure 5.9. The plots on the right-hand side of Figure 5.9 show that the control strategy given in Lemma 5.12 is conservative.

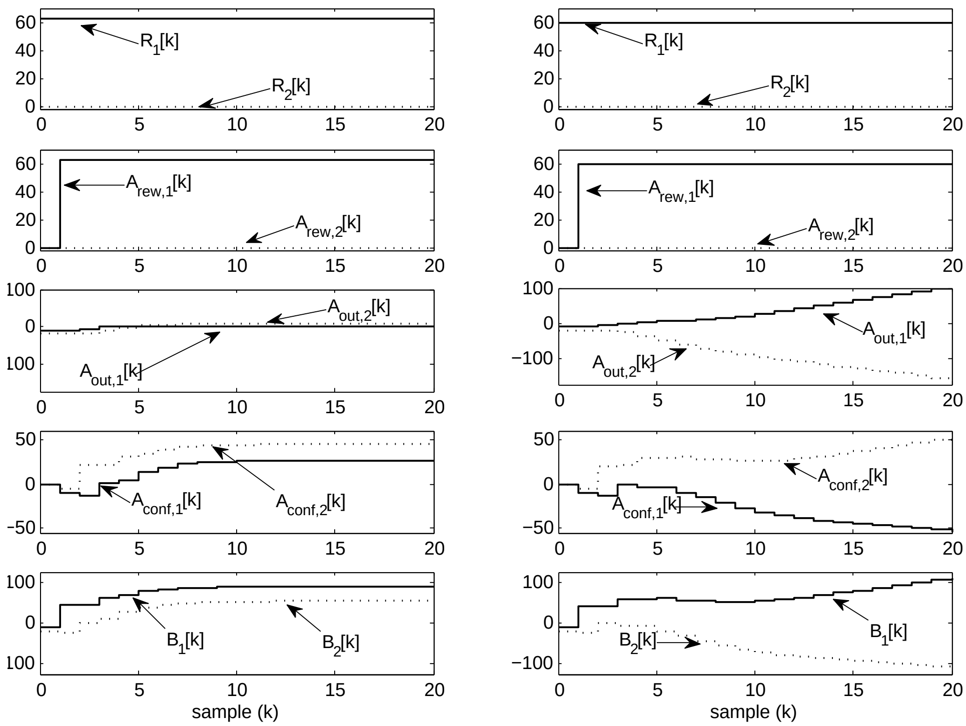

Figure 5.8: A controller of the form (5.9) is applied to the two-person system with the initial conditions given in (4.26), parameter values in Table 4.1, where $A_{o1} = -10$ *attitude units* and $A_{o2} = -20$ *attitude units* (thus the system begins in initial region V). The plots on the left-hand side show the simulation results for a reward magnitude that satisfies (5.18), specifically, $R_o = \$63$. This reward is sufficient to drive $B_1[1] \geq 0$ and $B_2[2] \geq 0$ and to keep both behaviours positive for $k \geq 2$. The plots on the right-hand side demonstrate that not only is Lemma 5.11 conservative because the reward magnitude, $R_o = \$60$, does not satisfy (5.18), but the strategy of driving $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ is also conservative, since in this simulation $B_2[k] < 0$ as k tends to infinity yet $B_1[k] \geq 0$ for $k \geq 1$. Initial region V is used for simulations to show that even in the case when person one experiences significant, negative conformity pressure, it is possible to drive $B_1[k] \geq 0$ as k tends to infinity.

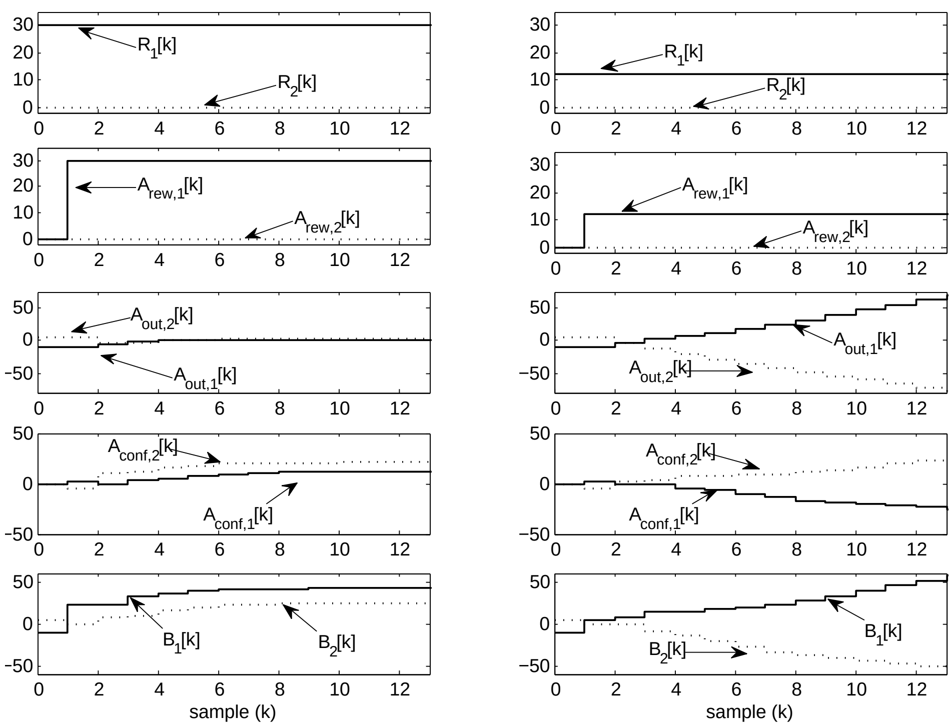

Figure 5.9: A controller of the form (5.9) is applied to the two-person system with the initial conditions given in (4.26), parameter values in Table 4.1, where $A_{o1} = -10$ *attitude units* and $A_{o2} = 4$ *attitude units* (thus the system begins in initial region VI). The plots on the left-hand side show the simulation results for a reward magnitude that satisfies (5.20), specifically, $R_o = \$30$. This reward is sufficient to drive $B_1[1] \geq 0$ and $B_2[2] \geq 0$ and to keep both behaviours positive for $k \geq 2$. The plots on the right-hand side demonstrate that Lemma 5.12 is conservative because $B_1[k] \geq 0$ for $k \geq 1$, even though the reward magnitude, $R_o = \$12$, does not satisfy (5.20). Moreover, the plots on the right-hand side also show that the strategy of driving $B_2[k] \geq 0$ for $k \geq 2$ is also conservative, since in this simulation $B_2[k] < 0$ as k tends to infinity yet $B_1[k] \geq 0$ for $k \geq 1$.

To conclude this section, Lemmas 5.11 and 5.12 provide sufficient conditions on R_o such that for $k \geq 1$, $B_1[k] \geq 0$. This goal is achieved by finding an R_o that is able to drive the behaviour of both people positive for $k \geq 2$. These two lemmas show that, provided the reward is high enough, not only are parents able to induce their child to play the piano when he does not want to, this reward is also enough to entice the child's friend to play the piano as well, even though the friend is not being offered a reward. This result is somewhat interesting and highlights the impact conformity pressure may have on an individual's behaviour. The conclusions of this section suggest that, provided a high enough reward is offered, parents can eventually induce their child to to play piano at some desired strength, $B_d > 0$; thus, in the next section, the main control objective is considered.

5.2.2 General Step-Reward Controller Design

This section investigates whether or not, for all $B_d > 0$, there exists an R_o such that a controller of the form given in (5.9) can drive $B_1[k] \geq B_d$ as $k \to \infty$. The strategy to find such an R_o builds on the results of the previous section. To obtain a closed-form expression for a sufficient reward to meet the control objective, first it is shown that it is possible for $B_1[k]$ to converge as k tends to infinity. Then, the expression for $B_1[\infty]$ can be used to find a reward that is sufficient to drive $B_1[k] \geq B_d$ as k tends to infinity.

Before determining whether or not an R_o exists such that $B_1[k]$ converges as k tends to infinity, a few preliminary results are needed. Since $B_1[k]$ depends on $A_{out,1}[k]$ and $A_{out,2}[k]$ (indirectly through $A_{conf,1}[k]$), it is necessary to characterize $A_{out,1}[k]$ and $A_{out,2}[k]$ as k tends to infinity. From Lemmas 5.11 and 5.12, it is possible to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. When both behaviours are positive, by Assumption 4.10, both attitudes are non-decreasing functions of k until $A_{out,1}[k]$ becomes positive (at which time, $A_{out,1}[k]$ decreases due to the overjustification effect). Similar to Lemma 3.3 from the one-person system, if both behaviours are positive, then each attitude converges as k tends to infinity. The lemma below formalizes this result.

Lemma 5.13. *For the two-person system with the initial conditions given by (4.26), if $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then $A_{out,1}[k]$ and $A_{out,2}[k]$ converge to some constants c_1 and c_2 respectively as k tends to infinity. Moreover,*

(a) if there exists some $\hat{k} > 0$ such that $A_{out,1}[\hat{k}] > 0$, then $c_1 = 0$, otherwise, $c_1 \leq 0$; and

(b) if there exists some $\tilde{k} > 2$ such that $A_{out,2}[\tilde{k}] > 0$, then $c_2 = A_{out,2}[\tilde{k}]$, otherwise, $c_2 \leq 0$.

Proof. See Appendix E.11. □

The above lemma states that if a reward drives $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then both attitudes converge as k tends to infinity. For $A_{out,2}[k]$, the condition that $\tilde{k} > 2$ (instead of $\tilde{k} > 0$) is a technical detail that accounts for the case when $P_2^{CD}[1] < 0$, and the delay between when the reward is first applied and the sample at which it first affects person two.

Since both attitudes converge as k tends to infinity, each attitude is bounded. This conclusion is promising because $B_1[k]$ and $B_2[k]$ depend on both attitudes, i.e.,

$$B_1[k] = A_{out,1}[k] + \mu_{11}R_o + \mu_{32}A_{out,2}[k-1] + \mu_{31}\mu_{32}B_1[k-1]$$
$$B_2[k] = A_{out,2}[k] + \mu_{31}A_{out,1}[k-1] + \mu_{31}\mu_{11}R_o + \mu_{31}\mu_{32}B_2[k-1]$$

In both of the above equations, the only signals that could possibly be unbounded are $B_1[k]$ and $B_2[k]$. Through a non-trivial analysis, Assumption 4.7 and Lemma 5.13 can be used to show that both behaviours are bounded. The lemma below formalizes this result.

Lemma 5.14. *If a control signal of the form (5.9) is applied to the two-person system with the initial conditions given by (4.26) and drives $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then $B_1[\cdot]$ and $B_2[\cdot]$ are bounded.*

Proof. See Appendix E.12. □

At this point, it has been shown that there exists an R_o such that a controller of the form given by (5.9) is able to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ and these two behaviours are bounded. To show that, in this situation, $B_1[k]$ converges as k tends to infinity, consider that by defining

$$\alpha[k] := A_{out,1}[k] + \mu_1 R_o + \mu_{32}A_{out,2}[k-1],$$

$B_1[k]$ can be expressed as

$$B_1[k] = \alpha[k] + \mu_{31}\mu_{32}B_1[k-2]. \tag{5.21}$$

From Lemma 5.13, $\alpha[k]$ is bounded, and, from Lemma 5.14, $B_1[k]$ is bounded. If it can be shown that for large enough k, $\alpha[k]$ is a non-decreasing function of k, then, since $B_1[k]$ is bounded and is given by (5.21), it follows (from Lemma A.2) that $B_1[k]$ must necessarily converge as k tends to infinity. Since $\alpha[k]$ depends on $A_{out,1}[k]$ and $A_{out,2}[k]$, showing $\alpha[k]$ is eventually a non-decreasing function of k requires knowing more information about the steady-state values of $A_{out,1}[k]$ and $A_{out,2}[k]$.

From Lemma 5.13, if $A_{out,1}[k]$ becomes positive, then $A_{out,1}[k]$ necessarily tends to zero as k tends to infinity. To determine the value to which $A_{out,1}[k]$ converges in the case when $A_{out,1}[k] < 0$ for $k \geq 0$, the dissonance pressure expression is considered. It turns out that since $B_2[k]$ is a bounded function of k, again $A_{out,1}[k]$ necessarily converges to zero, as stated in the lemma below.

Lemma 5.15. *If a control signal of the form (5.9) is applied to the two-person system with the initial conditions given by (4.26) and drives $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then*

$$A_{out,1}[k] \to 0 \text{ as } k \to \infty.$$

Proof. See Appendix E.13. □

Now, consider $A_{out,2}[k]$, which, from Lemma 5.13, converges as k tends to infinity. Since, by Lemma 5.13, the steady-state value of $A_{out,2}[k]$ is given in the case when $A_{out,2}[k]$ becomes positive at some $\tilde{k} > 2$, it only remains to determine the value for $A_{out,2}[\infty]$ in the case when $A_{out,2}[k] < 0$ for $k \geq 2$. Similar to $A_{out,1}[k]$, the dissonance pressure expression and the results Lemma 5.14 imply that $A_{out,2}[k]$ tends to zero as k tends to infinity. The lemma below formalizes these conclusions.

Lemma 5.16. *If a control signal of the form (5.9) is applied to the two-person system with the initial conditions given by (4.26) and drives $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then*

(a) if, at some $\tilde{k} > 2$, $A_{out,2}[\tilde{k}] \geq 0$, then $A_{out,2}[k] \to A_{out,2}[\tilde{k}]$ as $k \to \infty$;

(b) otherwise, $A_{out,2}[k] \to 0$ as $k \to \infty$.

Proof. See Appendix E.14. □

Now, the results of Lemmas 5.15 and 5.16 can be used to show that for large enough k, $\alpha[k]$ in (5.21) remains non-negative, implying that $B_1[k]$ converges as k tends to infinity. The steady-state value of $B_1[k]$ is given by the lemma below.

Lemma 5.17. *If a control signal of the form (5.9) is applied to the two-person system with the initial conditions given by (4.26) and drives $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then*

$$\lim_{k \to \infty} B_1[k] = \frac{\mu_{11} R_o + \mu_{32} c_2}{1 - \mu_{31}\mu_{32}}. \tag{5.22}$$

Proof. See Appendix E.15. □

In other words, the above lemma demonstrates that there exists an R_o such that when step reward is applied to the two-person system, $B_1[k]$ converges as k tends to infinity, provided this R_o is sufficient to drive both behaviours positive for $k \geq 2$. From Lemmas 5.11 and 5.12, there exists an R_o that is able to meet this condition. Such an R_o must satisfy (5.18) if the system begins in regions V, VII or VIII, or (5.20) if the system begins in region VI. Define R_o^1 as the reward magnitude at the right-hand side of (5.18) and R_o^2 as the reward magnitude at the right-hand side of (5.20). Now, it only remains to find an expression for an R_o that is able to drive $B_1[k] \geq B_d$ as k tends to infinity. This amounts to setting, by (5.22),

$$\frac{\mu_{11} R_o + \mu_{32} c_2}{1 - \mu_{31}\mu_{32}} \geq B_d,$$

and rearranging for R_o, giving the following inequality:

$$R_o \geq \frac{B_d \left(1 - \mu_{31}\mu_{32}\right) - \mu_{32} c_2}{\mu_{11}}. \tag{5.23}$$

With the exception of c_2, all of the terms in the above inequality are given directly by the parameter values. From Lemma 5.16, c_2 has two possible values: zero (in the case when $A_{out,2}[k]$ remains negative) and $A_{out,2}[\tilde{k}]$ (in the case when $A_{out,2}[k]$ becomes positive). From (5.23), a lower bound on c_2 provides a value that can replace c_2 in the above inequality. Of the two possible steady-state values for $A_{out,2}[k]$, zero is the smallest. Thus, if

$$R_o \geq \frac{B_d \left(1 - \mu_{31}\mu_{32}\right)}{\mu_{11}}, \tag{5.24}$$

then (5.23) also holds. If the system begins in region V, VII or VIII, then, provided $R_o \geq R_o^1$, the above inequality ensures $B_1[k] \geq B_d$ as k tends to infinity. Similarly, if the system begins in region VI, then, provided $R_o \geq R_o^2$, the above inequality ensures the control objective is met. Define R_o^3 as the reward magnitude at the right-hand side of (5.24). By combining the results of Lemmas 5.11, 5.12 and 5.17, the final theorem is obtained, which gives an expression for a sufficient R_o such that a step-reward controller is able to meet the control objective of driving $B_1[k] \geq B_d$ as k tends to infinity.

Theorem 5.1. *For all $B_d > 0$, if a control signal of the form (5.9) is applied to the two-person system with the initial conditions given by (4.26), then, provided*

$$R_o \geq \begin{cases} \max\left\{R_o^1, R_o^3\right\} & \textit{if (5.1), (5.7) or (5.8) hold,} \\ \max\left\{R_o^2, R_o^3\right\} & \textit{if (5.4) holds,} \end{cases} \tag{5.25}$$

$B_1[k] \geq B_d$ as k tends to infinity.

Proof. See Appendix E.16. □

The above theorem states that a step reward is able to meet the control objective of driving $B_1[k] \geq B_d$ as k tends to infinity. Instead of selecting the maximum value between R_o^1, R_o^2 and R_o^3, two cases are considered: the first case considers the situation in which $P_2^{CD}[1] \geq 0$, while the second case considers the situation in which $P_2^{CD}[1] < 0$. By distinguishing between these two situations, the reward most specific to the initial conditions can be obtained. In particular, R_o^2, which is specific to initial region VI, compensates for the effect of negative dissonance pressure experienced by person two at $k = 1$. If the system begins in region V, then no dissonance pressure arises (for either person) at $k = 1$; thus R_o^2 is higher than the value necessary to meet the control objective when the system begins in region V. Practically speaking, suppose, in the piano example, that the child and his friend both have initially negative attitudes. If the child's parents offer a reward that is at least the maximum value between R_o^1, R_o^2 and R_o^3, then these parents are spending more money than is required for convincing their child to play the piano at some desired level. Instead, the parents can save money by offering the maximum value between R_o^1 and R_o^3, as proposed by (5.25).

A second interesting consequence of Theorem 5.1 is that in some cases, the reward value necessary to drive $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k \geq 2$, i.e., R_o^1 or R_o^2, is greater than R_o^3. This is typically the case for highly negative initial attitudes, as the reward must overcome their negative effect on $B_1[k]$ and $B_2[k]$. In these situations, $B_1[k]$ settles to a steady-state value that is greater than B_d. The simulation results given in Figure 5.10 demonstrate this consequence along with the results of Theorem 5.1. A simulation is performed for two situations, both of which begin in region V. The plots on the left-hand side show the results of a simulation for attitudes that, relative to the desired behaviour, are weak; thus, the $R_o = R_o^3$. The plots on the right-hand side show the results of a simulation for attitudes that are strong enough to require $R_o = R_o^1$. In this situation, $B_1[k]$ converges to a value greater than B_d as k tends to infinity.

The strategy of driving $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ proves to be a useful stepping-stone for achieving the main control objective, as many key results arise from this strategy. These key results include determining $B_1[k]$ and $B_2[k]$ are bounded (Lemma 5.14), obtaining steady-state expressions for $A_{out,1}[k]$ and $A_{out,2}[k]$ (Lemmas 5.15 and 5.16), and finding a steady-state expression for $B_1[k]$ (Lemma 5.17), from which a range of R_o values that are able to meet the control objective is obtained. However, even though Theorem 5.1 states that is it possible to meet the control objective with a step reward, open-loop control has drawbacks and, therefore, a closed-loop control approach is considered.

5.3 Closed-Loop Investigation

Similar to the controller design for the one-person system, closed-loop control is considered because of its general benefits (i.e., the ability to more effectively fight disturbances and uncertainty). For the one-person system, two types of closed-loop controllers were designed: state-feedback and output-feedback. Since the output-feedback controller design

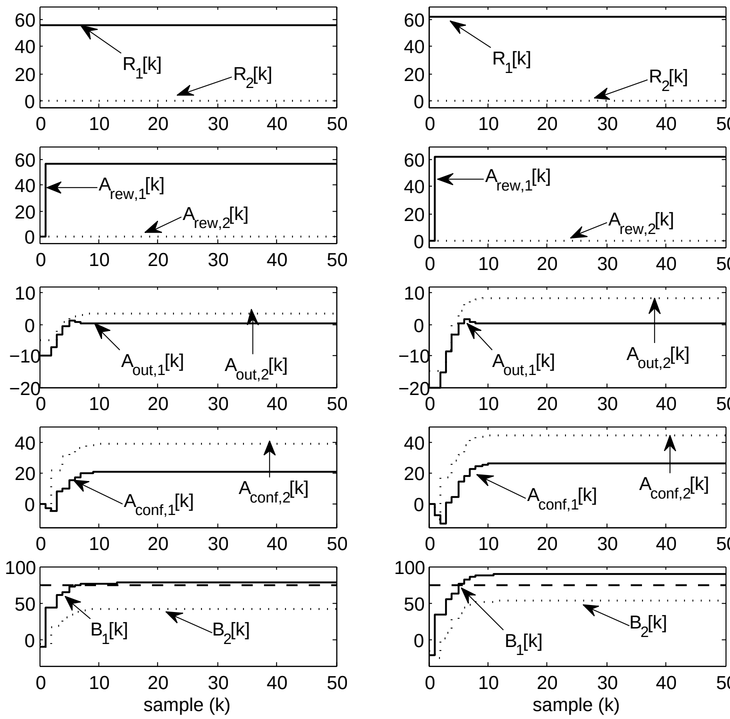

Figure 5.10: A controller of the form (5.9) is applied to the two-person system with the initial conditions given in (4.26), parameter values in Table 4.1 and $B_d = 75$ *attitude units*. The plots on the left-hand side show a situation in which $A_{o1} = -10$ *attitude units* and $A_{o2} = -5$ *attitude units*. In this situation, the attitude magnitudes are small relative to the desired behaviour, $B_d = 75$ *attitude units*, denoted by the dashed line in the bottom plot. From (5.25), $R_o \geq \max\{R_o^1, R_o^3\}$ (since the system begins in region V). For the parameter values used, $R_o^1 = \$25$ and $R_o^3 = \$56.25$; thus a step reward with magnitude $R_o = \$56.25$ is applied to the two-person system. The plots on the right-hand side show a situation in which $A_{o1} = -20$ *attitude units* and $A_{o2} = -15$ *attitude units*. In this situation, the attitude magnitudes are large enough, relative to the desired behaviour, $B_d = 75$ *attitude units*, such that $R_o^1 = \$62.5$ is the larger of the two possible reward values (note that R_o^3 remains the same as in the previous case); thus, a reward magnitude of $R_o = \$62.5$ is applied to the two-person system. In both situations, $A_{out,1}[k]$ converges to zero, $A_{out,2}[k]$ converges to $A_{out,2}[\tilde{k}]$, $B_2[k]$ is bounded and $B_1[k]$ converges to at least B_d as k tends to infinity, thus supporting Lemmas 5.14, 5.15, and 5.16, and Theorem 5.1.

for the one-person system was complicated, designing an output-feedback controller for the two-person system is not even considered due to the additional complexities arising from conformity pressure dynamics. Therefore, the state-feedback approach is the only closed-loop control strategy considered for the two-person system. Note that, as an item for future work, examining how the one-person output-feedback controller given in Section 3.2 performs when applied to the two-person system could provide insight to designing an output-feedback controller for the two-person system.

The general state-feedback controller developed for the one-person system (in Section 3.2.1) suggests a strategy that could be used for the two-person system. Recall from Section 3.2.1 that the general state-feedback controller for the one-person system is given by

$$R[k] = \max\left\{0, \frac{1}{\mu_1}\left(B_d - A_{out}[k] - \Delta A_{out}[k]\right)\right\}.$$

This controller was developed by rearranging the equation for $B[k]$ in terms of the control signal, $R[k]$, and limiting $R[k]$ to ensure it remains non-negative. From Theorem 3.8, the above state-feedback controller drives $B[k] \geq B_d$ for $k \geq 1$, while ensuring all signals converge as k tends to infinity. Since the control objective for the two-person system is to drive $B_1[k] \geq B_d$ as $k \to \infty$, a state-feedback controller with a form similar to the one-person state-feedback controller above is considered. From (4.24), $B_1[k]$ is given by

$$B_1[k] = A_{out,1}[k] + A_{rew,1}[k] + A_{conf,1}[k].$$

It follows from this equation that if, for $k \geq 1$, $A_{out,1}[k] + A_{rew,1}[k] + A_{conf,1}[k] \geq B_d$, then the control objective is met, not only as $k \to \infty$, but for all $k \geq 1$. Observe that because $A_{rew,1}[k] = \mu_{11} R_1[k-1]$, and $R_1[k]$ equals the control signal, $R[k]$, the control signal enters the above inequality in a straight-forward manner. As a result, the above inequality can be arranged as

$$R[k-1] \geq \frac{1}{\mu_{11}}\left(B_d - A_{out,1}[k] - A_{conf,1}[k]\right).$$

Using the expression for $A_{out,1}[k]$ given in (4.1) and the expression for $A_{conf,1}[k]$ given in (4.28), this inequality can be re-written as

$$R[k-1] \geq \frac{1}{\mu_{11}}\left(B_d - A_{out,1}[k-1] - \Delta A_{out,1}[k-1] - \mu_{32} B_2[k-1]\right).$$

Accounting for the requirement that $R[k] \geq 0$ and shifting samples suggests a state-feedback controller given by

$$R[k] = \max\left\{0, \frac{1}{\mu_{11}}\left(B_d - A_{out,1}[k] - \Delta A_{out,1}[k] - \mu_{32} B_2[k]\right)\right\}, \tag{5.26}$$

which, from the theorem below, meets the control objective.

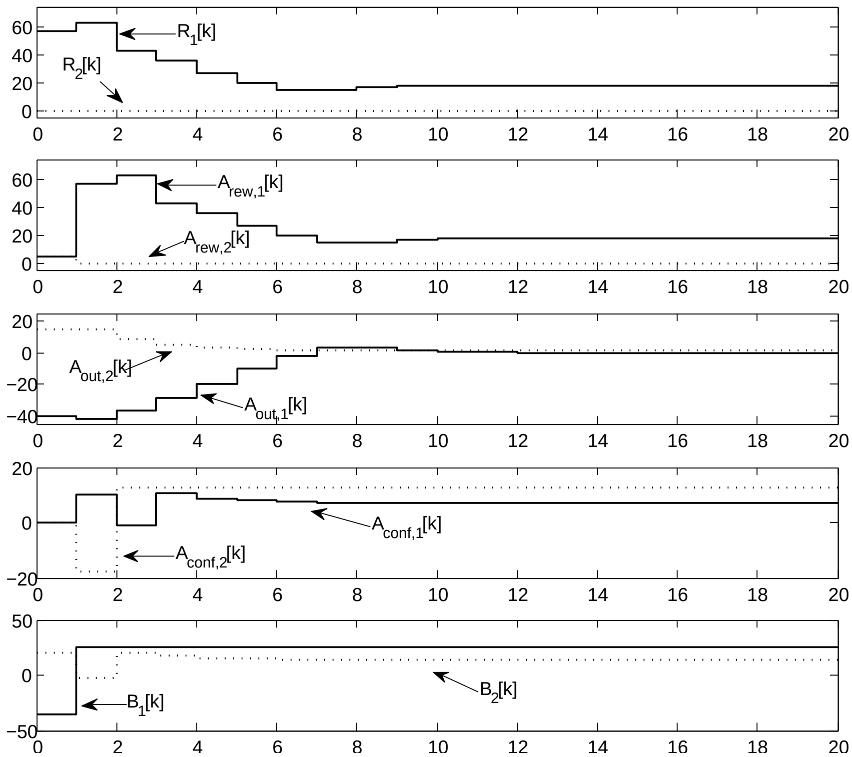

Figure 5.11: A controller of the form (5.26) is applied to the two-person system with parameter values in Table 4.1. In this simulation, $A_{o1} = -40$ *attitude units*, $A_{o2} = 15$ *attitude units*, $B_d = 25$ *attitude units* and the initial conditions do not match those given in (4.26). Note that the initial conditions used for this simulation result in dissonance pressure at $k = 0$ and thus attitude change begins at $k = 1$ for both person one and person two. Nevertheless, the controller given by (5.26) is able to drive $B_1[k] \geq B_d$ for $k \geq 1$.

Theorem 5.2. *For all $B_d > 0$, if the control signal given in (5.26) is applied to the two-person system, then $B[k] \geq B_d$ for $k \geq 1$.*

Proof. See Appendix E.17. □

Note that Theorem 5.2 does not depend on the initial conditions. That is, the state-feedback controller given in (5.26) meets the control objective for any set of initial conditions (due to how the reward enters the expression for $B_1[k]$). The simulation results in Figure 5.11 demonstrate that the controller works for initial conditions that do not satisfy (4.26).

Although this state-feedback controller does, indeed, meet the control objective of driving person one's behaviour to at least B_d, it may result in unbounded states. For example, suppose the control signal in (5.26) is applied to the two-person system when both initial attitudes are negative (and all other initial conditions are given by (4.26)). Although the controller can ensure $B_1[k] \geq B_d$ for $k \geq 1$, it may not drive $B_2[k] \geq 0$. In this situation, person two experiences negative dissonance pressure, which decreases his attitude, and, in turn, his behaviour. In fact, person two's behaviour could be decreasing indefinitely, while person one's behaviour could be increasing indefinitely, and thus, these signals could be unbounded. The simulation shown in Figure 5.12 proves this situation may arise. Although the controller given by (5.26) meets the control objective for $k \geq 1$, bounded signals are desirable over unbounded signals. Therefore, a slight modification is made to (5.26), ensuring all signals are bounded and giving a stable system.

Recall from the previous section that if $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then both attitudes converge as k tends to infinity (see Lemma 5.13). The controller given in (5.26) ensures $B_1[k] \geq B_d$ for $k \geq 1$ and, thus, the behaviour of person one is positive for $k \geq 1$. However, as shown in the simulation results of Figure 5.12, this controller does not guarantee that the behaviour of person two is positive for $k \geq 2$. By slightly modifying (5.26) at $k = 0$, and imposing some mild, technical conditions on B_d, it is possible to force $B_2[k] \geq 0$ for $k \geq 2$. Since the reward first affects $B_2[k]$ at $k = 2$, $R[0]$ must be large enough to ensure $B_2[2] \geq 0$. From Section 5.2, there exists a reward value for each initial region that guarantees $B_2[2] \geq 0$. Therefore, the above state-feedback controller is modified by setting $R_1[0]$ to either R_o^1 or R_o^2 (depending on the initial region), while remaining of the form given by (5.26) for $k \geq 1$, i.e.,

$$R[k] = \begin{cases} R_o^1 & \text{if (5.1), (5.7)} \\ & \text{or (5.8) hold, and } k = 0 \\ R_o^2 & \text{if (5.4) holds, and } k = 0 \\ \max\left\{0, \frac{1}{\mu_{11}}\left(B_d - A_{out,1}[k] - \Delta A_{out,1}[k] - \mu_{32} B_2[k]\right)\right\} & \text{if } k \geq 1. \end{cases} \tag{5.27}$$

Theorem 5.1 implies that the above controller ensures $B_1[1] \geq 0$ and $B_2[2] \geq 0$. Moreover, it can be shown that $B_1[k] \geq B_d$ for $k \geq 2$. As the goal of the above controller is to also ensure $B_2[k] \geq 0$ for $k \geq 2$, it may be necessary, in some initial regions, to impose conditions on B_d. The theorem below summarizes these results and provides the expression for a sufficiently large B_d for all initial regions that may arise from the initial conditions.

Theorem 5.3. *For all $B_d \geq \frac{1}{\mu_{31}} \max\left\{|A_{o2}|, \frac{K_{12}}{2}\right\}$, if the control signal given by (5.27) is applied to the two-person system with the initial conditions given by (4.26), then*

(a) $B_1[k] \geq B_d$ for $k \geq 2$,

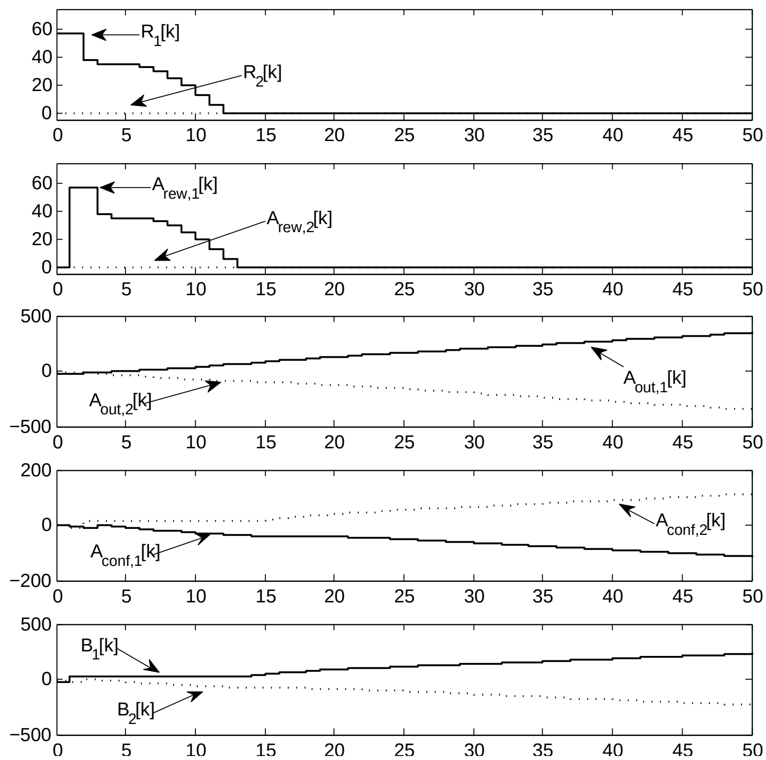

Figure 5.12: A controller of the form (5.26) is applied to the two-person system with the initial conditions given in (4.26) and parameter values in Table 4.1. In this simulation, $A_{o1} = -20$ *attitude units*, $A_{o2} = -15$ *attitude units* and $B_d = 30$ *attitude units*. The controller fails to ensure $B_2[k] \geq 0$, and thus, the attitude of person two decreases due to the resulting negative dissonance pressure, causing $B_2[k]$ to continuing decreasing. The increasingly negative conformity pressure exerted by person two results in person one experiencing dissonance pressure which causes $A_{out,1}[k]$ to increase. As a result, $B_1[k]$ increases and the control signal, $R[k]$, decreases (as less reward is required to meet the control objective due to the increased internal attitude). Eventually, no reward is necessary, and each behaviour and internal attitude continues to grow in magnitude as k tends to infinity.

(b) For $i = 1, 2$, $A_{out,i}[k]$, $\Delta A^{CD}_{out,i}[k]$, $\Delta A^{OJ}_{out,i}[k]$, $P^{CD}_i[k]$, $P^{OJ}_i[k]$ converge as k tends to infinity, and

(c) $R[\cdot]$, $B_1[\cdot]$ and $B_2[\cdot]$ are bounded.

Proof. See Appendix E.18. □

The above theorem provides a controller that is able to meet the control objective for $k \geq 2$. Note that, in contrast to Theorem 5.2, Theorem 5.3 requires the initial conditions to be those specified in (4.26). Moreover, if the system encounters some disturbance that causes $B_2[k] < 0$, there is no guarantee that the above controller will ensure $B_2[k]$ will eventually become positive again (since, after $k = 1$, the controller above simplifies to that given in (5.26), which does not guarantee $B_2[k] \geq 0$). The condition that B_d be sufficiently large is to ensure that the desired behaviour is large enough to overcome a negative value for $A_{out,2}[2]$. This condition is relatively conservative, as it is derived from known bounds on $A_{out,2}[2]$ instead of the exact expression for $A_{out,2}[2]$ for each initial region. The simulation results in Figure 5.13 demonstrate that the condition on B_d is conservative, as the inequality given in Theorem 5.3 is not satisfied, yet the control objective is met and all signals converge as k tends to infinity. Although Theorem 5.3 says that $B_1[k]$, $B_2[k]$ and $R[k]$ are bounded signals of k, the simulation results shown in Figure 5.13 suggest that, at least under some conditions, these three signals converge as k tends to infinity.

To summarize, this chapter considered open-loop and closed-loop control strategies for driving $B_1[k] \geq B_d$ as k tends to infinity. Before designing these controllers, the zero-input system response was studied to not only confirm the consistency between our model and the psychology, but to motivate the open-loop control approach of driving both behaviours positive as soon as possible. From this strategy, several intermediate results arose that were useful for meeting the control objective with a step reward. Then, due to the benefits of closed-loop controllers, a state-feedback controller was developed to drive $B_1[k] \geq B_d$ for $k \geq 1$.

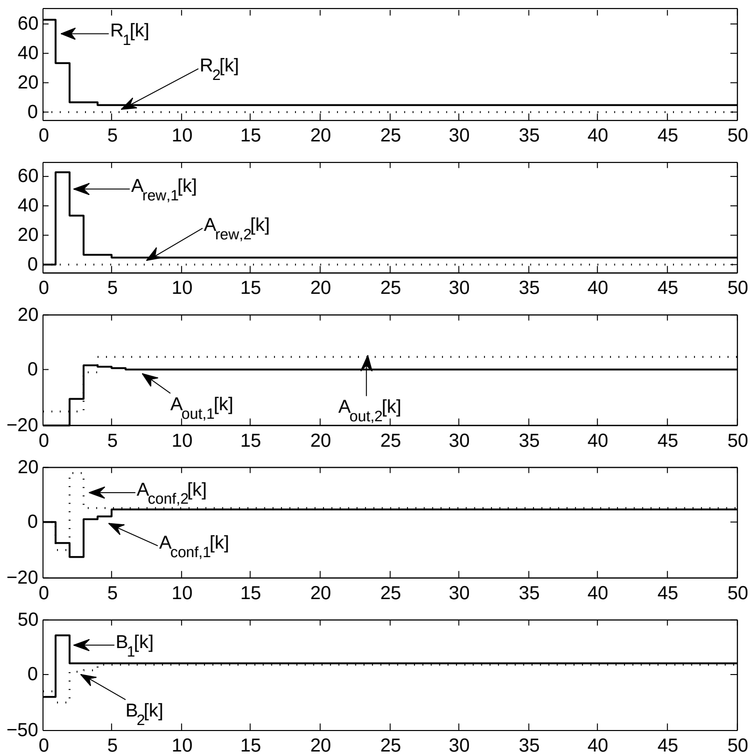

Figure 5.13: A controller of the form (5.27) is applied to the two-person system with the initial conditions given in (4.26) and parameter values in Table 4.1. In this simulation, $A_{o1} = -20$ *attitude units*, $A_{o2} = -15$ *attitude units* and $B_d = 10$ *attitude units*. Since the system begins in region V, $R[0] = R_o^1 = \$62.5$, which has the effect of driving $B_1[1] > B_d$; then, when the state-feedback expression is applied for $k \geq 1$, $B_1[k] = B_d$ for $k \geq 2$. Moreover, confirming the results of Theorem 5.3, $R[\cdot]$, $B_1[\cdot]$ and $B_2[\cdot]$ are bounded and the remaining signals converge as k tends to infinity.

Chapter 6

Summary and Future Work

This thesis introduced a new application of control systems engineering; specifically, the effects of rewards on an individual's attitudes and behaviour were modelled and control strategies were developed. The fundamental problem considered was whether or not there exists a sequence of reward values that is able to drive a person's behaviour to a desired, positive amount ($B_d > 0$). This problem was examined in the context of two discrete-time nonlinear models, the one-person system (Figure 1.2) and the two-person system (Figure 1.3).

To model the one-person system, three well-established theories from social psychology were used: the theory of planned behaviour, which models how attitudes combine to form a behaviour, cognitive dissonance theory, which models (among other things) how behaviour can influence attitudes, and the overjustification effect, which models the effect a reward has on a person with a positive attitude. The one-person system was then augmented to incorporate the effect of two additional external influences (social pressure and conformity pressure), which provided the framework for the the two-person system. In particular, by connecting two instances of the augmented model (through the conformity pressure exerted by each person's behaviour) the two-person system was formed. The main conclusion arising from the modelling stage is that each model has predictive power. In both the one-person and the two-person systems, the known conclusion that a medium-sized reward produces the most amount of immediate attitude change is extended to predict that this reward subsequently causes the most amount of future attitude and behaviour change. In the two-person system, known psychological phenomena arising from group dynamics were shown to be produced by dissonance pressure, a mechanism that is traditionally not linked to these phenomena. Even though the models developed in this thesis were shown to be qualitatively consistent with the relevant psychology, several simplifying assumptions were made. To make the model more realistic and, perhaps, to allow for experimental validation, future work can include:

- Differentiating between behaviour and intent to behave, possibly by incorporating randomness into the model;
- Using discrete values for $B[k]$;
- Generalizing the model for N people, including determining whether or not the nature of conformity pressure changes with respect to group size;
- Incorporating various dissonance reduction methods into the model, including an overall scheme for determining when each technique should be applied;
- Modelling how attitudes are influenced by factors such as motivation and emotion;
- Considering more than one behavioural outcome (which would generate more than one $A_{out}[k]$ for a single person);
- Modelling the effect of punishments; and
- Validating the model, possibly by using data from smoking cessation studies.

To study the effect of rewards on a person's attitudes and behaviour, a particular control objective was considered: finding a sequence of non-negative reward values that can ensure $B[k] \geq B_d$ as k tends to infinity (for the one-person system) or $B_1[k] \geq B_d$ as k tends to infinity (for the two-person system). Various open-loop and closed-loop controllers were designed to meet this control objective for both systems. The most notable result arising from these controller designs is that driving a person's behaviour positive as soon as possible is an effective strategy for achieving the overall control objective. Although this strategy proved effective, in most controller designs considered, it was shown to be only sufficient. Additionally, other intermediate steps taken throughout the various controller designs relied on sufficient conditions and simplifying assumptions. Furthermore, in some of the analysis, conjectures about the system response were made. Open problems include, specifically, the following:

- For the one-person system output-feedback controller, find necessary and sufficient conditions guaranteeing the system never returns to stage I after leaving stage I;
- For the one-person system output-feedback controller, confirm or refute Conjectures 3.1 and 3.2;
- For the two-person system zero-input response analysis, confirm or refute Conjecture 5.1;
- For the two-person system step-reward controller, find necessary and sufficient conditions to achieve the control objective;

- For the two-person system, for any set of initial conditions, find a state-feedback controller that meets the control objective while guaranteeing each signal is bounded; and
- Investigate the performance of the one-person output-feedback controller on the two-person system by considering the additional attitudes, $A_{others}[k]$ and $A_{conf}[k]$, as "disturbances" to the one-peson output-feedback system.

Finally, although one particular control problem was considered in this thesis, studying a dynamic attitude-behaviour model demonstrates the existence of many interesting control problems, including:

- Examining how multiple reward signals can be used to drive $B_1[k] \geq B_{d,1}$ and $B_2[k] \geq B_{d,2}$ as k tends to infinity;
- Controlling the internal attitude, $A_{out}[k]$, rather than the behaviour $B[k]$;
- Studying optimal control problems such as minimizing the total required reward while ensuring $B[k] \geq B_d$ as k tends to infinity; and
- Investigating robust closed-loop controllers that specifically fight disturbances and accommodate uncertainty in the model.

APPENDICES

Appendix A

Useful Mathematical Tools

In the analysis of both the one-person and the two-person systems, the following Lemmas are required. The first states that a non-increasing function that is bounded below converges and the second states that a non-decreasing function that is bounded above also converges. These two lemmas are standard results [1].

Lemma A.1. *Let* $a[\cdot]$ *be a sequence with* $a[i] \leq 0$ *for* $i \geq 0$. *Let*

$$h[k] = \sum_{i=0}^{k} a[i].$$

If $h[k]$ *is bounded below by* M_1, *then* $h[k]$ *converges to some constant,* $c_1 \geq M_1$.

Lemma A.2. *Let* $b[\cdot]$ *be a sequence with* $b[i] \geq 0$ *for* $i \geq 0$. *Let*

$$g[k] = \sum_{i=0}^{k} b[i].$$

If $g[k]$ *is bounded above by* M_2, *then* $g[k]$ *converges to some constant,* $c_2 \leq M_2$.

Appendix B

Summary of Assumptions

Throughout the thesis, several key assumptions are made, which are collected and summarized below.

B.1 Psychological Assumptions

Assumption 2.1: Behaviour and intent to behave are equal.

Assumption 2.2: There is exactly one behavioural outcome.

Assumption 2.3: No doubts or obstacles arise that would reduce the behavioural intent.

Assumption 2.4: Dissonance pressure reduction occurs through change in $A_{out}[k]$.

Assumption 4.1: Any type of social pressure can produce overjustification pressure.

Assumption 4.2: Overjustification pressure cannot arise due to conformity pressure.

B.2 Parameter Value Assumptions

B.2.1 One-Person System

The following assumptions hold for all analysis and simulation performed on the one-person system.

Assumption 2.5: Gains reflecting how the dissonance and overjustification pressures affect attitude change are strictly positive, i.e., $K_1 > 0$ and $K_2 > 0$.

Assumption 2.6: The value assigned to one dollar is strictly positive, i.e., $\mu_1 > 0$.

Assumption 2.7: The mental processing pole location for reward attitude formation in (2.3) is zero, i.e., $r_1 = 0$.

Assumption 2.8: The mental processing pole locations for dissonance and overjustification pressures in (2.11) and (2.14) respectively, are contained in the range $[0, 1)$, i.e., $0 \leq r_2, r_3 < 1$.

B.2.2 Two-Person System

The following assumptions hold for all analysis and simulation performed on the two-person system.

Assumption 4.3: $A_{rew,2}[k] = 0$ for $k \geq 0$.

Assumption 4.4: $A_{others,i}[k] = 0$ for $k \geq 0$, $i = 1, 2$.

Assumption 4.5: Gains reflecting how the dissonance and overjustification pressures affect attitude change are strictly positive, i.e., $K_{1i} > 0$, $K_{2i} > 0$ and $K_{3i} > 0$.

Assumption 4.6: The value assigned to one dollar is strictly positive, i.e., $\mu_{1i} > 0$.

Assumption 4.7: The extent to which each person in the two-person system experiences conformity pressure cannot be more than 100%, i.e., $0 \leq \mu_{31} < 1$ and $0 \leq \mu_{32} < 1$.

Assumption 4.8: The mental processing pole locations for reward attitude formation and the conformity attitude formation in (4.2) and (4.4) are zero, i.e., $r_{1i} = r_{3i} = 0$.

Assumption 4.9: The mental processing pole locations for dissonance and overjustification pressures in (4.16) and (4.20) and (4.21), are contained in the range $[0, 1)$, i.e., $0 \leq r_{4i}, r_{5i}, r_{6i} < 1$.

B.3 Other Simplifying Assumptions

B.3.1 One-Person System

The following assumptions are used to simplify the analysis of the one-person *output-feedback system*. These assumptions do not necessarily hold for simulations.

Assumption 3.1: The mental processing pole location for dissonance pressure in (2.11) is zero, i.e., $r_2 = 0$.

Assumption 3.2: The controller given in (3.21) can be tuned such that upon entering stage II, the one-person output-feedback system never returns to stage I, i.e., $B[k] > 0$ for $\hat{k} \leq k < \bar{k}$ and $B[k] \geq 0$ for $k \geq \bar{k}$.

B.3.2 Two-Person System

The following assumption is used to simplify the analysis of the two-person system. This assumption does not necessarily hold for simulations.

Assumption 4.10: The mental processing pole location for dissonance pressure in (4.16) is zero, i.e., $r_{4,i} = 0$ for $i = 1, 2$.

Appendix C

Detailed Proofs for Chapter 2

C.1 Proof of Lemma 2.1

Proof. To find the maximum dissonance pressure magnitude, all cases in which dissonance occurs are considered. For the given initial conditions, there are two situations in which dissonance occurs. To see this, note that if $A_{rew}[k] = 0$, then from (2.10), $P^{CD}_{raw}[k] = 0$; thus for raw, unprocessed dissonance pressure to arise, it is necessary that $A_{rew}[k] > 0$. Furthermore, if $A_{out}[k] \geq 0$ and $A_{rew}[k] > 0$, then from (2.4), $B[k] \geq 0$ and thus both cognitive pairs are consistent, i.e., $P^{CD}_{raw}[k] = 0$. Since $A_{rew}[k] > 0$, it follows that $A_{out}[k] < 0$ must hold to guarantee $P^{CD}_{raw}[k] \neq 0$. This leaves the following two situations in which dissonance pressure arises:

(i) $A_{rew}[k] > 0$, $A_{out}[k] < 0$ and $B[k] < 0$; and

(ii) $A_{rew}[k] > 0$, $A_{out}[k] < 0$ and $B[k] \geq 0$

Consider, first, the case when $A_{rew}[k] > 0$, $A_{out}[k] < 0$ and $B[k] < 0$. Since, from (2.4), $B[k] = A_{out}[k] + A_{rew}[k]$, and $B[k] < 0$, it follows that

$$A_{rew}[k] < -A_{out}[k]. \tag{C.1}$$

Furthermore, from (2.10), the *magnitude* of the raw, unprocessed dissonance pressure is

$$|P^{CD}_{raw}[k]| = \frac{M_{incon}[k]}{M_{incon}[k] + M_{con}[k]}.$$

Substituting in the expressions for $M_{incon}[k]$ and $M_{con}[k]$ from (2.6)–(2.9) and simplifying yields

$$|P^{CD}_{raw}[k]| = \frac{|A_{rew}[k]|}{|A_{rew}[k]| + |A_{out}[k]|}.$$

Additionally, since $A_{out}[k] < 0$, it follows that $|A_{out}[k]| = -A_{out}[k]$; therefore

$$|P_{raw}^{CD}[k]| = \frac{|A_{rew}[k]|}{|A_{rew}[k]| - A_{out}[k]}.$$

Using the inequality in (C.1), the denominator in the above expression can be replaced, giving

$$|P_{raw}^{CD}[k]| < \frac{|A_{rew}[k]|}{2|A_{rew}[k]|} = \frac{1}{2}.$$

Consider, instead, the second case: $A_{rew}[k] > 0$, $A_{out}[k] < 0$ and $B[k] \geq 0$. From the expression for $B[k]$ and the condition that $B[k] \geq 0$, it follows that

$$A_{rew}[k] \geq -A_{out}[k] = |A_{out}[k]|. \tag{C.2}$$

In this case, the *magnitude* of the raw, unprocessed dissonance pressure is calculated as above and, after some simplification, is given by

$$|P_{raw}^{CD}[k]| = \frac{|A_{out}[k]|}{|A_{out}[k]| + |A_{rew}[k]|}.$$

Using the inequality in (C.2), the denominator in the above expression can be replaced, giving

$$P_{raw}^{CD}[k] \leq \frac{|A_{out}[k]|}{2|A_{out}[k]|} = \frac{1}{2}.$$

In both cases, the magnitude of the raw, unprocessed dissonance pressure is always no more than 0.5, as predicted by Festinger. □

C.2 Proof of Lemma 2.2

Proof. From the expression for $A_{out}[k]$ and $\Delta A_{out}[k]$, given in (2.2) and (2.5) respectively, $A_{out}[k+1]$ can be expressed as

$$A_{out}[k+1] = A_{out}[k] + \Delta A_{out}^{CD}[k] + \Delta A_{out}^{OJ}[k].$$

If $P^{CD}[\bar{k}] \geq 0$, it follows from (2.12) that $\Delta A_{out}^{CD}[\bar{k}] \geq 0$. Therefore, it is evident that if $A_{out}[\bar{k}] + \Delta A_{out}^{OJ}[\bar{k}] \geq 0$, then $A_{out}[\bar{k}+1] \geq 0$.

From the conditions of the lemma, $A_{out}[\bar{k}] \geq 0$. Furthermore, from (2.15), $\Delta A_{out}^{OJ}[\bar{k}]$ has three possible expressions:

(i) If $P^{OJ}[\bar{k}] > 0$ and $K_2 P^{OJ}[\bar{k}] \leq A_{out}[\bar{k}]$, then $\Delta A_{out}^{OJ}[\bar{k}] = -K_2 P^{OJ}[\bar{k}]$.

(ii) If $P^{OJ}[\bar{k}] > 0$ and $K_2 P^{OJ}[\bar{k}] > A_{out}[\bar{k}]$, then $\Delta A_{out}^{OJ}[\bar{k}] = -A_{out}[\bar{k}]$.

(iii) If $P^{OJ}[\bar{k}] = 0$, then $\Delta A_{out}^{OJ}[\bar{k}] = 0$.

Note that due to (2.13) and (2.14), negative values of $P^{OJ}[\bar{k}]$ are impossible.

First, suppose that $P^{OJ}[\bar{k}] > 0$ and $K_2 P^{OJ}[\bar{k}] \leq A_{out}[\bar{k}]$; then,

$$\Delta A_{out}^{OJ}[\bar{k}] = -K_2 P^{OJ}[\bar{k}] \geq -A_{out}[\bar{k}].$$

Rearranging the above inequality gives $A_{out}[\bar{k}] + \Delta A_{out}^{OJ}[\bar{k}] \geq 0$, as required.

Instead, suppose that $P^{OJ}[\bar{k}] > 0$ and $K_2 P^{OJ}[\bar{k}] > A_{out}[\bar{k}]$; thus,

$$\Delta A_{out}^{OJ}[\bar{k}] = -A_{out}[\bar{k}];$$

hence $A_{out}[\bar{k}] + \Delta A_{out}^{OJ}[\bar{k}] = 0$, as required.

Finally, suppose for $P^{OJ}[\bar{k}] = 0$; then,

$$\Delta A_{out}^{OJ}[\bar{k}] = 0,$$

and thus, $A_{out}[\bar{k}] + \Delta A_{out}^{OJ}[\bar{k}] = A_{out}[\bar{k}] \geq 0$, as required.

This analysis shows that $A_{out}[\bar{k}+1] \geq 0$, which implies that $B[\bar{k}+1] \geq 0$ (owing to the fact that the model is only valid for non-negative rewards). Since $B[\bar{k}+1] \geq 0$, it follows that $P_{raw}^{CD}[\bar{k}+1] \geq 0$, implying $P^{CD}[\bar{k}+1] \geq 0$ and thus, $\Delta A_{out}^{CD}[\bar{k}+1] \geq 0$. Thus, conditions of the lemma again hold, and therefore, it must follow that $A_{out}[k] \geq 0$ for $k \geq \bar{k}$. □

Appendix D

Detailed Proofs for Chapter 3

D.1 Proof of Theorem 3.1

Proof. From (3.3), for $k \geq 2$, $B[k] = A_{out}[k]$. Thus $B[k]$ is a decreasing (increasing) function of k if and only if $A_{out}[k]$ is a decreasing (increasing) function of k. From (2.2)

$$A_{out}[k] = A_{out}[k-1] + \Delta A_{out}[k-1].$$

Along with (3.4), the above expression implies that if $\Delta A_{out}^{CD}[k] < 0$ for $k \geq 2$ then $A_{out}[k]$ is a decreasing function of k for $k \geq 2$. Moreover, if $\Delta A_{out}[k] > 0$ for $k \geq 2$ then $A_{out}[k]$ is an increasing function of k for $k \geq 2$.

First, suppose $R_o < \frac{-A_o}{\mu_1}$, implying $B[1] < 0$; then, to show $B[k]$ is a decreasing function for $k \geq 2$, it is enough to show that $\Delta A_{out}^{CD}[k] < 0$ for $k \geq 2$. If $B[1] < 0$, then it follows from the dissonance pressure equations, i.e., (2.10) and (2.11), that $P^{CD}[1] < 0$. Moreover, since for $k \geq 2$, $A_{rew}[k] = 0$, it follows that $P_{raw}^{CD}[k] = 0$ for $k \geq 2$, which, for $k \geq 2$, reduces the expression for $P^{CD}[k]$ in (2.11) to

$$P^{CD}[k] = r_2 P^{CD}[k-1].$$

Consequently, $P^{CD}[k] < 0$ for $k \geq 1$. This inequality implies that $\Delta A_{out}^{CD}[k] < 0$ for $k \geq 1$ as required; thus if $B[1] < 0$, then $B[k]$ is a decreasing function of k for $k \geq 2$.

Suppose instead that $R_o \geq \frac{-A_o}{\mu_1}$, implying $B[1] \geq 0$; then, to show $B[k]$ is an increasing function for $k \geq 2$, it is enough to show that $\Delta A_{out}^{CD}[k] > 0$ for $k \geq 2$. If $B[1] \geq 0$, then it follows from the dissonance pressure equations, i.e., (2.10) and (2.11), that $P^{CD}[1] > 0$. Using an argument similar to the previous case, this inequality implies that $P^{CD}[k] > 0$ for $k \geq 2$ and thus, $\Delta A_{out}^{CD}[k] > 0$ for $k \geq 2$, as required. □

D.2 Proof of Lemma 3.2

Proof. First, the existence of $A_{out}[\infty]$ is proven; then, using the expression for $A_{out}[\infty]$, the maximum steady-state value of $A_{out}[k]$ over values of R_o is determined, subject to the condition that $R_o \geq \frac{-A_o}{\mu_1}$.

Part 1

To show $A_{out}[k]$ converges as k tends to infinity, the expression for $A_{out}[k]$, given in (2.2), is studied. Since, from (3.4), the $\Delta A_{out}[k]$ term in (2.2) only depends on cognitive dissonance effects, the expression for $\Delta A_{out}^{CD}[k]$ can be directly substituted into (2.2). This substitution yields $A_{out}[k] = A_{out}[k-1] + K_1 P^{CD}[k-1]$. The above expression can be re-written as

$$A_{out}[k] = A_o + \sum_{i=0}^{k-1} K_1 P^{CD}[i].$$

From the initial conditions given in (2.16), $P^{CD}[0] = 0$; hence, the above expression becomes

$$A_{out}[k] = A_o + \sum_{i=1}^{k-1} K_1 P^{CD}[i]. \tag{D.1}$$

To show $A_{out}[k]$ converges as k tends to infinity, the expression for $P^{CD}[k]$ is first obtained. The expression for the raw, unprocessed dissonance pressure, given in (2.10) states that $P_{raw}^{CD}[k]$ is non-zero when $A_{out}[k] < 0$ and $A_{rew}[k] > 0$. However, $A_{rew}[k] = 0$ for $k \geq 2$; thus for $k \geq 2$, $P_{raw}^{CD}[k] = 0$ and $P^{CD}[k] = r_2 P^{CD}[k-1]$.

Since at $k = 1$, the raw, unprocessed dissonance pressure is non-zero, it follows that $P^{CD}[1] = (1 - r_2)\, P_{raw}^{CD}[1]$. From this expression, the dissonance pressure equation for $k \geq 1$ can be obtained in terms of $P_{raw}^{CD}[1]$; specifically,

$$P^{CD}[k] = r_2^{k-1}\,(1 - r_2)\, P_{raw}^{CD}[1].$$

Substituting this into (D.1) gives

$$A_{out}[k] = A_o + K_1\,(1 - r_2)\, P_{raw}^{CD}[1] \sum_{i=1}^{k-1} r_2^{i-1},$$

which, after rearranging the summation limits, produces

$$A_{out}[k] = A_o + K_1\,(1 - r_2)\, P_{raw}^{CD}[1] \sum_{i=0}^{k-2} r_2^{i}.$$

The summation term above is simply the sum of geometric series, which can be simplified as

$$\sum_{i=0}^{k-2} r_2^i = \frac{1 - r_2^{k-1}}{1 - r_2}.$$

Substituting the equation above into the expression for $A_{out}[k]$ results in the following equation, valid for $k \geq 2$:

$$A_{out}[k] = A_o + \left(K_1 \left(1 - r_2\right) P_{raw}^{CD}[1]\right) \left(\frac{1 - r_2^{k-1}}{1 - r_2}\right).$$

From Table 2.2, $r_2 \in [0, 1)$ and hence, the right-hand side of the above expression converges as k tends to infinity, thereby allowing limits to be taken of the above expression:

$$\begin{aligned} &\lim_{k \to \infty} A_{out}[k] = A_o + \left(K_1 P_{raw}^{CD}[1]\right) \lim_{k \to \infty} \left(1 - r_2^{k-1}\right) \\ \Rightarrow &\lim_{k \to \infty} A_{out}[k] = A_o + K_1 P_{raw}^{CD}[1]. \end{aligned} \tag{D.2}$$

Thus, $A_{out}[k]$ converges as k tends to infinity.

Part 2

Solving $\max_{\substack{Ro \\ R_o \geq \frac{-A_o}{\mu_1}}} A_{out}[\infty]$ follows simply from the expression for $P_{raw}^{CD}[1]$ and the inequality $R_o \geq \frac{-A_o}{\mu_1}$. Note that to guarantee a maximum value for $A_{out}[\infty]$, the condition that $R_o \geq \frac{-A_o}{\mu_1}$ must hold, since from Theorem 3.1, if $R_o < \frac{-A_o}{\mu_1}$, then (due to $A_{out}[k]$ decreasing) $B[k]$ is a decreasing function of k. From this result, it is clear that $A_{out}[k]$ only increases if $R_o \geq \frac{-A_o}{\mu_1}$. From the condition that $R_o \geq \frac{-A_o}{\mu_1}$, it follows that $B[1] \geq 0$ and thus, from the model equations that the raw, unprocessed dissonance pressure at $k = 1$ is given by

$$P_{raw}^{CD}[1] = \frac{|A_{out}[1]|}{|A_{out}[1]| + |A_{rew}[1]|}.$$

Since there is no attitude change at $k = 0$, it follows that $A_{out}[1] = A_o$. By substituting this value for $A_{out}[1]$ into $P_{raw}^{CD}[1]$ in (D.2), the problem becomes determining

$$\max_{\substack{Ro \\ R_o \geq -\frac{A_o}{\mu_1}}} \left\{ A_o + K_1 \frac{|A_o|}{|A_o| + |\mu_1 R_o|} \right\}.$$

Since a smaller R_o yields a larger $A_{out}[\infty]$ and since R_o must be greater than some minimum value, it follows that $A_{out}[\infty]$ is maximized subject to the given conditions when

$R_o = -\frac{A_o}{\mu_1}$. Note that $R_o > 0$ holds because $A_o < 0$ and $\mu_1 > 0$. Substituting the above expression for R_o into (D.2) gives

$$A_{out}[\infty] = A_o + K_1 \frac{|A_o|}{|A_o| + |\mu_1 \left(-\frac{A_o}{\mu_1}\right)|}.$$

Rearranging the above expression presents the conclusion that

$$\max_{\substack{Ro \\ R_o \geq \frac{-A_o}{\mu_1}}} A_{out}[\infty] = A_o + \frac{K_1}{2}.$$

□

D.3 Proof of Theorem 3.2

Proof. To determine whether or not an impulse-reward controller is able to meet the control objective, (3.3) and Lemma 3.2 are used. From (3.3), $B[k] = A_{out}[k]$ for $k \geq 2$. Furthermore, if $R_o \geq \frac{-A_o}{\mu_1}$, then from Lemma 3.2, $A_{out}[k]$ converges as k tends to infinity, and therefore $B[k]$ converges as k tends to infinity. From (3.3), these two signals converge to the same value. From Lemma 3.2, the maximum value to which $A_{out}[k]$ can converge is $A_o + \dfrac{K_1}{2}$. If B_d is larger than this maximum value, then it is impossible for $B[k] \geq B_d$ as k tends to infinity. Consequently, it is possible to drive $B[k] \geq B_d$ as k tends to infinity if and only if

$$A_o + \frac{K_1}{2} \geq B_d.$$

To find the set of R_o values that are able to meet the control objective, two conditions are examined: $R_o \geq \frac{-A_o}{\mu_1}$ and $B[\infty] \geq B_d$. The first condition is necessary for producing a positive $B[k]$ as k tends to infinity (but is not sufficient). The second condition is the control objective.

From (3.3), the second condition is equivalent to $A_{out}[\infty] \geq B_d$. Using the expression for $A_{out}[\infty]$, obtained from Lemma 3.2, the second condition is re-stated as

$$A_o + \frac{K_1|A_o|}{|A_o| + |\mu_1 R_o|} \geq B_d.$$

Rearranged for R_o gives

$$R_o \leq \frac{|A_o|}{\mu_1}\left(\frac{K_1}{B_d - A_o} - 1\right). \tag{D.3}$$

Since both $R_o \geq \frac{-A_o}{\mu_1}$ and (D.3) must hold simultaneously, the set of R_o values that are able to meet the control objective is given by

$$\frac{|A_o|}{\mu_1} \leq R_o \leq \frac{|A_o|}{\mu_1}\left(\frac{K_1}{B_d - A_o} - 1\right).$$

□

D.4 Proof of Theorem 3.3

Proof. The proof is presented in two parts, one for $R_o < \frac{-A_o}{\mu_1}$ and the other for $R_o \geq \frac{-A_o}{\mu_1}$.

Part 1

This part of the proof uses induction to show that if $R_o < \frac{-A_o}{\mu_1}$, then $B[k]$ tends to negative infinity as k tends to infinity. First, suppose $R_o < \frac{-A_o}{\mu_1}$, implying $B[1] < 0$; then, from the model equations and the reward attitude expression given (3.12), $B[1]$ and $B[2]$ are given by

$$\begin{aligned} B[1] &= A_{out}[1] + \mu_1 R_o, \\ B[2] &= A_{out}[1] + \Delta A_{out}[1] + \mu_1 R_o. \end{aligned}$$

Since $B[1] < 0$ and $A_{rew}[1] > 0$, dissonance pressure arises and thus, from (2.10) and (2.11), $P^{CD}[1] < 0$, which then implies that $\Delta A_{out}^{CD}[1] < 0$. From the initial condition that $B[0] < 0$ and from the supposition that $B[1] < 0$, Lemma 3.1 applies for $0 \leq k \leq 1$; thus $\Delta A_{out}[1] = \Delta A_{out}^{CD}[1] < 0$. By comparing the expressions for $B[1]$ and $B[2]$, clearly $B[2] < B[1]$, which implies $B[2] < 0$.

More generally, suppose $B[k] < B[k-1]$ for $1 \leq k < \bar{k}$, which implies $B[k] < 0$ for $0 \leq k < \bar{k}$. It follows from the dissonance pressure and attitude change equations that $\Delta A_{out}^{CD}[k] < 0$ for $0 \leq k < \bar{k}$. Furthermore, from Lemma 3.1, $\Delta A_{out}[k] = \Delta A_{out}^{CD}[k] < 0$ for $0 \leq k < \bar{k}$.

From (2.2) and (2.4),

$$\begin{aligned} B[\bar{k}-1] &= A_{out}[\bar{k}-1] + \mu_1 R_o, \text{ and} \\ B[\bar{k}] &= A_{out}[\bar{k}-1] + \Delta A_{out}[\bar{k}-1] + \mu_1 R_o. \end{aligned}$$

Thus, $B[\bar{k}] = B[\bar{k}-1] + \Delta A_{out}[\bar{k}]$. Due to the inductive assumption, $B[\bar{k}-1] < 0$ and thus from (2.10), (2.11), (2.12) and Lemma 3.1, $\Delta A_{out}[\bar{k}] < 0$; hence, $B[\bar{k}] < B[\bar{k}-1]$. Therefore, by induction, it has been shown that if $R_o < \frac{-A_o}{\mu_1}$, then $B[k]$ is a decreasing function of k for $k \geq 1$, which implies that either $B[k]$ tends to negative infinity as k tends to infinity, or $B[k]$ converges as k tends to infinity.

To show $B[k]$ does not converge, contradiction is used; thus, suppose $B[k]$ converges as k tends to infinity. From (2.4), this implies that for some $c \leq 0$, $\lim_{k\to\infty} A_{out}[k] = c$. The expression for $A_{out}[k]$, given in (2.2), can be re-written as

$$A_{out}[k] = A_o + \sum_{i=0}^{k-1} \Delta A_{out}[i]. \tag{D.4}$$

Since $A_{out}[k] < 0$ for $k \geq 0$, Lemma 3.1 applies, i.e., $\Delta A_{out}[k] = \Delta A_{out}^{CD}[k]$. By substituting the expression for $\Delta A_{out}^{CD}[k]$ given in (2.12) into (D.4), the following expression is obtained:

$$A_{out}[k] = A_o + K_1 \sum_{i=0}^{k-1} P^{CD}[i].$$

Since the left-hand side of the above expression converges to c as k tends to infinity, limits can be taken on both sides of the equation, giving

$$\lim_{k\to\infty} A_{out}[k] = A_o + K_1 \lim_{k\to\infty} \sum_{i=0}^{k-1} P^{CD}[i].$$

Given that $A_{out}[k]$ converges as k tends to infinity, it necessarily follows that $P^{CD}[k]$ tends to zero as k tends to infinity. This conclusion will be used shortly.

Since $B[1] < 0$ and $B[k]$ is a decreasing function, $B[k] < 0$ for $k \geq 0$. Likewise, $A_{out}[k] < 0$ for $k \geq 0$. Since these two inequalities hold for $k \geq 0$, the raw, unprocessed dissonance pressure, $P_{raw}^{CD}[k]$, has a specific form for $k \geq 0$. In particular,

$$P_{raw}^{CD}[k] = -\frac{\mu_1 R_o}{|A_{out}[k]| + \mu_1 R_o}.$$

Substituting this equation into the expression for $P^{CD}[k]$ gives

$$P^{CD}[k] = r_2 P^{CD}[k-1] - \frac{(1-r_2)\,\mu_1 R_o}{|A_{out}[k]| + \mu_1 R_o}.$$

Furthermore, since $P^{CD}[k]$ tends to zero as k tends to infinity, limits can be taken on both sides of the above equation, giving

$$\begin{aligned}
&\lim_{k\to\infty} P^{CD}[k] = \lim_{k\to\infty} \left(r_2 P^{CD}[k-1] - \frac{(1-r_2)\mu_1 R_o}{|A_{out}[k]| + \mu_1 R_o} \right). \\
\Rightarrow\,& 0 = \lim_{k\to\infty} \frac{(1-r_2)\mu_1 R_o}{|A_{out}[k]| + \mu_1 R_o}, \\
\Rightarrow\,& \lim_{k\to\infty} \frac{(1-r_2)\mu_1 R_o}{|A_{out}[k]| + \mu_1 R_o} = 0, \\
\Rightarrow\,& \lim_{k\to\infty} |A_{out}[k]| = \infty.
\end{aligned}$$

This violates the assumption that $A_{out}[k]$ settles to a constant as k tends to infinity. Therefore, $A_{out}[k]$ does not converge as k tends to infinity. Since $A_{out}[k]$ is a decreasing function that does not converge, $A_{out}[k]$ tends to $-\infty$ as $k \to \infty$.

Therefore, if $B[1] < 0$ then $B[k]$ tends to negative infinity as k tends to infinity.

Part 2

This part of the proof uses induction to show that if $R_o \geq \frac{-A_o}{\mu_1}$, then $B[k]$ is an increasing function for $0 \leq k \leq T$. From the initial conditions, $B[0] = A_o < 0$. Suppose $R_o \geq \frac{-A_o}{\mu_1}$, implying $B[1] \geq 0$; then, $B[1] > B[0]$. Moreover, as in the previous part, $B[1]$ and $B[2]$ can be expressed as

$$\begin{aligned} B[1] &= A_{out}[1] + \mu_1 R_o, \\ B[2] &= A_{out}[1] + \Delta A_{out}[1] + \mu_1 R_o. \end{aligned}$$

Since $B[1] \geq 0$ and $A_{out}[1] < 0$, dissonance pressure arises and thus, $P^{CD}[1] > 0$, which then implies that $\Delta A^{CD}_{out}[1] > 0$. It follows directly from the initial conditions that for $0 \leq k \leq 1$, $A_{out}[k] = A_o < 0$; hence, Lemma 3.1 applies over this range of k values. In particular, $\Delta A_{out}[1] = \Delta A^{CD}_{out}[1]$. Since $\Delta A^{CD}_{out}[1] > 0$, it follows from the expressions for $B[1]$ and $B[2]$ that $B[2] > B[1]$.

More generally, suppose $B[k] > B[k-1]$ for $1 \leq k < \hat{k}$ where $\hat{k} \leq T$; then, it follows that $B[k] > 0$ for $1 \leq k \leq \hat{k}$; thus $P^{CD}[k] > 0$ and $\Delta A^{CD}_{out}[k] > 0$. Furthermore, since $\hat{k} \leq T$ and T is defined as the largest value of k such that $A_{out}[k] < 0$ for $0 \leq k \leq T$, it follows that $A_{out}[k] < 0$ for $0 \leq k < \hat{k}$, and therefore, Lemma 3.1 applies, i.e, $\Delta A_{out}[k] = \Delta A^{CD}_{out}[k]$ for $0 \leq k < \hat{k}$.

Now use induction. Start with

$$\begin{aligned} B[\hat{k}-1] &= A_{out}[\hat{k}-1] + \mu_1 R_o, \\ B[\hat{k}] &= A_{out}[\hat{k}-1] + \Delta A_{out}[\hat{k}-1] + \mu_1 R_o. \end{aligned}$$

Since $\Delta A_{out}[\hat{k}-1] = \Delta A^{CD}_{out}[\hat{k}-1]$ and $\Delta A_{out}[\hat{k}-1] > 0$, it follows that $B[\hat{k}] > B[\hat{k}-1]$. Hence, by induction, $B[k]$ is an increasing function of k for $0 \leq k \leq T$. □

D.5 Proof of Lemma 3.3

Proof. Suppose $R_o \geq \frac{-A_o}{\mu_1}$, implying $B[k] \geq 0$ for $k \geq 1$. Then, from (2.10), $P^{CD}_{raw}[k] > 0$ for $k \geq 1$, implying $P^{CD}[k] > 0$ and thus $A^{CD}_{out}[k] > 0$ for $k \geq 1$. Consider, first, the case when

$A_{out}[k] \leq 0$ for all $k \geq 0$. Then, $P^{OJ}[k] = 0$ and thus $\Delta A_{out}^{OJ}[k] = 0$ for $k \geq 0$, i.e., $\Delta A_{out}^{OJ}[k]$ converges to zero as k tends to infinity. Furthermore, since $A_{out}[k]$ can be expressed as

$$A_{out}[k] = A_o + \sum_{i=0}^{k-1} K_1 P^{CD}[i],$$

$A_{out}[k]$ is an increasing function of k for $k \geq 1$. From the condition that $A_{out}[k] \leq 0$ for $k \geq 0$ and given that $A_{out}[k]$ is an increasing function of k, Lemma A.2 applies; hence $A_{out}[k]$ converges as k tends to infinity. From the above expression for $A_{out}[k]$, this conclusion implies that $P^{CD}[k]$ tends to zero as k tends to infinity and, as a result, $\Delta A_{out}^{CD}[k]$ tends to zero as k tends to infinity. In summary, if $A_{out}[k] \leq 0$ for $k \geq 1$, then, as k tends to infinity, $A_{out}[k]$ converges and $\Delta A_{out}^{CD}[k]$ and $\Delta A_{out}^{OJ}[k]$ tend to zero.

Suppose instead, that $A_{out}[k] > 0$ at some $k = \bar{k}$. Then, from Lemma 2.2, $A_{out}[k] \geq 0$ for all $k \geq \bar{k}$; hence, overjustification pressure may arise and, thus, must be considered in the expression for $A_{out}[k]$, which, for $k > \bar{k}$, can be stated as

$$A_{out}[k] = A_{out}[\bar{k}] + \sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{CD}[i] + \sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{OJ}[i]. \tag{D.5}$$

To show $A_{out}[k]$ converges as k tends to zero, each summation term in the above expression is examined.

First, consider the summation term related to $P^{CD}[k]$, i.e., the one containing $\Delta A_{out}^{CD}[k]$. Like the case when $A_{out}[k] \leq 0$, $B[k] \geq 0$ for $k \geq 1$ implies $P^{CD}[k] > 0$ and $\Delta A_{out}^{CD}[k] > 0$ for $k \geq 1$. Furthermore, since for $k \geq \bar{k}$, $A_{rew}[k] \geq 0$ and $A_{out}[k] \geq 0$, it follows that $P_{raw}^{CD}[k] = 0$ over this sample range; thus, for $k \geq \bar{k}$, $P^{CD}[k] = r_2 P^{CD}[k-1]$. Solving this equation in terms of $P^{CD}[\bar{k}]$ gives, for $k \geq \bar{k}$,

$$P^{CD}[k] = r_2^{k-\bar{k}} P^{CD}[\bar{k}]. \tag{D.6}$$

In other words, after attitude becomes positive, no more raw, unprocessed dissonance pressure arises and therefore, the dissonance pressure is decaying according to r_2. Since $r_2 \in [0, 1)$, it follows that $P^{CD}[k]$ converges to zero as k tends to infinity; thus $\Delta A_{out}^{CD}[k]$ tends to zero as k tends to infinity. Moreover, the first summation term in (D.5) can also be shown to converge. Substituting (D.6) gives the following for the first summation term:

$$\begin{aligned}
\sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{CD}[i] &= K_1 \sum_{i=\bar{k}}^{k-1} r_2^{i-\bar{k}} P^{CD}[\bar{k}] \\
&= K_1 P^{CD}[\bar{k}] \sum_{i=\bar{k}}^{k-1} r_2^{i-\bar{k}} \\
&= K_1 P^{CD}[\bar{k}] \frac{1 - r_2^{k-\bar{k}}}{1 - r_2}.
\end{aligned}$$

Since $|r_2| < 1$, $\frac{1-r_2^{k-\bar{k}}}{1-r_2}$ converges as k tends to infinity and, therefore the first summation term in (D.5) converges as k tends to infinity.

Now, the convergence of the second summation term in (D.5) is considered. Since $A_{out}[k] \geq 0$ for $k \geq \bar{k}$ and (D.5) gives an expression for $A_{out}[k]$ for $k > \bar{k}$, set the $A_{out}[k] \geq 0$ in (D.5), giving

$$A_{out}[\bar{k}] + \sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{CD}[i] + \sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{OJ}[i] \geq 0.$$

Rearranging the above inequality gives

$$\sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{OJ}[i] \geq -\left(A_{out}[\bar{k}] + \sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{CD}[i] \right).$$

From before, the summation term on the right-hand side of the above inequality can be written in a different form, giving

$$\sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{OJ}[i] \geq -\left(A_{out}[\bar{k}] + K_1 P^{CD}[\bar{k}] \frac{1-r_2^{k-\bar{k}}}{1-r_2} \right).$$

As k tends to infinity, the last term in the above inequality converges. Therefore, both terms in the right-hand side of the above inequality are constant as k tends to infinity, and thus, it follows that the summation term on the left-hand side is bounded from below. From (2.15) the summation term on the left-hand side is also a non-increasing function. Therefore, Lemma A.1 applies since the left-hand side is a non-increasing function bounded from below; hence, this sum converges as k tends to infinity. As a result, the term inside the sum, $\Delta A_{out}^{OJ}[k]$, converges to zero as k tends to infinity.

Since the two sums in (D.5) converge as k tends to infinity, $A_{out}[k]$ converges as k tends to infinity. □

D.6 Proof of Lemma 3.4

Proof. The proof to show attitude converges to zero as k tends to infinity is done in two parts. The first part considers the case when $A_{out}[k] \leq 0$ for all k; the second part considers the case when $A_{out}[k] > 0$ at some $k = \bar{k}$.

Part 1

Suppose $A_{out}[k] \leq 0$ for $k \geq 0$; then, Lemma 3.1 applies and, as a result, $\Delta A_{out}[k] = \Delta A_{out}^{CD}[k]$ for $k \geq 0$. The equation for $A_{out}[k]$, given in (2.2), then simplifies to

$$A_{out}[k] = A_o + K_1 \sum_{i=0}^{k-1} P^{CD}[i].$$

Since $B[1] \geq 0$, Theorem 3.3 states that $B[k]$ is an increasing function for $0 \leq k \leq T$. Given the definition of T, clearly $T = \infty$ and, therefore, $B[k]$ is an increasing function of k for $k \geq 0$ and thus, $B[1] \geq 0$ implies $B[k] \geq 0$ for $k \geq 1$. As a result, Lemma 3.3 applies, stating that $A_{out}[k]$ converges to some constant, c, as k tends to infinity. Given that $A_{out}[k] \leq 0$ for $k \geq 1$, naturally, $c \leq 0$. Moreover, Lemma 3.3 states that $\Delta A_{out}^{CD}[k]$ tends to zero as k tends to infinity, which implies $P^{CD}[k]$ tends to zero as k tends to infinity. By examining the steady-state expression for $P^{CD}[k]$, the value of c is obtained.

Since $A_{out}[k] \leq 0$, $B[k] \geq 0$ and $A_{rew}[k] = \mu_1 R_o > 0$ for $k \geq 1$, from (2.10) and (2.11), $P^{CD}[k]$ can be expressed as

$$P^{CD}[k] = r_2 P^{CD}[k-1] + \frac{(1-r_2)|A_{out}[k]|}{|A_{out}[k]| + \mu_1 R_o}. \tag{D.7}$$

Since $P^{CD}[k]$ tends to zero and $A_{out}[k]$ tends to c as k tends to infinity, both sides of the above equation are well behaved, thus taking limits gives

$$0 = \lim_{k\to\infty} \frac{(1-r_2)|A_{out}[k]|}{|A_{out}[k]| + \mu_1 R_o}.$$

Therefore, it follows that $A_{out}[k]$ necessarily converges to zero as k tends to infinity if $A_{out}[k] \leq 0$ for $k \geq 0$.

Part 2

Now, suppose at some $k = \bar{k}$, $A_{out}[k] > 0$. Since $A_{out}[k]$ is an increasing function of k for $0 \leq k \leq T$, where T is the last sample at which $A_{out}[k] < 0$, it follows that $\bar{k} = T + 1$. Furthermore, since $A_{out}[k]$ is increasing, and since the overjustification effect does not arise for negative values of $A_{out}[k]$, it must be true that at $k = \bar{k}$, $P^{CD}[k] > 0$. Furthermore, since $A_{rew}[k] > 0$ for $k \geq 1$, the conditions of Lemma 2.2 hold at $k = \bar{k}$, i.e., $A_{out}[k] \geq 0$ for $k \geq \bar{k}$. As a result, it follows that $B[k] > 0$ for $k \geq \bar{k}$. Combining this with the results from Part 1 shows that $B[k] \geq 0$ for $k \geq 1$ and, thus, Lemma 3.3 can be applied. Therefore, $A_{out}[k]$ converges as k tends to infinity and both $\Delta A_{out}^{CD}[k]$ and $\Delta A_{out}^{OJ}[k]$ tend to zero as k tends to infinity. To find the steady-state value of $A_{out}[k]$, the expression for $\Delta A_{out}^{OJ}[k]$ is examined. From (2.15), there are three possible expressions for $\Delta A_{out}^{OJ}[k]$ for $k \geq \bar{k}$:

$$\Delta A_{out}^{OJ}[k] = -K_2 P^{OJ}[k], \tag{D.8}$$

$$\Delta A_{out}^{OJ}[k] = -A_{out}[k], \text{ and} \tag{D.9}$$

$$\Delta A_{out}^{OJ}[k] = 0. \tag{D.10}$$

Clearly, there are several possible expression sequences for $\Delta A_{out}^{OJ}[k]$, as switching between the three cases above may occur. However, since at $k = \bar{k}$, $A_{out}[k] > 0$, $A_{rew}[k] > 0$ and $B[k] > 0$, it follows directly from the model equations that $P^{OJ}[k] > 0$. As a result, $P^{OJ}[k] > 0$ for all $k \geq \bar{k}$ (since $P^{OJ}[k] = r_3 P^{OJ}[k-1] + (1 - r_3)\, P_{raw}^{OJ}[k]$ and $P_{raw}^{OJ}[k]$ cannot be negative). From the expression for $\Delta A_{out}^{OJ}[k]$, if $P^{OJ}[k] > 0$ for $k \geq \bar{k}$, then (D.10) cannot hold, as it requires $P^{OJ}[k] = 0$ and hence, only (D.8) and (D.9) are permissible expressions for $\Delta A_{out}^{OJ}[k]$ for $k \geq \bar{k}$.

The two remaining expressions for $\Delta A_{out}^{OJ}[k]$, i.e., (D.8) and (D.9), can combine in three possible ways:

(i) For $k \geq \bar{k}$, (D.8) applies;

(ii) For $\bar{k} \leq k < \hat{k}$, (D.8) applies, and for $k \geq \hat{k}$, (D.9) applies; and

(iii) For $k \geq \bar{k}$, (D.9) applies.

To see this, note that if at $k = \bar{k}$, (D.9) holds true, then $A_{out}[k{+}1] = 0$. Since $P^{OJ}[k{+}1] > 0$, it follows that $K_2 P^{OJ}[k+1] > A_{out}[k+1] = 0$ and thus, (D.9) holds again at $k = \bar{k} + 1$. Therefore, it is impossible to switch from (D.9) to (D.8). It is however, possible to switch from the (D.8) to (D.9).

Finally, it only remains to show that the convergence of $\Delta A_{out}^{OJ}[k]$ to zero as k tends to infinity implies that for each of the possible combinations above, $A_{out}[k]$ necessarily converges to zero as k tends to infinity. Suppose the first of the three combinations occur. Then,

$$\begin{aligned} \Delta A_{out}^{OJ}[k] \to 0 \text{ as } k \to \infty &\Rightarrow -K_2 P^{OJ}[k] \to 0 \text{ as } k \to \infty, \\ &\Rightarrow P^{OJ}[k] \to 0 \text{ as } k \to \infty. \end{aligned}$$

Since $A_{out}[k] \geq 0$, $B[k] > 0$, and $A_{rew}[k] > 0$ for $k \geq \bar{k}$, $P^{OJ}[k]$ is given by

$$P^{OJ}[k] = r_3 P^{OJ}[k-1] + (1 - r_3) A_{out}[k] A_{rew}[k]$$

for $k \geq \bar{k}$. Since the left-hand side of the above expression is well-behaved as k tends to infinity, the right-hand is also well-behaved. Therefore, limits can be taken on of both sides, giving

$$\lim_{k\to\infty} P^{OJ}[k] = \lim_{k\to\infty} \left(r_3 P^{OJ}[k-1] + (1 - r_3)\, A_{out}[k] A_{rew}[k] \right)$$

Given that $A_{rew}[k]$ is a constant and $P^{OJ}[k]$ converges to zero as k tends to infinity, the above expression implies $A_{out}[k]$ tends to zero as k tends to infinity.

Suppose instead that the second of the three combinations occur. Given that the steady-state trend of $A_{out}[k]$ is being considered, only values for $k \geq \hat{k}$ need to be considered (because as previously shown for $\bar{k} \leq k < \hat{k}$, $A_{out}[k]$ is well-behaved). Thus, suppose $k \geq \hat{k}$, then, using (D.9) as the expression for $\Delta A_{out}^{OJ}[k]$,

$$\Delta A_{out}^{OJ}[k] \rightarrow 0 \text{ as } k \rightarrow \infty \Rightarrow -A_{out}[k] \rightarrow 0 \text{ as } k \rightarrow \infty,$$

i.e., $A_{out}[k]$ necessarily tends to zero as k tends to infinity.

Finally, suppose the last of the three combinations occur; then, similar to the previous case, $A_{out}[k]$ tends to zero as k tends to infinity. Since $A_{out}[k]$ tends to zero as k tends to infinity for any admissible sequences of $\Delta A_{out}^{OJ}[k]$

$$\Delta A_{out}^{OJ}[k] \rightarrow 0 \text{ as } k \rightarrow \infty \Rightarrow A_{out}[k] \rightarrow 0 \text{ as } k \rightarrow \infty.$$

By combining the results of Part 1 and Part 2, $A_{out}[k]$ necessarily tends to zero as k tends to infinity. □

D.7 Proof of Theorem 3.4

Proof. The expression for $B[k]$, from (2.4) and (3.12), is given by $B[k] = A_{out}[k] + \mu_1 R_o$. $B[\infty]$ exists if and only if $A_{out}[\infty]$ exists. Suppose $R_o \geq \frac{-A_o}{\mu_1}$. From Lemma 3.4, $A_{out}[k]$ tends to zero as k tends to infinity, and therefore $B[k]$ converges to $\mu_1 R_o$. Thus, if $R_o \geq \frac{-A_o}{\mu_1}$, then the control objective is met if and only if $R_o \geq \frac{B_d}{\mu_1}$.

On the other hand, if $R_o < \frac{-A_o}{\mu_1}$, then Theorem 3.3 implies that the control objective cannot be achieved. Hence, it is achieved if and only if

$$R_o \geq \max\left(\frac{-A_o}{\mu_1}, \frac{B_d}{\mu_1}\right).$$

□

D.8 Proof of Theorem 3.5

Proof. First, to show the given range of R_o is necessary, suppose

$$R_o < \frac{-A_o}{\mu_1}. \tag{D.11}$$

From the model equations and the initial conditions, these values of R_o produce a negative value for $B[1]$. To see this, note $B[1] = A_o + \mu_1 R[0]$. Since $A_{out}[1] = A_o < 0$, it follows from (3.14) that $R[0] = R_o$; hence,

$$B[1] < A_o - \mu_1 \frac{A_o}{\mu_1} < 0.$$

Since $B[1] < 0$, $A_{out}[1] < 0$ and $A_{rew}[1] > 0$, it follows from the model equations that $P^{CD}_{raw}[1] < 0$ and therefore, $P^{CD}[1] < 0$ and $\Delta A^{CD}_{out}[1] < 0$. Furthermore, from Lemma 3.1, $\Delta A^{OJ}_{out}[1] = 0$ and therefore,

$$A_{out}[2] = A_{out}[1] + \Delta A^{CD}_{out}[1] < A_{out}[1],$$

i.e., the internal attitude decreases. This decrease causes $B[2] < B[1]$ owing to the fact that $A_{rew}[k]$ is a constant as long as $A_{out}[k] < 0$, which, in this case, it is. This argument applies at each sample $k \geq 1$ and therefore, if R_o satisfies (D.11), then the reward is simply a step reward and, from Theorem 3.3, $B[k]$ decreases and tends to negative infinity as k tends to infinity. Clearly, the reward is not able to drive $B[k]$ positive and therefore, $R_o \geq \frac{-A_o}{\mu_1}$ is a necessary condition.

Now, we show that $R_o \geq \frac{-A_o}{\mu_1}$ is sufficient for driving $B[k] \geq 0$ for $k \geq 1$. From the controller equation, $R[k] = R_o$ when $A_{out}[k]+\Delta A_{out}[k] < 0$. Since $A_{out}[0]+\Delta A_{out}[0] = A_o < 0$, it follows that $R[0] = R_o$. Suppose $R_o \geq \frac{-A_o}{\mu_1}$; then $B[1] = A_{out}[0]+\Delta A_{out}[0]+\mu_1 R_o \geq 0$. As a result, $P^{CD}[1] \geq 0$. Note that since $A_{out}[1] = A_o < 0$, $\Delta A^{OJ}_{out}[1] = 0$; hence $\Delta A_{out}[1] = \Delta A^{CD}_{out} \geq 0$. Therefore

$$A_{out}[2] = A_{out}[1] + \Delta A_{out}[1] \geq A_{out}[1].$$

If $A_{out}[1] + \Delta A_{out}[1] < 0$, then $R[1] = R_o$, in which case

$$\begin{aligned} B[2] &= A_{out}[2] + \mu_1 R_o \\ &\geq A_{out}[1] + \mu_1 R_o \\ &= B[1] \\ &\geq 0. \end{aligned}$$

Repeating this argument for each $k < T$, i.e., each k such that $A_{out}[k] + \Delta A_{out}[k] < 0$, implies that for $1 \leq k \leq T$, $B[k] \geq 0$. When $k = T$,

$$A_{out}[T] + \Delta A_{out}[T] \geq 0,$$

and therefore, the controller equation forces $R[T] = 0$; hence,

$$B[T+1] = A_{out}[T] + \Delta A_{out}[T] + 0 \geq 0.$$

Moreover, by Lemma 2.2, $A_{out}[k]+\Delta A_{out}[k] \geq 0$ for $k \geq T$. Therefore, for $k \geq T$, $R[k] = 0$ and $B[k+1] = A_{out}[k] + \Delta A_{out}[k] \geq 0$.

Combining the results above, $R_o \geq \frac{-A_o}{\mu_1}$ is a necessary and sufficient condition to ensure $B[k] \geq 0$ for $k \geq 1$. □

D.9 Proof of Theorem 3.6

Proof. The proof to show $B[k] \geq B_d$ for $k \geq 1$ is straight-forward. Suppose $R[k-1] = \frac{1}{\mu_1}\left(B_d - A_{out}[k-1] - \Delta A_{out}[k-1]\right)$. Then,

$$\begin{aligned} B[k] &= A_{out}[k-1] + \Delta A_{out}[k-1] + \mu_1 R[k-1] \\ &= A_{out}[k-1] + \Delta A_{out}[k-1] + B_d - A_{out}[k-1] - \Delta A_{out}[k-1] \\ &= B_d. \end{aligned}$$

Alternatively, suppose $R[k-1] = 0$, which occurs when $B_d - A_{out}[k-1] - \Delta A_{out}[k-1] < 0$; hence, $B_d \leq A_{out}[k-1] + \Delta A_{out}[k-1]$. Substituting this inequality and the control equation into the expression for $B[k]$ gives

$$\begin{aligned} B[k] &= A_{out}[k-1] + \Delta A_{out}[k-1] + \mu_1 R[k-1] \\ &= A_{out}[k-1] + \Delta A_{out}[k-1] + 0 \\ &\geq B_d. \end{aligned}$$

Thus, the controller satisfies $B[k] \geq B_d$ for $k \geq 1$. □

D.10 Proof of Theorem 3.7

Proof. The proof to show $B[k] \geq B_d[k-1]$ for $k \geq 1$ is straight-forward. First, suppose $R[k-1] = \frac{1}{\mu_1}\left(B_d[k-1] - A_{out}[k-1] - \Delta A_{out}[k-1]\right)$. Then,

$$\begin{aligned} B[k] &= A_{out}[k-1] + \Delta A_{out}[k-1] + \mu_1 R[k-1] \\ &= A_{out}[k-1] + \Delta A_{out}[k-1] + B_d[k-1] - A_{out}[k-1] - \Delta A_{out}[k-1] \\ &= B_d[k-1]. \end{aligned}$$

Alternatively, suppose $R[k-1] = 0$, which occurs when $B_d[k-1] - A_{out}[k-1] - \Delta A_{out}[k-1] < 0$; hence, $B_d[k-1] \leq A_{out}[k-1] + \Delta A_{out}[k-1]$. Substituting this inequality and the control equation into the expression for $B[k]$ gives

$$\begin{aligned}B[k] &= A_{out}[k-1] + \Delta A_{out}[k-1] + \mu_1 R[k-1] \\ &= A_{out}[k-1] + \Delta A_{out}[k-1] + 0 \\ &\geq B_d[k-1].\end{aligned}$$

Thus, the controller satisfies $B[k] \geq B_d[k-1]$ for $k \geq 1$. □

D.11 Proof of Theorem 3.8

Proof. From Theorem 3.7, $B[k] \geq B_d[k-1]$ for $k \geq 1$. As a result, $B[k]$ is positive for $k \geq 1$ implying (from Lemma 3.3) the following states converge as k tends to infinity: $A_{out}[k]$, $\Delta A^{CD}_{out}[k]$ and $\Delta A^{OJ}_{out}[k]$. It follows directly (2.12) that the convergence of $\Delta A^{CD}_{out}[k]$ implies $P^{CD}[k]$ also converges in steady-state. Before showing $P^{OJ}[k]$ converges as k tends to infinity, $B[k]$ and $A_{rew}[k]$ are considered.

To show $A_{rew}[\cdot]$ is bounded, the expression for the controller is examined. From (3.17), $A_{rew}[k]$ can be expressed as

$$A_{rew}[k] = \begin{cases} 0, & \text{if } B_d[k] - A_{out}[k] \leq 0 \\ B_d[k] - A_{out}[k], & \text{otherwise.} \end{cases}$$

Since $A_{out}[k]$ converges as k tends to infinity, it is bounded. Along with the fact that $B_d[\cdot]$ is bounded, this conclusion implies that the second expression in the above equation is bounded. Moreover, the first expression in the above equation is also bounded. As a result, $A_{rew}[\cdot]$ is bounded. Finally, since $B[k] = A_{out}[k] + A_{rew}[k]$, and both $A_{out}[\cdot]$ and $A_{rew}[\cdot]$ are bounded, it follows that $B[\cdot]$ is also bounded.

For the special case when $B_d[k] = B_d$ for $k \geq 0$, $A_{rew}[k]$ can be expressed as

$$A_{rew}[k] = \begin{cases} 0, & \text{if } B_d - A_{out}[k] \leq 0 \\ B_d - A_{out}[k], & \text{otherwise.} \end{cases}$$

Since $A_{out}[k]$ converges as k tends to infinity, it follows that $B_d - A_{out}[k]$ also converges as k tends to infinity. If $B_d - A_{out}[k]$ converges to some negative value, then $A_{rew}[k]$ converges to zero as k tends to infinity. If, instead, $B_d - A_{out}[k]$ converges to some positive value, then $A_{rew}[k]$ converges to $B_d - A_{out}[k]$ as k tends to infinity. Finally, if $B_d - A_{out}[k]$ converges to zero as k tends to infinity, then both possible expressions for $A_{rew}[k]$ tend to zero as k tends to infinity and, thus, $A_{rew}[k]$ tends to zero as k tends to infinity. As a result, $A_{rew}[k]$ converges as k tends to infinity. Finally, since $B[k] = A_{out}[k] + A_{rew}[k]$, and both $A_{out}[k]$

and $A_{rew}[k]$ converge as k tends to infinity, it follows that $B[k]$ also converges as k tends to infinity.

Now, it only remains to show that $P^{OJ}[k]$ converges as k tends to infinity. Consider first the case when $A_{out}[k] \leq 0$ for $k \geq 0$. It follows directly from Lemma 3.1 that $P^{OJ}[k] = 0$ for $k \geq 0$ and, thus, $P^{OJ}[k]$ converges to zero as k tends to infinity.

Now, suppose $A_{out}[k] > 0$ at $k = \bar{k}$ (but $A_{out}[k] \leq 0$ at $k = \bar{k} - 1$). Note that since $A_{out}[k] < 0$ at $k = \bar{k} - 1$, it follows from (3.17) that $R[\bar{k} - 1] > 0$ and hence $A_{rew}[\bar{k}] > 0$. As a result, $P^{OJ}[\bar{k}] > 0$, implying $P^{OJ}[k] > 0$ for $k \geq \bar{k}$. Therefore, for $k \geq \bar{k}$, only two expressions are possible for $\Delta A_{out}^{OJ}[k]$:

$$\Delta A_{out}^{OJ}[k] = \begin{cases} -K_2 P^{OJ}[k], & \text{if } P^{OJ}[k] > 0 \text{ and } K_2 P^{OJ}[k] \leq A_{out}[k], \\ -A_{out}[k], & \text{if } P^{OJ}[k] > 0 \text{ and } K_2 P^{OJ}[k] > A_{out}[k]. \end{cases}$$

It has previously been argued that these two expressions can combine in three possible ways:

(i) For $k \geq \bar{k}$, $\Delta A_{out}^{OJ}[k] = -K_2 P^{OJ}[k]$;

(ii) For $\bar{k} \leq k < \hat{k}$, $\Delta A_{out}^{OJ}[k] = -K_2 P^{OJ}[k]$, and for $k \geq \hat{k}$, $\Delta A_{out}^{OJ}[k] = -A_{out}[k]$; and

(iii) For $k \geq \bar{k}$, $\Delta A_{out}^{OJ}[k] = -A_{out}[k]$.

First, suppose for $k \geq \bar{k}$, $\Delta A_{out}^{OJ}[k] = -K_2 P^{OJ}[k]$. Then, since, by Lemma 3.3, $\Delta A_{out}^{OJ}[k]$ tends to zero as k tends to infinity, it follows that $P^{OJ}[k]$ necessarily converges to zero as k tends to infinity.

Now, suppose that for $k \geq \hat{k}$, $\Delta A_{out}^{OJ}[k] = -A_{out}[k]$ (which applies in both of the remaining combinations for $\Delta A_{out}^{OJ}[k]$ above). Then, for $k \geq \hat{k}$,

$$A_{out}[k+1] = A_{out}[k] + \Delta A_{out}^{OJ}[k] + \Delta A_{out}^{CD}[k] = \Delta A_{out}^{CD}[k].$$

Since, by Lemma 3.3, $\Delta A_{out}^{CD}[k]$ converges to zero as k tends to infinity, it follows that $A_{out}[k]$ tends to zero as k tends to infinity. Additionally, note that for $k \geq \hat{k}$,

$$P^{OJ}[k] = r_3 P^{OJ}[k-1] + (1 - r_3)\, A_{out}[k] A_{rew}[k].$$

Since the above equation can be re-written as

$$P^{OJ}[k] - r_3 P^{OJ}[k-1] = (1 - r_3)\, A_{out}[k] A_{rew}[k],$$

Because $A_{out}[k]$ converges as k tends to infinity and $A_{rew}[\cdot]$ is bounded, the above equation can be viewed as a LTI system with an input that converges as k tends to infinity. Specifically, by defining $u[k]$ as the right-hand side of the above equation and taking z-transforms gives

$$\frac{P^{OJ}[z]}{u[z]} = \frac{z}{z - r_3}.$$

The above system is stable because $|r_3| < 1$ and, thus, since the input, $u[k]$ converges as k tends to infinity, the output, $P^{OJ}[k]$ also converges as k tends to infinity. □

D.12 Proof of Theorem 3.9

Proof. Define the numerator and denominator of the plant and the controller by

$$P[z] = \frac{N_p[z]}{D_p[z]} := \frac{\mu_1}{z},$$
$$C_{nom}[z] = \frac{N_c[z]}{D_c[z]} := \frac{K_c\,(z-a)}{(z-1)^2}.$$

Then, the characteristic polynomial given by $\Delta[z] = N_c[z]N_p[z] + D_c[z]D_p[z]$ is converted to continuous-time via a bilinear transformation, giving

$$\begin{aligned}
\Delta(s) &= \Delta(z)|_{z=\frac{1+s}{1-s}} \\
&= K_c\,(z-a)\,\mu_1 + z\left(z^2 - 2z + 1\right)|_{z=\frac{1+s}{1-s}} \\
&= z^3 - 2z^2 + z\,(1+\mu_1 K_c) - \mu_1 K_c a|_{z=\frac{1+s}{1-s}} \\
&= \left(\frac{1+s}{1-s}\right)^3 - 2\left(\frac{1+s}{1-s}\right)^2 + \left(\frac{1+s}{1-s}\right)(1+\mu_1 K_c) - \mu_1 K_c a \\
&= \frac{(1+s)^3 - 2\,(1+s)^2\,(1-s) + (1+\mu_1 K_c)\,(1+s)\,(1-s)^2 - \mu_1 K_c a\,(1-s)^3}{(1-s)^3}.
\end{aligned}$$

Since the bilinear transformation maintains stability properties after transformation, the roots of $\Delta[z]$ are in the unit circle if and only if the roots of $\Delta(s)$ are in the OLHP. The numerator of $\Delta(s)$ is

$$s^3\,(4 + \mu_1 K_c\,(1+a)) + s^2\,(4 - \mu_1 K_c\,(1+3a)) - s\,(\mu_1 K_c\,(1-3a)) + \mu_1 K_c\,(1-a)\,.$$

Thus, the Routh Array of the numerator of $\Delta(s)$ is given by

$$\begin{array}{c|ccc}
s^3 & 4 + \mu_1 K_c\,(1+a) & -\mu_1 K_c\,(1-3a) & 0 \\
s^2 & 4 - \mu_1 K_c\,(1+3a) & \mu_1 K_c\,(1-a) & 0 \\
s^1 & A & 0 & 0 \\
s^0 & \mu_1 K_c\,(1-a) & 0 & 0,
\end{array}$$

where

$$A = \frac{-\mu_1 8K_c\left(\mu_1 K_c a^2 - 2a + 1\right)}{4 - \mu_1 K_c\left(1 + 3a\right)}.$$

The Routh-Hurwitz criterion states if there are no sign changes in the first column of the Routh array, then all of the roots of $\Delta(s)$ are in the OLHP.

Now the conditions on K_c and a that guarantee stability of the closed-loop system are found. To simply things, consider only the case when the sign of all terms in the first column are positive. From the last element of the first column, notice that since $\mu_1 > 0$ and $K_c > 0$, $(1 - a) > 0$ must hold, thus restricting the controller zero to $a < 1$.

Now consider the case when $a < 0$. Suppose $a < 0$; then, since the second element of the Routh Array must remain positive,

$$K_c < \frac{4}{1 + 3a}$$

must hold. However, $K_c > 0$ must also hold and therefore, the right-hand side of the above inequality must necessarily be positive. To ensure the right-hand side is positive, a must be greater than $\frac{-1}{3}$. Furthermore, from the first element of the Routh Array,

$$K_c > \frac{-4}{\left(1 + a\right)\mu_1}.$$

Since $\mu_1 K_c > 0$, it follows that if $\frac{-1}{3} < a < 1$, then the right-hand side of the above inequality is negative and therefore, this element does not provide additional information on the range of permissible values for a and K_c.

Now, only one element in the Routh Array remains:

$$\frac{-\mu_1 8K_c\left(\mu_1 K_c a^2 - 2a + 1\right)}{4 - \mu_1 K_c\left(1 + 3a\right)} > 0.$$

To ensure the above inequality holds, the numerator and the denominator must have the same sign. Notice that the denominator is the same as the second element of the Routh Array and has previously been shown to be positive; thus the numerator must also be positive. Since $\mu_1 > 0$ and $K_c > 0$, it only remains to find conditions on K_c and a to guarantee

$$\mu_1 K_c a^2 - 2a + 1 < 0.$$

Rearranging gives,

$$K_c < \frac{2a - 1}{\mu_1 a^2}.$$

Since $K_c > 0$, the right-hand side of the above inequality must necessarily be positive. This further restricts the range of a to $0.5 < a < 1$.

In summary, from the fourth element of the Routh Array, an upper bound on a of 1 is obtained. From the first element of the Routh Array, an upper bound is obtained for K_c and an initial lower bound of $\frac{-1}{3}$ is obtained for a. Finally, the third element of the Routh Array further bounds a to yield the final permissible range of a values to $0.5 < a < 1$. Moreover, this third element provides a second upper bound on K_c. Combining these two upper bounds with the restriction that $K_c > 0$ gives

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

as required. □

D.13 Proof of Lemma 3.5

Proof. The proof to show that $B[k]$ is an increasing function of k for at least the interval $T^* \leq k \leq \hat{k}$, eventually becoming non-negative at $k = \hat{k}$, is done in three parts. The first part shows that there exists a T^*, i.e., *eventually* $B[k+1] > B[k]$. The second part shows that $B[k]$ is an increasing function of k for at least the interval $T^* \leq k \leq \hat{k}$. The final part shows that $B[k]$ must necessarily become positive.

Part 1

First, another expression for the reward is derived for the case when $B[k] \leq 0$ for $0 \leq k < \hat{k}$, i.e., when the system is in stage I. This expression is then shown to imply that T^* exists. Thus, suppose $B[k] \leq 0$ for $0 \leq k < \hat{k}$. Then, over this sample range,

$$R[k] = K_c e[k-1] - K_c a e[k-2] + 2R[k-1] - R[k-2] = K_c \sum_{i=0}^{k-1} e[i]\left(\left(k-i\right)\left(1-a\right)+a\right).$$

To show that the above expression can be used as the controller expression for $0 \leq k < \hat{k}$, induction is used. Since $B[0] = A_o < 0$, it follows that $e[0] := B_d - B[0] > 0$ and, therefore,

$$K_c e[0] + 2R[0] > 0,$$

i.e., $R[1] = \max\left\{0, K_c e[0] + 2R[0]\right\} > 0$. In other words at $k = 1$, $R[k]$ is not saturated. More generally, suppose, for $1 \leq k < \tilde{k}$, $B[k] \leq 0$ and $R[k] > 0$. Then, for the given range of k, the controller is not saturated. As a result, $R[k] = K_c e[k-1] - K_c a e[k-2] + 2R[k-1] - R[k-2]$, which can alternatively be expressed as

$$R[k] = K_c \sum_{i=0}^{k-1} e[i]\left(\left(k-i\right)\left(1-a\right)+a\right). \quad \text{(D.12)}$$

Since $e[k] = B_d - B[k]$, and $B[k] \leq 0$ for $1 \leq k < \tilde{k}$, $e[k] > 0$ for $1 \leq k < \tilde{k}$ and thus, the reward is guaranteed to be positive for the given range of k. Since $R[\tilde{k}]$ only depends on values of k in the range $0 \leq k < \tilde{k}$, it follows that the expression given in (D.12) can be used for the second of the two possible expressions for $R[\tilde{k}]$, i.e.,

$$R[\tilde{k}] = \max\left\{0, K_c \sum_{i=0}^{\tilde{k}-1} e[i] \left(\left(\tilde{k} - i\right)(1-a) + a\right)\right\}.$$

It only remains to show that the second expression is larger than the first. Since for $0 \leq k < \tilde{k}$, $B[k] \leq 0$ and given that $e[k]$ is defined as the difference $B_d - B[k]$ for all $k \geq 0$, it follows that for $0 \leq k < \tilde{k}$, $e[k] > 0$. Due to the summation range, for the second expression to be positive, $e[k] > 0$ must hold for $0 \leq k < \tilde{k}$, which is exactly the case. Note that the range of permissible values for a implies $1 - a > 0$ and $a > 0$. Moreover, since i can be at most $\tilde{k} - 1$, it follows that $\tilde{k} - i > 0$ also and, therefore, all terms in the summation are positive. In other words, $R[\tilde{k}] > 0$ and, therefore, when $B[k] \leq 0$, (D.12) can be used as the controller expression.

This controller expression is now used to show T^* exists. We first show that $R[k]$ is an increasing function of k as long as $B[k] \leq 0$; note that

$$R[k] - R[k-1] = K_c \left(e[k-1] + \sum_{i=0}^{k-2} e[i](1-a)\right) \quad \text{(D.13)}$$

on the interval $0 \leq k < \hat{k}$ (since $B[k] \leq 0$ for $0 \leq k < \hat{k}$). Moreover, $e[k] > 0$ over this interval. Therefore $R[k] - R[k-1] > 0$ for $0 \leq k < \hat{k}$, i.e., $R[k]$ is an increasing function of k over this interval.

Suppose k is in the range $0 \leq k < \hat{k}$, i.e., $B[k] \leq 0$ and the system is in stage I. Then, no overjustification pressure arises and therefore, from (2.2), (2.4), (2.5), (2.12) and (D.12), $B[k]$ can be expressed, for $0 \leq k < \hat{k}$, as

$$B[k] = A_o + \sum_{i=0}^{k-1} K_1 P^{CD}[i] + \mu_1 K_c \sum_{i=0}^{k-2} e[i]\left((k-1-i)(1-a) + a\right). \quad \text{(D.14)}$$

To show T^* exists, we look at the difference $B[k+1] - B[k]$, which, for $0 \leq k < \hat{k}$, is given by

$$B[k+1] - B[k] = K_1 P^{CD}[k] + \mu_1 K_c \left(e[k-2] + \sum_{i=0}^{k-3} e[i](1-a)\right). \quad \text{(D.15)}$$

Suppose T^* does not exist. Then, $B[k+1] - B[k] \leq 0$ for all $k \geq 0$, which implies that $e[k+1] - e[k] \geq 0$ for all $k \geq 0$ (because $e[k] = B_d - B[k]$ for all $k \geq 0$). Therefore, $e[k]$ is an non-decreasing function of k for $k \geq 0$ and, thus, either $e[k]$ tends to positive infinity as k tends to infinity, or $e[k]$ converges as k tends to infinity. We show now that $e[k]$ does not converge as k tends to infinity by showing $B[k]$ does not converge as k tends to infinity (by contradiction). Thus, suppose $B[k]$ is a non-increasing function of k for $k \geq 0$, converging as k tends to infinity. Since $B[k] \leq 0$ for $k \geq 0$, it follows that $P^{CD}[k] < 0$ for $k \geq 1$. Therefore, $A_{out}[k] < A_{out}[k-1]$ for $k \geq 1$. Moreover, since $B[k]$ is a non-increasing function of k for $k \geq 1$, the reward is necessarily increasing at a slower rate than the internal attitude is decreasing, i.e.,

$$\frac{\mu_1 R[k-1]}{|A_{out}[k]|} < \frac{\mu_1 R[k-2]}{|A_{out}[k-1]|}.$$

Since $P^{CD}[k] = -\frac{1}{1+\frac{\mu_1 R[k-1]}{|A_{out}[k]|}}$ for $k \geq 0$, $P^{CD}[k] < P^{CD}[k-1]$. Therefore, $P^{CD}[k]$ is a decreasing function of k for $k \geq 0$, which, by Lemma 2.1, is lower bounded by -0.5. From Lemma A.1, this implies that $P^{CD}[k]$ converges to some non-zero, negative value as k tends to infinity. Therefore, $A_{out}[k]$ does not converge as k tends to infinity, since

$$A_{out}[k] = A_o + \sum_{i=0}^{k-1} K_1 P^{CD}[i].$$

Since $B[k] = A_{out}[k] + \mu_1 R[k-1]$, $A_{out}[k]$ does not converge as k tends to infinity and $R[k]$ is an increasing function of k for $k \geq 0$, $B[k]$ does not converge as k tends to infinity, a result that contradicts the initial supposition that $B[k]$ converges as k tends to infinity. Therefore, $B[k]$ does not converge as k tends to infinity; consequently, $e[k]$ does not converge as k tends to infinity.

Since $e[k]$ does not converge as k tends to infinity and $e[k]$ is an non-decreasing function of k for $k \geq 0$ (from the supposition that T^* does not exist) there exists a $\check{k}$ such that $\mu_1 K_c e[\check{k}-2] > \frac{K_1}{2}$. Substituting this relationship into (D.15) gives

$$B[k+1] - B[k] > K_1 P^{CD}[k] + \frac{K_1}{2} + \mu_1 K_c \left(\sum_{i=0}^{k-3} e[i] \, (1-a) \right).$$

By Lemma 2.1, $P^{CD}[k] > \frac{-1}{2}$; thus,

$$B[k+1] - B[k] > \mu_1 K_c \left(\sum_{i=0}^{k-3} e[i] \, (1-a) \right).$$

Moreover, since $e[k] \geq 0$ for $0 \leq k < \hat{k}$ and $1 - a > 0$, it follows that the right-hand side of the above inequality is positive, thus $B[k+1] - B[k] > 0$, which contradicts the supposition that T^* does not exist. Therefore, T^* exists.

Part 2

Now, to show that $B[T^*+1] > B[T^*]$ implies $B[k]$ remains an increasing function for at least the interval $T^* \leq k \leq \hat{k}$, we show that for any $\grave{k}$ in the interval $T^* \leq \grave{k} \leq \hat{k}$, $B[\grave{k}+1] - B[\grave{k}] > 0$.

Note from (2.2), (2.4) and (2.12),

$$B[\grave{k}+1] - B[\grave{k}] = A_{out}[\grave{k}] + K_1 P^{CD}[\grave{k}] - A_{out}[\grave{k}] + \mu_1 \left(R[\grave{k}] - R[\grave{k}-1] \right).$$

Since $B[\grave{k}] \leq 0$ and $B[\grave{k}-1] \leq 0$, (D.12) can be used for $R[\grave{k}]$ and $R[\grave{k}-1]$; therefore, (D.13) can be used for the difference $R[\grave{k}] - R[\grave{k}-1]$, giving

$$B[\grave{k}+1] - B[\grave{k}] = K_1 P^{CD}[\grave{k}] + \mu_1 K_c \left(e[\grave{k}-1] + \sum_{i=0}^{\grave{k}-2} e[i]\,(1-a) \right).$$

By Assumption 3.1 and Lemma 2.1, $|P^{CD}[\grave{k}]| < 0.5$. Since $B[\grave{k}] \leq 0$, it follows that $P^{CD}[\grave{k}] < 0$ and therefore, $P^{CD}[\grave{k}] > \frac{-1}{2}$. Furthermore, $e[\grave{k}-1] = B_d - B[\grave{k}-1]$, and thus, substituting (3.22) into the expression for $B[\grave{k}+1] - B[\grave{k}]$ gives

$$B[\grave{k}+1] - B[\grave{k}] > \frac{-K_1}{2} + \mu_1 K_c \left(\frac{K_1}{2\mu_1 K_c} \right) - \mu_1 K_c B[\grave{k}-1] + \mu_1 K_c \sum_{i=0}^{\grave{k}-2} e[i]\,(1-a).$$

Simplifying the above expression yields

$$\begin{aligned} B[\grave{k}+1] - B[\grave{k}] &> \frac{-K_1}{2} + \frac{K_1}{2} - \mu_1 K_c B[\grave{k}-1] + \mu_1 K_c \sum_{i=0}^{\grave{k}-2} e[i]\,(1-a) \\ &= -\mu_1 K_c B[\grave{k}-1] + \mu_1 K_c \sum_{i=0}^{\grave{k}-2} e[i]\,(1-a). \end{aligned}$$

Since $B[\grave{k}-1] \leq 0$, and, for the summation range, $e[\grave{k}] > 0$, the right-hand side of the above expression is positive, i.e., $B[\grave{k}+1] - B[\grave{k}] > 0$. Therefore, for some arbitrary $T^* \leq \grave{k} \leq \hat{k}$, if $B[\grave{k}] > B[\grave{k}-1]$ then, provided $B_d > \frac{K_1}{2\mu_1 K_c}$, $B[\grave{k}+1] > B[\grave{k}]$; thus, $B[k]$ remains an increasing function of k, at least for $T^* \leq k \leq \hat{k}$.

Part 3

Since $B[k]$ is an increasing function it will either go positive or, by Lemma A.2, it will converge to some non-positive constant. Suppose the latter case is true, i.e., $B[k]$ tends to $c \leq 0$ as k tends to infinity; then the error, $e[k]$, converges to $B_d - c > 0$ as k tends to infinity. Since (D.12) holds over the interval $0 \le k < \hat{k}$, it follows that $B[k]$ converging to c as k tends to infinity implies $\hat{k} = \infty$. Thus, for $k \ge 0$,

$$R[k] = K_c e[k-1] - K_c a e[k-2] + 2R[k-1] - R[k-2].$$

Notice that the expression for $B[k+1]$ given in (2.4) can be rearranged as $\mu_1 R[k] = B[k+1] - A_{out}[k+1]$; hence, the control signal can be written as

$$B[k+1] - A_{out}[k+1] = \mu_1 K_c e[k-1] - \mu_1 K_c a e[k-2] + 2B[k] - 2A_{out}[k] - B[k-1] + A_{out}[k-1].$$

Using the fact that $A_{out}[k] = A_{out}[k-1] + K_1 P^{CD}[k-1]$, the above expression can be simplified and rearranged to

$$B[k+1] - 2B[k] + B[k-1] = \mu_1 K_c e[k-1] - \mu_1 K_c a e[k-2] + K_1 P[k] - K_1 P[k-1].$$

From the assumption that $B[k]$ tends to $c \le 0$ as k tends to infinity, the left-hand side of the above expression is well-behaved. Therefore, the right-hand side of the above expression is also well-behaved, meaning limits can be taken on both sides, giving

$$c - 2c + c = \lim_{k\to\infty} \left(\mu_1 K_c e[k-1] - \mu_1 K_c a e[k-2] + K_1 P[k] - K_1 P[k-1] \right). \qquad \text{(D.16)}$$

Since $B[k]$ converges as k tends to infinity, $e[k]$ also converges as k tends to infinity.

Note though, that since $B[k]$ is an increasing function of k, the reward is necessarily increasing at a faster rate than the internal attitude is decreasing and therefore,

$$\frac{\mu_1 R[k-1]}{|A_{out}[k]|} > \frac{\mu_1 R[k-2]}{|A_{out}[k-1]|}.$$

From Assumption 3.1, $P^{CD}[k] = P^{CD}_{raw}[k]$ and therefore, $P^{CD}[k] = -\frac{1}{1+\frac{\mu_1 R[k-1]}{|A_{out}[k]|}}$. In other words, the above inequality implies $P^{CD}[k] > P^{CD}[k-1]$, i.e., dissonance pressure is an increasing function, which is bounded from above by 0 (since $B[k] \le 0$ for $k \ge 0$). Therefore, by Lemma A.2, $P^{CD}[k]$ converges as k tends to infinity. Therefore, (D.16) becomes

$$0 = \lim_{k\to\infty} \mu_1 K_c e[k-1] (1-a),$$

implying $e[k]$ converges to zero as k tends to infinity. However, since $e[k] := B_d - B[k]$, it must be true that $B_d - B[k]$ tends to zero as k tends to infinity. Given that $\lim_{k\to\infty} B[k] = c$, it follows that

$$B_d = c.$$

The above conclusion contradicts the restriction that $c \le 0$, because $B_d > 0$; thus, the assumption that $B[k]$ converges to a non-positive number as k tends to infinity is incorrect. Therefore, $B[k]$ must become positive. $\square$

D.14 Proof of Lemma 3.6

Proof. By Assumption 3.2, $B[k] \geq 0$ for $k \geq \hat{k}$; therefore, by Assumption 3.1, $P^{CD}[k] \geq 0$ for $k \geq \hat{k}$. As a result, $A_{out}[k]$ is an increasing function over this sample range due to the positive dissonance pressure. Since $A_{out}[k] < 0$ and is an increasing function of k, by Lemma A.2, it converges to some non-positive constant as k tends to infinity.

Notice that for $k \geq \hat{k}$, $A_{out}[k]$ can be written as

$$A_{out}[k] = A_{out}[\hat{k}] + \sum_{i=\hat{k}}^{k-1} K_1 P^{CD}[i].$$

Since $A_{out}[k]$ converges as k tends to infinity, the summation term converges to zero as k tends to infinity and therefore, $P^{CD}[k]$ necessarily converges to zero as k tends to infinity. This conclusion has implications on the control signal. In particular, given that $B[k] \geq 0$ for $k \geq \hat{k}$ and $A_{out}[k] < 0$ over this sample range, $P^{CD}[k]$, by Assumption 3.1, is given by

$$P^{CD}[k] = \frac{|A_{out}[k]|}{|A_{out}[k]| + \mu_1 R[k-1]} = \frac{1}{1 + \frac{\mu_1 R[k-1]}{|A_{out}[k]|}}.$$

Taking limits of both sides of the above expression implies that, as $k \to \infty$, $R[k] \to \infty$, $|A_{out}[k]| \to 0$, or both.

Suppose $R[k]$ tends to infinity as k tends to infinity; then, since $B[k] = A_{out}[k] + \mu_1 R[k-1]$ and $A_{out}[k]$ converges to some non-positive constant, $B[k]$ must also tend to infinity as k tends to infinity. Therefore, $e[k]$ must tend to negative infinity as k tends to infinity. Moreover, if $R[k]$ tends to infinity as k tends to infinity, then as k gets sufficiently large, the reward is in the form given by (D.12); thus, for large k,

$$R[k] - R[k-1] = K_c \left(e[k-1] + \sum_{i=0}^{k-2} e[i]\,(1-a) \right).$$

From the conclusion that the error tends to negative infinity as k tends to infinity, it follows that $R[k] - R[k-1]$ also tends to negative infinity as k tends to infinity. Therefore $R[k]$ must be decreasing as k tends to infinity. However, this violates the supposition that $R[k]$ tends to infinity as k tends to infinity and thus, the supposition is incorrect. As a result, $A_{out}[k]$ necessarily tends to zero as k tends to infinity. □

D.15 Proof of Lemma 3.7

Proof. The lemma is proven by contradiction. Suppose, for some $\epsilon > 0$, there exists a $\tilde{k}$ such that for all $k \geq \tilde{k}$, $B[k] \leq B_d - \epsilon$; thus, for all $k \geq \tilde{k}$

$$e[k] = B_d - B[k] \geq B_d - B_d + \epsilon = \epsilon.$$

Now, to aid the analysis, consider the expanded controller given in Figure 3.13. In particular, the characteristics of the following signals are discussed for $k \geq \tilde{k}$: $e_1[k]$, $e_2[k]$ and $R_{us}[k]$.

Since $e[k] \geq \epsilon$ for $k \geq \tilde{k}$ and $e_1[k] = e[k-1] + e_1[k-1]$, it follows that $e_1[k] \geq \epsilon + e_1[k-1]$. In other words, $e_1[k]$ is the sum of its previous sample, $e_1[k-1]$, and some positive number, ϵ; thus, $e_1[k]$ is an increasing function with the property that it increases by at least ϵ at each sample. As a result, $e_1[k]$ is unbounded and, hence, there exists a $\acute{k}$ such that for all $k \geq \acute{k}$, $e_1[k] \geq \epsilon$. Repeating this argument for $e_2[k]$ we conclude that there exists a $\grave{k}$ such that for all $k \geq \grave{k}$, $e_2[k] \geq \epsilon$ and $e_2[k+1] \geq e_2[k] + \epsilon$.

Now, taking the inverse z-transform of the relationship between $R_{us}[z]$ and $e_2[z]$ yields

$$R_{us}[k] = K_c e_2[k+1] - K_c a e_2[k].$$

Use $e_2[k+1] > e_2[k]$, for $k \geq \grave{k}$ to deduce

$$R_{us}[k] > K_c \left(e_2[k] - a e_2[k]\right) = K_c e_2[k] \left(1 - a\right). \tag{D.17}$$

Since $K_c > 0$, $1 - a > 0$ and $e_2[k] \geq \epsilon$, it follows that $R_{us}[k] > \epsilon$ and therefore, $R[k]$ is unsaturated for $k \geq \grave{k}$.

Because $e_2[k]$ is increasing by at least ϵ at each time sample, there exists a sample $\bar{N} \geq \grave{k}$ such that

$$e_2[\bar{N}] > \frac{B_d - \epsilon - A_{out}[\bar{N}+1]}{\left(1-a\right)\mu_1 K_c}.$$

Thus, from (D.17),

$$R_{us}[\bar{N}] > K_c \frac{B_d - \epsilon - A_{out}[\bar{N}+1]}{\left(1-a\right)\mu_1 K_c}\left(1-a\right) = \frac{B_d - \epsilon - A_{out}[\bar{N}+1]}{\mu_1}.$$

From the fact that $R[k]$ is unsaturated for $k \geq \grave{k}$, it follows that $R[k] = R_{us}[k]$. Using the above inequality for $R_{us}[k]$ in the expression for $B[\bar{N}+1]$ gives

$$\begin{aligned} B[\bar{N}+1] &= A_{out}[\bar{N}+1] + \mu_1 R_{us}[\bar{N}] \\ &> A_{out}[\bar{N}+1] + \mu_1 \frac{B_d - \epsilon - A_{out}[\bar{N}+1]}{\mu_1} \\ &= A_{out}[\bar{N}+1] + B_d - \epsilon - A_{out}[\bar{N}+1] \\ &= B_d - \epsilon. \end{aligned}$$

Therefore, $B[\bar{N}+1] > B_d - \epsilon$, which is a contradiction of the original supposition. Therefore, there does not exists a $\tilde{k}$ such that for all $k \geq \tilde{k}$, $B[k] \leq B_d - \epsilon$. □

D.16 Proof of Lemma 3.8

Proof. First, $A_{out}[k]$ is shown to converge to some non-negative constant as k tends to infinity. From the system equations, for $k > \bar{k}$,

$$A_{out}[k] = A_{out}[\bar{k}] + \sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{OJ}[i].$$

It follows from the construction of $\Delta A_{out}^{OJ}[k]$ that $A_{out}[k]$ is a non-increasing function lower bounded by zero. Therefore, it must converge to come non-negative constant, c, i.e., $\lim_{k\to\infty} A_{out}[k] = c \geq 0$. Since $A_{out}[k] < \frac{K_1}{2}$, it follows that $c < \frac{K_1}{2}$ and since $B_d > \frac{K_1}{2}$, it follows that $c < B_d$.

Now, contradiction is used to show that $R[k]$ cannot converge to zero as k tends to infinity; thus, suppose $\lim_{k\to\infty} R[k] = 0$. Since $A_{out}[k]$ converges to c as k tends to infinity, the right-hand side of the expression for $B[k]$, given by $B[k] = A_{out}[k] + A_{rew}[k]$, is well-behaved as k tends to infinity. As a result, limits can be taken on both sides of the above expression as follows:

$$\lim_{k\to\infty} B[k] = \lim_{k\to\infty} \left(A_{out}[k] + \mu_1 R[k-1]\right).$$

From the assumption that $R[k]$ tends to zero as k tends to infinity, it follows that $B[k]$ tends to c as k tends to infinity. But $c < B_d$, so $B[k]$ settles to some value c that is less than B_d, which contradicts Lemma 3.7. Therefore, the supposition that $R[k]$ converges to zero as k tends to infinity is incorrect. □

D.17 Proof of Lemma 3.9

Proof. For $k \geq \bar{k}$, $A_{out}[k] \geq 0$, $A_{rew}[k] \geq 0$ and $B[k] \geq 0$ and, therefore, no raw, unprocessed dissonance pressure arises for $k \geq \bar{k}$. By Assumption 3.1, $P^{CD}[k] = 0$ for $k \geq \bar{k}$, implying $\Delta A_{out}^{CD}[k] = 0$ for $k \geq \bar{k}$. Therefore, for $k > \bar{k}$,

$$A_{out}[k] = A_{out}[\bar{k}] + \sum_{i=\bar{k}}^{k-1} \Delta A_{out}^{OJ}[i].$$

Since $A_{out}[k] \geq 0$ for $k \geq \bar{k}$ and $\Delta A_{out}^{OJ}[k] \leq 0$ for $k \geq \bar{k}$, it follows that $A_{out}[k]$ is a non-increasing function of k for $k \geq \bar{k}$, lower bounded by zero. Therefore, by Lemma A.1, $A_{out}[k]$ converges as k tends to infinity, which implies that $\Delta A_{out}^{OJ}[k]$ tends to zero as k tends to infinity.

To determine the effect this conclusion has on the steady-state value for $A_{out}[k]$, consider the three possible expressions for $\Delta A_{out}^{OJ}[k]$, given by (2.15). Since from Lemma 3.8, $R[k]$ cannot converge to zero as k tends to infinity, it follows that at some $\dot{k} \geq \bar{k}$, $R[\dot{k}] > 0$; therefore, $P^{OJ}[\dot{k}] > 0$ and $P^{OJ}[k] > 0$ for $k \geq \dot{k}$. Therefore, for $k \geq \dot{k}$, there are three possible expressions for $\Delta A_{out}^{OJ}[k]$, each depending on the magnitude of $P^{OJ}[k]$:

1. $\Delta A_{out}^{OJ}[k] = -K_2 P^{OJ}[k]$ for $k \geq \dot{k}$;
2. $\Delta A_{out}^{OJ}[k] = -K_2 P^{OJ}[k]$ for $\dot{k} \leq k < \tilde{k}$ and $\Delta A_{out}^{OJ}[k] = -A_{out}[k]$ for $k \geq \tilde{k}$; and
3. $\Delta A_{out}^{OJ}[k] = -A_{out}[k]$ for $k \geq \dot{k}$.

From the above three possible sequences for $\Delta A_{out}^{OJ}[k]$, if $\Delta A_{out}^{OJ}[k]$ tends to zero as k tends to infinity, either $P^{OJ}[k]$ tends to zero as k tends to infinity or $A_{out}[k]$ tends to zero as k infinity (or both). Since the objective is to show $A_{out}[k]$ tends to zero as k tends to infinity, is suffices to show that if $P^{OJ}[k]$ tends to zero as k tends to infinity, then $A_{out}[k] \to 0$ as $k \to \infty$. Thus, suppose $P^{OJ}[k] \to \infty$ as $k \to \infty$; then

$$\begin{aligned}
&\lim_{k\to\infty} P^{OJ}[k] = r_3 \lim_{k\to\infty} P^{OJ}[k-1] + (1-r_3) \lim_{k\to\infty} A_{out}[k] A_{rew}[k] \\
\Rightarrow &0 = 0 + (1-r_3) \lim_{k\to\infty} A_{out}[k] A_{rew}[k].
\end{aligned}$$

The above implies that either $\lim_{k\to\infty} A_{out}[k] = 0$ or $\lim_{k\to\infty} A_{rew}[k] = 0$. However, from Lemma 3.8, $A_{rew}[k]$ does not converge to zero as k tends to infinity and $A_{out}[k]$ necessarily converges to zero as k tends to infinity. □

D.18 Proof of Lemma 3.10

Proof. Similar to Lemma 3.7, this proof uses contradiction. Suppose there exists a $\tilde{k}$ such that for all $k \geq \tilde{k}$, $B[k] \geq B_d + \epsilon$. Thus, for all $k \geq \tilde{k}$

$$e[k] = B_d - B[k] \leq B_d - B_d - \epsilon = -\epsilon.$$

Now, to aid the analysis, the controller is again expanded as shown in Figure 3.13 and the following intermediate signals are studied for $k \geq \tilde{k}$: $e_1[k]$, $e_2[k]$ and $R_{us}[k]$.

Since $e[k] \leq -\epsilon$ for $k \geq \tilde{k}$ and $e_1[k] = e[k-1]+e_1[k-1]$, it follows that $e_1[k] \leq e_1[k-1]-\epsilon$ and thus, $e_1[k]$ is a decreasing function with the additionally property that it decreases by at least $-\epsilon$ at each step. As a result, $e_1[k]$ is an unbounded decreasing function and hence, there exists a $\acute{k}$ such that for all $k \geq \acute{k}$, $e_1[k] \leq -\epsilon$. Repeating this argument for $e_2[k]$, there exists a $\grave{k}$ such that for all $k \geq \grave{k}$, $e_2[k] \leq -\epsilon$ and $e_2[k+1] \leq e_2[k] - \epsilon$.

Now, taking the inverse z-transform of the relationship between $R_{us}[z]$ and $e_2[z]$ yields

$$R_{us}[k] = K_c e_2[k+1] - K_c a e_2[k].$$

Use $e_2[k+1] < e_2[k]$ for $k \geq \grave{k}$ to deduce

$$R_{us}[k] < K_c \left(e_2[k] - a e_2[k] \right) = K_c e_2[k] \left(1 - a \right).$$

Since $K_c > 0$, $1 - a > 0$ and $e_2[k] \leq -\epsilon$, it follows that $R_{us}[k] < 0$ and hence, $R[k]$ is saturated for $k \geq \grave{k}$. However, from Lemma 3.8, $R[k]$ cannot converge to zero as k tends to infinity and therefore, $R[k]$ cannot be saturated for all $k \geq \grave{k}$. Thus, $B[k]$ cannot remain above $B_d + \epsilon$ as k tends to infinity. □

Appendix E

Detailed Proofs for Chapter 5

E.1 Proof of Lemma 5.1

Proof. Since $A_{rew,i}[k] = 0$ for $k \geq 0$ and $i = 1, 2$, (4.25) implies $\Delta A^{OJ}_{out,i}[k] = 0$ for $k \geq 0$ and $i = 1, 2$. To show dissonance pressure does not occur, induction is used. First, consider $B_1[1]$, which, from the initial conditions and (4.24) and (4.28), is given by $B_1[1] = A_{o1} + \mu_{32} A_{o2}$. From the initial conditions, $A_{o1} < 0$ and $A_{o2} < 0$, and therefore, all cognitions are negative; thus $P^{CD}_1[1] = 0$, implying $\Delta A^{CD}_{out,1}[1] = 0$. Similarly, consider $B_2[1]$, which is given by $B_2[1] = A_{o2} + \mu_{31} A_{o1}$. From the initial conditions, $A_{o1} < 0$ and $A_{o2} < 0$, and therefore, all cognitions are negative, thus $P^{CD}_2[1] = 0$, implying $\Delta A^{CD}_{out,2}[1] = 0$. Moreover, $A_{out,i}[2] = A_{out,i}[1] < 0$ for $i = 1, 2$.

More generally, suppose for $1 \leq k < \bar{k}$, $A_{out,1}[k] < 0$ and $A_{out,2}[k] < 0$; then, $B_1[k] < 0$ and $B_2[k] < 0$, implying $P^{CD}_1[k] = 0$, $P^{CD}_2[k] = 0$, $\Delta A^{CD}_{out,1}[k] = 0$ and $\Delta A^{CD}_{out,2}[k] = 0$ for $1 \leq k < \bar{k}$. From (4.1), (4.24) and (4.28), $B_1[\bar{k}]$ is given by

$$B_1[\bar{k}] = A_{out,1}[\bar{k} - 1] + \Delta A_{out,1}[\bar{k} - 1] + \mu_{32} B_2[\bar{k} - 1].$$

Since $\Delta A^{OJ}_{out,i}[k] = 0$ for $k \geq 0$, it follows that $\Delta A_{out,1}[\bar{k} - 1] = \Delta A^{CD}_{out,1}[\bar{k} - 1]$. By the inductive hypothesis, $A_{out,1}[\bar{k} - 1] < 0$, $\Delta A^{CD}_{out,1}[\bar{k} - 1] = 0$ and $B_2[\bar{k} - 1] < 0$, implying $P^{CD}_1[\bar{k}] = 0$ and thus, $\Delta A^{CD}_{out,1}[\bar{k}] = 0$. Similarly, $\Delta A^{CD}_{out,2}[\bar{k}] = 0$. Therefore, if the system begins in region V, then $\Delta A^{CD}_{out,i}[k] = 0$ for $k \geq 0$ and $i = 1, 2$. □

E.2 Proof of Lemma 5.2

Proof. For the zero-input system, $A_{rew,i}[k] = 0$ for all $k \geq 0$ and $i = 1, 2$. Additionally, since for $i = 1, 2$ and all $k \geq 0$, $\Delta A^{CD}_{out,i}[k] = 0$ and $\Delta A^{OJ}_{out,i}[k] = 0$, from (4.1), (4.24) and

(4.28), the expressions for $B_1[k]$ and $B_2[k]$ are given by

$$B_1[k] = A_{o1} + \mu_{32} B_2[k-1] \text{ and } B_2[k] = A_{o2} + \mu_{31} B_1[k-1],$$

for $k \geq 1$. Iterating through the system equations yields

$$B_1[k] = A_{o1} \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} (\mu_{31}\mu_{32})^j + \mu_{32} A_{o2} \sum_{j=0}^{\lfloor \frac{k-1}{2} \rfloor} (\mu_{31}\mu_{32})^j$$
$$B_2[k] = A_{o2} \sum_{j=0}^{\lfloor \frac{k}{2} \rfloor} (\mu_{31}\mu_{32})^j + \mu_{31} A_{o1} \sum_{j=0}^{\lfloor \frac{k-1}{2} \rfloor} (\mu_{31}\mu_{32})^j .$$

The summation terms are geometric series and therefore, for $k \geq 1$, the above equations can be given by

$$B_1[k] = A_{o1} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{k}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right) + \mu_{32} A_{o2} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{k-1}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right)$$
$$B_2[k] = \mu_{31} A_{o1} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{k-1}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right) + A_{o2} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{k}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right),$$

as required. □

E.3 Proof of Lemma 5.3

Proof. From Lemma 5.1, $\Delta A_{out,i}^{CD}[k] = 0$ and $\Delta A_{out,i}^{OJ}[k] = 0$ for $k \geq 0$ and $i = 1, 2$ and therefore, Lemma 5.2 applies; thus, the expression for each behaviour is given by (5.2) and (5.3). From these expressions, Assumption 4.7, which says that $|\mu_{31}| < 1$ and $|\mu_{32}| < 1$, and the fact that $A_{o1} < 0$ and $A_{o2} < 0$, it follows that $B_i[k] < 0$ for $i = 1, 2$ and all $k \geq 0$, converging as k tends to infinity.

To show each behaviour is a decreasing function of k, consider the behaviour expressions given by (5.2) and (5.3). First examine (5.2) at two consecutive samples:

$$B_1[\bar{k}] = A_{o1} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{\bar{k}}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right) + \mu_{32} A_{o2} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{\bar{k}-1}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right)$$
$$B_1[\bar{k}+1] = A_{o1} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{\bar{k}+1}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right) + \mu_{32} A_{o2} \left(\frac{1 - (\mu_{31}\mu_{32})^{1+\lfloor \frac{\bar{k}}{2} \rfloor}}{1 - (\mu_{31}\mu_{32})} \right).$$

By Assumption 4.7, $|\mu_{31}\mu_{32}| < 1$; hence each numerator term is positive. Furthermore, note that

$$1 - (\mu_{31}\mu_{32})^{1+\lfloor\frac{\bar{k}}{2}\rfloor} < 1 - (\mu_{31}\mu_{32})^{1+\lfloor\frac{\bar{k}+1}{2}\rfloor} \text{ and}$$
$$1 - (\mu_{31}\mu_{32})^{1+\lfloor\frac{\bar{k}-1}{2}\rfloor} < 1 - (\mu_{31}\mu_{32})^{1+\lfloor\frac{\bar{k}}{2}\rfloor}.$$

Combining this result with the fact that $A_{o1} < 0$ and $A_{o2} < 0$ leads to the conclusion that $B_1[\bar{k}+1] < B_1[\bar{k}]$, i.e., $B_1[k]$ is a decreasing function of k for $k \geq 0$ (since $\bar{k}$ can be any $k \geq 0$). A similar analysis can be performed on (5.3), and hence, $B_2[k]$ is also a decreasing function of k for $k \geq 0$. □

E.4 Proof of Lemma 5.4

Proof. To show that $B_1[k] < 0$ and $B_2[k] < 0$ for $k \geq 1$, induction is used. First, both behaviours are shown to be negative for $k = 1, 2, 3$. From the initial conditions, (4.24) and (4.28), the two behaviours are given by

$$B_1[1] = A_{o1} + \mu_{32}A_{o2} \text{ and } B_2[1] = A_{o2} + \mu_{31}A_{o1}.$$

From (5.4), $A_{o2} < \mu_{31}|A_{o1}|$ and $A_{o1} < 0$. Thus,

$$B_1[1] < A_{o1} + \mu_{32}\mu_{31}|A_{o1}| = A_{o1}(1 - \mu_{31}\mu_{32})$$
$$B_2[1] < \mu_{31}|A_{o1}| + \mu_{31}A_{o1} = 0$$

Since $|\mu_{31}| < 1$ and $|\mu_{32}| < 1$, it follows that $B_1[1] < 0$. Therefore, from (4.29), $P_i^{CD}[1] < 0$ for $i = 1, 2$. Now, consider $B_1[2]$: $B_1[2] = A_{out,1}[1] + K_{11}P_1^{CD}[1] + \mu_{32}B_2[1]$. From the previous step, $P_1^{CD}[1] < 0$ and $B_2[1] < 0$, and from the initial conditions, $A_{out,1}[1] = A_{o1} + \Delta A_{out,1}[0] = A_{o1} < 0$; thus, all terms in the expression for $B_1[2]$ are negative and therefore, $B_1[2] < 0$. Additionally, from (5.6), $B_2[2]$ is given by

$$B_2[2] = B_2[1] + K_{12}P_2^{CD}[1] + \mu_{31}\mu_{32}A_{o2}.$$

From the approximation that $|\mu_{31}^n\mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$, the above equation is approximated by

$$B_2[2] \approx B_2[1] + K_{12}P_2^{CD}[1].$$

From the previous step, both terms in the above expression are negative and therefore, $B_2[2] < 0$. Since both $B_1[2] < 0$ and $B_2[2] < 0$, it follows from (4.29) that $P_1^{CD}[2] < 0$ and $P_2^{CD}[2] < 0$.

Finally, consider $B_1[3]$: $B_1[3] = A_{out,1}[2] + K_{11}P_1^{CD}[2] + \mu_{32}B_2[2]$. From the $k = 2$ step, $A_{out,1}[2] < 0$ because $A_{out,1}[1] + P_1^{CD}[1] < 0$. Moreover, $B_2[2] < 0$ and $P_1^{CD}[2] < 0$; therefore, $B_1[3] < 0$. Additionally, from (5.6), $B_2[3]$ can be expressed as

$$B_2[3] = B_2[2] + K_{12}P_2^{CD}[2] + \mu_{31}\left(K_{11}P_1^{CD}[1] + \mu_{31}\mu_{32}A_{o1}\right).$$

From the $k = 2$ step, $B_2[2] < 0$, $P_2^{CD}[2] < 0$. From the $k = 1$ step, $P_1^{CD}[1] < 0$ and from the initial conditions, $A_{o1} < 0$; therefore, all the terms in the above expression negative, hence $B_2[3] < 0$. Consequently, $P_1^{CD}[3] < 0$ and $P_2^{CD}[3] < 0$.

More generally, suppose, for $3 \le k < \bar{k}$, $B_1[k] < 0$ and $B_2[k] < 0$; then from (4.29), $P_1^{CD}[k] < 0$ and $P_2^{CD}[k] < 0$ over this sample range. From (4.1), (4.24) and (4.28), $B_1[\bar{k}]$ can be expressed as

$$B_1[\bar{k}] = A_{o1} + \sum_{i=0}^{\bar{k}-1} K_{11}P_1^{CD}[i] + \mu_{32}B_2[\bar{k}-1].$$

The inductive hypothesis implies $P_1^{CD}[k] < 0$ for $3 \le k < \bar{k}$. Additionally, $P_1^{CD}[0] < 0$ from the initial conditions, $P_1^{CD}[1] < 0$ and $P_1^{CD}[2] < 0$ from previous steps. Finally, from the inductive hypothesis, $B_2[\bar{k}-1] < 0$. Thus all of the terms in the above equation are negative, therefore, $B_1[\bar{k}] < 0$. Therefore, by induction, $B_1[k] < 1$ for $k \ge 1$.

Equations (5.6) provides and expression for $B_2[k]$ for $k \ge 3$. Using the approximation $|\mu_{31}^n\mu_{32}^m| \approx 0$ for any m and n with $m + n \ge 2$, $B_2[k]$ can be approximated as

$$B_2[k] \approx B_2[k-1] + K_{12}P_2^{CD}[k-1] + \mu_{31}K_{11}P_1^{CD}[k-2],$$

, for $k \ge 3$. Recall that from the inductive hypothesis, $B_2[\bar{k}-1] < 0$, $P_2^{CD}[\bar{k}-1] < 0$ and $P_1^{CD}[\bar{k}-2] < 0$. Thus, all the terms in the above equation are negative and, thus, $B_2[\bar{k}] < 0$. Therefore, by induction, $B_2[k] < 0$ for $k \ge 1$. □

E.5 Proof of Lemma 5.5

Proof. From Lemma 5.4, $B_1[k] < 0$ and $B_2[k] < 0$ for $k \ge 1$ and thus, from (4.29), $P_i^{CD}[k] \le 0$ for $i = 1, 2$ and all $k \ge 1$. Along with the initial conditions on $A_{out,1}[k]$, this conclusion implies $A_{out,1}[k] < 0$ for all $k \ge 0$. Moreover, $A_{out,2}[k]$ is a non-increasing function of k for $k \ge 1$. As a result, either $A_{out,2}[k] \ge 0$ for $k \ge 0$, converging as k tends to infinity (by Lemma A.1), or there exists a k^* such that for $0 \le k < k^*$, $A_{out,2}[k] \ge 0$ and for $k \ge k^*$, $A_{out,2}[k] < 0$. Therefore, conclusions (a) and (b) of the lemma hold.

To show conclusion (c), i.e., $B_1[k]$ and $B_2[k]$ are decreasing functions of k for $k \geq 3$, consider the expressions for $B_1[k]$ and $B_2[k]$ given by (5.5) and (5.6). For $k \geq 3$, using $|\mu_{31}^n \mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$, $B_1[k]$ can be approximated by

$$B_1[k] \approx B_1[k-1] + K_{11} P_1^{CD}[k-1] + \mu_{32} K_{12} P_2^{CD}[k-2].$$

Since $B_1[k] < 0$ and $B_2[k] < 0$ for $k \geq 1$, it follows that $P_1^{CD}[k-1] \leq 0$ and $P_2^{CD}[k-2] \leq 0$ for $k \geq 3$. Therefore, $B_1[k]$ is approximately equal to the sum of its previous sample, and two non-positive terms, which means that $B_1[k]$ is a non-increasing function of k for $k \geq 3$. The exact same argument holds for $B_2[k]$ and thus, $B_2[k]$ is a non-increasing function of k for $k \geq 3$.

To show $B_1[k]$ and $B_2[k]$ converge as $k \to \infty$, note that since $B_2[k] < 0$ for $k \geq 1$, it follows that for $k \geq 2$, $A_{conf,1}[k] < 0$. Along with conclusion (a) and (4.29), this implies that for $k \geq 2$, $P_1^{CD}[k] = 0$. This result means that $A_{out,1}[k]$ remains constant for $k \geq 2$ and $B_1[k]$ can be expressed as

$$B_1[k] = A_{out,1}[2] + \mu_{32} B_2[k-1],$$

for $k \geq 2$. If $B_2[k-1]$ converges as k tends to infinity, then $B_1[k]$ also converges as k tends to infinity.

Consider now $B_2[k]$. If $A_{out,2}[k] < 0$ at some sample k^*, then for $k \geq k^*$, $A_{out,2}[k] < 0$, $A_{conf,2}[k] < 0$ and $B_2[k] < 0$, implying $P_2^{CD}[k] = 0$ for $k \geq k^*$. Since, for $k \geq 3$, $B_2[k]$ can be approximated by

$$B_2[k] \approx B_2[k-1] + K_{12} P_2^{CD}[k-1] + \mu_{31} K_{11} P_1^{CD}[k-2],$$

this conclusion, along with the fact that $P_1^{CD}[k] = 0$ for $k \geq 2$, implies that $B_2[k] \approx B_2[k-1]$ for $k \geq 3$ and thus, $B_2[k]$ converges as k tends to infinity. On the other hand, if $A_{out,2}[k] \geq 0$ for $k \geq 0$, then, for $k \geq 2$, $B_2[k]$ can be approximated as

$$B_2[k] \approx B_2[k-1] + K_{12} P_2^{CD}[k-1],$$

which can alternatively be expressed, for $k \geq 2$ as

$$B_2[k] \approx B_2[1] + K_{12} \sum_{i=1}^{k-1} P_2^{CD}[i].$$

To show $B_2[k]$ converges as k tends to infinity, we must show that $\sum_{i=1}^{k-1} P_2^{CD}[i]$ converges as k tends to infinity. Recall that since $A_{out,2}[k] \geq 0$ and is a non-increasing function of k for $k \geq 1$, Lemma A.1 applies; thus, $A_{out,2}[k]$ converges as $k \to \infty$. Since, for $k \geq 0$,

$$A_{out,2}[k] = A_{o2} + \sum_{i=0}^{k-1} K_{12} P_2^{CD}[i],$$

it follows that $\sum_{i=0}^{k-1} K_{12} P_2^{CD}[i]$ necessarily converges as k tends to infinity. As a result, $B_2[k]$ converges as k tends to infinity. Moreover, $P_2^{CD}[k]$ necessarily converges to zero as k tends to infinity.

Both in the case when $A_{out,2}[k] < 0$ for $k \geq k^*$ and when $A_{out,2}[k] \geq 0$ for $k \geq 0$, $B_2[k]$ converges as k tends to infinity. As a result, $B_1[k]$ also converges as k tends to infinity as previously argued. □

E.6 Proof of Lemma 5.7

Proof. First, the signs of $B_1[1]$, $B_1[2]$, $B_2[1]$ and $B_2[2]$ are established. The rest of the proof consists of two parts. Part 1 proves conclusion (a), whereas Part 2 proves conclusion (b).

$B_1[1]$ and $B_2[1]$ are given by

$$B_1[1] = A_{o1} + \mu_{32} A_{o2} \text{ and } B_2[1] = A_{o2} + \mu_{31} A_{o1}.$$

Since the initial conditions satisfy (5.8), it follows that

$$\begin{aligned} A_{o1} + \mu_{32} A_{o2} &< A_{o1} + |A_{o1}| \\ A_{o2} + \mu_{31} A_{o1} &\geq \mu_{31} |A_{o1}| + \mu_{31} A_{o1}. \end{aligned}$$

Since $A_{o1} < 0$, it follows that $B_1[1] < 0$ and $B_2[1] \geq 0$. This result implies that $P_1^{CD}[1] \leq 0$ and $P_2^{CD}[1] \geq 0$.

$B_1[2]$ and $B_2[2]$ are given by

$$\begin{aligned} B_1[2] &= A_{o1} + K_{11} P_1^{CD}[1] + \mu_{32} B_2[1] \\ B_2[2] &= A_{o2} + K_{12} P_2^{CD}[1] + \mu_{31} B_1[1]. \end{aligned}$$

Substituting the expressions for $B_1[1]$ and $B_2[1]$ into the above equations gives

$$\begin{aligned} B_1[2] &= A_{o1} + K_{11} P_1^{CD}[1] + \mu_{32} \left(A_{o2} + \mu_{31} A_{o1}\right) \\ B_2[2] &= A_{o2} + K_{12} P_2^{CD}[1] + \mu_{31} \left(A_{o1} + \mu_{32} A_{o2}\right). \end{aligned}$$

Note that $A_{o1} + \mu_{32} A_{o2} = B_1[1] < 0$. The remaining terms in the first equation, $K_{11} P_1^{CD}[1]$ and $\mu_{31}\mu_{32} A_{o1}$ are non-positive and, thus, $B_1[2] < 0$. Similarly, $B_2[2] \geq 0$. This result implies that $P_1^{CD}[2] \leq 0$ and $P_2^{CD}[2] \geq 0$.

Part 1

To show $B_1[3] \geq 0$ and $B_2[3] \geq 0$ implies $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k \geq 3$, induction is used with $k = 3$ as the base step. Suppose for $3 \leq k < \bar{k}$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$. Then, from (4.29), $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$ for $3 \leq k < \bar{k}$. From (5.5) and the approximation that $|\mu_{31}^n \mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$,

$$B_1[\bar{k}] \approx B_1[\bar{k} - 1] + K_{11} P_1^{CD}[\bar{k} - 1] + \mu_{32} K_{12} P_2^{CD}[\bar{k} - 2],$$

for $k \geq 3$. From the inductive hypothesis, $B_1[\bar{k} - 1] \geq 0$, and $P_1^{CD}[\bar{k} - 1] \geq 0$. Note that $P_2^{CD}[k] \geq 0$ for $0 \leq k < 3$ (from the initial conditions, and the $k = 1$ and $k = 2$ steps). Moreover, from the inductive hypothesis, $P_2^{CD}[k] \geq 0$ for $3 \leq k < \bar{k}$; therefore, $P_2^{CD}[\bar{k} - 2] \geq 0$. Therefore, all terms in the above equation are positive, and $B_1[\bar{k}] \geq 0$.

Additionally, $B_2[\bar{k}]$ can be expressed as

$$B_2[\bar{k}] = A_{o2} + K_{12} \sum_{i=0}^{\bar{k}-1} P_2^{CD}[i] + \mu_{31} B_[\bar{k} - 1],$$

for $k \geq 1$. Since $P_2^{CD}[k] \geq 0$ for $0 \leq k < \bar{k}$, $A_{o2} \geq 0$, and the inductive hypothesis that $B_1[\bar{k} - 1] \geq 0$, it follows that $B_2[\bar{k}] \geq 0$. Hence, by induction, for $k \geq 3$,

$$B_1[k] \geq 0 \text{ and}$$
$$B_2[k] \geq 0.$$

Part 2

To show $B_1[3] < 0$ and $B_2[3] < 0$ implies $B_1[k] < 0$ and $B_2[k] < 0$ for $k \geq 3$, induction is used. Suppose for $3 \leq k < \bar{k}$, $B_1[k] < 0$ and $B_2[k] < 0$. Then, from (4.29), $P_1^{CD}[k] \leq 0$ and $P_2^{CD}[k] \leq 0$ for $3 \leq k < \bar{k}$. From (5.6), and the approximation that $|\mu_{31}^n \mu_{32}^m| \approx 0$ for any m and n with $m + n \geq 2$,

$$B_2[\bar{k}] \approx B_2[\bar{k} - 1] + K_{12} P_2^{CD}[\bar{k} - 1] + \mu_{31} K_{11} P_1^{CD}[\bar{k} - 2]$$

for $k \geq 3$. From the inductive hypothesis, $B_2[\bar{k} - 1] < 0$ and $P_2^{CD}[\bar{k} - 1] \leq 0$. Note that $P_1^{CD}[k] \leq 0$ for $0 \leq k < 3$ (from the initial conditions and the $k = 1$ and $k = 2$ steps). Moreover, from the inductive hypothesis, $P_1^{CD}[k] \leq 0$ for $3 \leq k < \bar{k}$; therefore, $P_1^{CD}[\bar{k} - 2] \leq 0$, and thus, $B_2[\bar{k}] < 0$.

Additionally, $B_1[\bar{k}]$ can be expressed as

$$B_1[\bar{k}] = A_{o1} + K_{11} \sum_{i=0}^{\bar{k}-1} P_1^{CD}[i] + \mu_{32} B_[\bar{k} - 1],$$

for $k \geq 1$. Since $P_1^{CD}[k] \leq 0$ for $0 \leq k < \bar{k}$, $A_{o1} < 0$, and the inductive hypothesis that $B_2[\bar{k}-1] < 0$, it follows that $B_1[\bar{k}] < 0$. Hence, by induction, for $k \geq 3$,

$$B_1[k] < 0 \text{ and}$$
$$B_2[k] < 0.$$

□

E.7 Proof of Lemma 5.8

Proof. From Lemma 5.7, $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k \geq 3$. This conclusion, along with (4.29), implies that $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$ for $k \geq 3$. To show that $A_{out,2}[k] \geq 0$ for $k \geq 3$, it is first shown that $A_{out,2}[3] \geq 0$. Then, since $P_2^{CD}[k] \geq 0$ for $k \geq 3$, it follows that $A_{out,2}[k] \geq 0$ for $k \geq 3$. From the initial conditions, $A_{o2} \geq 0$. Moreover, it has previously been shown that if the zero-input, two-person system begins in region VII, then $B_2[1] \geq 0$ and $B_2[2] \geq 0$. As a result, $P_2^{CD}[1] \geq 0$ and $P_2^{CD}[2] \geq 0$ and hence, $\Delta A_{out,2}[1] \geq 0$ and $\Delta A_{out,2}[2] \geq 0$. Additionally, from the initial conditions, $\Delta A_{out,2}[0] = 0$. Thus,

$$A_{out,2}[3] = A_{o2} + K_{12}\left(P_2^{CD}[0] + P_2^{CD}[1] + P_2^{CD}[2]\right) \geq 0.$$

Therefore, $A_{out,2}[k] \geq 0$ for $k \geq 3$ as previously argued.

Moreover, since $P_1^{CD}[k] \geq 0$ for $k \geq 3$, $A_{out,1}[k]$ is a non-decreasing function of k for $k \geq 3$. To prove conclusion (b), the sign of $A_{out,1}[3]$ must be determined. Since, $A_{o1} < 0$, and, as previously shown, $B_1[1] < 0$ and $B_1[2] < 0$, it follows that $\Delta A_{out,1}[1] \leq 0$ and $\Delta A_{out,1}[2] \leq 0$. Moreover, from the initial conditions, $\Delta A_{out,1}[0] = 0$. Therefore,

$$A_{out,1}[3] = A_{o1} + K_{11}\left(P_1^{CD}[0] + P_1^{CD}[1] + P_1^{CD}[2]\right) < 0.$$

Since $A_{out,1}[k]$ is a non-decreasing function of k for $k \geq 3$, it follows that either $A_{out,1}[k] < 0$ for all $k \geq 0$, which implies that $A_{out,1}[k]$ converges as k tends to infinity, or, there exists some k^* such that for $0 \leq k < k^*$, $A_{out,1}[k] < 0$ and for all $k \geq k^*$, $A_{out,1}[k] \geq 0$.

To show $B_1[k]$ and $B_2[k]$ converge as $k \to \infty$, note that since $B_1[k] \geq 0$ for $k \geq 3$, it follows that for $k \geq 4$, $A_{conf,2}[k] \geq 0$. Along with conclusion (a) and (4.29), this implies that for $k \geq 4$, $P_2^{CD}[k] = 0$. This result means that $A_{out,2}[k]$ remains constant for $k \geq 4$, and $B_2[k]$ can be expressed as

$$B_2[k] = A_{out,2}[4] + \mu_{31} B_1[k-1]$$

for $k \geq 4$. If $B_1[k-1]$ converges as k tends to infinity, then $B_2[k]$ also converges as k tends to infinity.

Consider now $B_1[k]$. If $A_{out,1}[k] \geq 0$ at some sample k^*, then for $k \geq k^*$, $A_{out,1}[k] \geq 0$, $A_{conf,1}[k] \geq 0$ and $B_1[k] \geq 0$, implying $P_1^{CD}[k] = 0$. Since $B_1[k]$ can be approximated by

$$B_1[k] \approx B_1[k-1] + K_{11}P_1^{CD}[k-1] + \mu_{32}K_{12}P_2^{CD}[k-2],$$

this conclusion implies that $B_1[k] \approx B_1[k-1]$ and thus, $B_1[k]$ converges as k tends to infinity. On the other hand, if $A_{out,1}[k] < 0$ for $k \geq 0$, then, for $k \geq 4$, $B_1[k]$ can be approximated as

$$B_1[k] \approx B_1[k-1] + K_{11}P_1^{CD}[k-1],$$

which can alternatively be expressed, for $k \geq 4$ as

$$B_1[k] \approx B_1[3] + K_{11}\sum_{i=3}^{k-1} P_1^{CD}[i].$$

To show $B_1[k]$ converges as k tends to infinity, we must show that $\sum_{i=3}^{k-1} P_1^{CD}[i]$ converges as k tends to infinity. Recall that since $A_{out,1}[k] < 0$ and is a non-decreasing function of k for $k \geq 3$, Lemma A.2 applies; thus, $A_{out,1}[k]$ converges as $k \to \infty$. Since, for $k \geq 0$,

$$A_{out,1}[k] = A_{o1} + \sum_{i=0}^{k-1} K_{11}P_1^{CD}[i],$$

it follows that $\sum_{i=0}^{k-1} K_{11}P_1^{CD}[i]$ necessarily converges as k tends to infinity. As a result, $B_1[k]$ converges as k tends to infinity. Moreover, $P_1^{CD}[k]$ necessarily converges to zero as k tends to infinity.

Both in the case when $A_{out,1}[k] \geq 0$ for $k \geq k^*$ and when $A_{out,1}[k] < 0$ for $k \geq 0$, $B_1[k]$ converges as k tends to infinity. As a result, $B_2[k]$ converges as k tends to infinity as previously argued. □

E.8 Proof of Lemma 5.10

Proof. First, we show that for $k \geq \bar{k}$, $A_{out,1}[k] \geq 0$, which is done by induction, with $k = \bar{k}$ as the base step. Thus, suppose, for $\bar{k} \leq k < \tilde{k}$, $A_{out,1}[k] \geq 0$. Since $A_{rew,1}[k] \geq 0$ for all $k \geq 0$ and since $A_{conf,1}[k] \geq 0$ for $k \geq \bar{k}$, it follows from the expression for $B_1[k]$, i.e.,

$$B_1[k] = A_{out,1}[k] + A_{rew,1}[k] + A_{conf,1}[k],$$

that $B_1[k] \geq 0$ for $\bar{k} \leq k < \tilde{k}$. It is now shown that $A_{out,1}[\tilde{k}] \geq 0$. Recall that $A_{out,1}[\tilde{k}]$ can be expressed as

$$A_{out,1}[\tilde{k}] = A_{out,1}[\tilde{k}-1] + \Delta A^{CD}_{out,1}[\tilde{k}-1] + \Delta A^{OJ,rew}_{out,1}[\tilde{k}-1].$$

Since at $\tilde{k}-1$, all cognitions are non-negative, from (4.29), $P^{CD}_1[\tilde{k}-1] = 0$, which implies that $\Delta A^{CD}_{out,1}[\tilde{k}-1] = 0$. Therefore, the above equation simplifies to

$$A_{out,1}[\tilde{k}] = A_{out,1}[\tilde{k}-1] + \Delta A^{OJ,rew}_{out,1}[\tilde{k}-1].$$

From (4.22), there are three possible expressions for $\Delta A^{OJ,rew}_{out,1}[\tilde{k}-1]$, each of which is examined below.

First suppose $P^{OJ,rew}_1[\tilde{k}-1] > 0$ and $A_{out,1}[\tilde{k}-1] \geq K_{21} P^{OJ,rew}_1[\tilde{k}-1]$. Then, from (4.22)

$$\begin{aligned} A_{out,1}[\tilde{k}] &= A_{out,1}[\tilde{k}-1] + \Delta A^{OJ,rew}_{out,1}[\tilde{k}-1] \\ \Rightarrow A_{out,1}[\tilde{k}] &= A_{out,1}[\tilde{k}-1] - K_{21} P^{OJ,rew}_1[\tilde{k}-1] \geq 0. \end{aligned}$$

Instead, suppose $P^{OJ,rew}_1 \tilde{k}-1] > 0$ and $K_{21} P^{OJ}_1[\tilde{k}-1] > A_{out,1}[\tilde{k}-1]$. Then, from (4.22)

$$\begin{aligned} A_{out,1}[\tilde{k}] &= A_{out,1}\tilde{k}-1] + \Delta A^{OJ,rew}_{out,1}[\tilde{k}-1] \\ \Rightarrow A_{out,1}[\tilde{k}] &= A_{out,1}[\tilde{k}-1] - A_{out,1}[\tilde{k}-1] = 0. \end{aligned}$$

Finally, suppose $P^{OJ,rew}_1[\tilde{k}-1] = 0$. Then, from (4.22)

$$\begin{aligned} A_{out,1}[\tilde{k}1] &= A_{out,1}[\tilde{k}-1] + \Delta A^{OJ}_{out,1}[\tilde{k}-1] \\ \Rightarrow A_{out,1}[\tilde{k}] &= A_{out,1}[\tilde{k}-1] + 0. \end{aligned}$$

Since $A_{out,1}[\tilde{k}-1] \geq 0$, $A_{out,1}[\tilde{k}] \geq 0$. Therefore, for any possible expression of $\Delta A^{OJ,rew}_{out,1}[\tilde{k}-1]$, $A_{out,1}[\tilde{k}] \geq 0$. Therefore, by induction, $A_{out,1}[k] \geq 0$ for $k \geq \bar{k}$.

The fact that $A_{out,1}[k]$ is a non-increasing function of k for $k \geq \bar{k}$ follows directly from the equation for $A_{out,1}[k]$ for $k \geq \bar{k}$. Since, for $k \geq \bar{k}$, $A_{out,1}[k] \geq 0$, $A_{rew,1}[k] \geq 0$ and $A_{conf,1}[k] \geq 0$, all cognitions are non-negative, and therefore, $P^{CD}_1[k] = 0$. As a result, for $k > \bar{k}$,

$$A_{out,1}[k] = A_{out,1}[k-1] + \Delta A^{OJ,rew}_{out,1}[k-1].$$

From (4.22), $\Delta A^{OJ,rew}_{out,1}[k-1] \leq 0$, and thus, $A_{out,1}[k]$ is a non-increasing function of k for $k \geq \bar{k}$. □

E.9 Proof of Lemma 5.11

Proof. The proof to show (5.18) is sufficient for a step reward to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ is split into two main parts. In Part A, the first few samples of $B_1[k]$ and $B_2[k]$ are examined. In Part B, a more general expression for each behaviour is used, along with induction, to show that $B_1[2] \geq 0$ and $B_2[2] \geq 0$ implies $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k \geq 2$. Part B considers two separate cases: one when $A_{out,1}[k] < 0$ for all $k \geq 0$ and the other when $A_{out,1}[k]$ becomes positive as some sample $\bar{k}$.

Part A

From (5.11), the expression for $B_1[1]$ is given by $B_1[1] = A_{out,1}[1] + \mu_{32}B_2[0] + \mu_{11}R_o$. Using the first relationship given in (5.18) and initial conditions given in (4.26) gives

$$\begin{aligned} B_1[1] &\geq A_{o1} + \mu_{32}A_{o2} - A_{o1}\left(1 + \mu_{31}\mu_{32}\right) - \mu_{32}A_{o2} \\ &= -\mu_{31}\mu_{32}A_{o1}. \end{aligned}$$

Since $A_{o1} < 0$, it follows that $B_1[1] \geq 0$. Note from (4.29), this implies $P_1^{CD}[1] \geq 0$ and since $A_{out,1}[1] < 0$ and $B_1[1] \geq 0$, overjustification pressure does not arise at $k = 1$, i.e., $\Delta A_{out,1}^{OJ,rew}[1] = 0$.

Additionally, from (5.12), $B_2[1]$ is given by $B_2[1] = A_{out,2}[1] + \mu_{31}B_1[0]$. Since the system can begin in region V, VII or VIII, the sign of $B_2[1]$ is unknown. However, for region V, $A_{o1} < 0$ and $A_{o2} < 0$ and therefore, $B_2[1] < 0$. On the other hand, in region VII and VIII, it has previously been shown that $B_2[1] \geq 0$. These conclusions, together, imply $P_2^{CD}[1] \geq 0$. Furthermore, recall that $P_2^{OJ,rew}[k] = 0$ for all $k \geq 0$ by Assumptions 4.2 and 4.4.

Now consider $k = 2$. From (5.11), $B_1[2]$ can be given by

$$\begin{aligned} B_1[2] &= A_{out,1}[2] + \mu_{11}R_o + \mu_{32}B_2[1] \\ &= A_{o1} + K_{11}P_1^{CD}[1] + \Delta A_{out,1}^{OJ,rew}[1] + \mu_{11}R_o + \mu_{32}B_2[1] \end{aligned}$$

From before, $\Delta A_{out,1}^{OJ,rew}[1] = 0$. Additionally, using the first expression for R_o given in (5.18) and substituting the expression for $B_2[1]$ gives

$$\begin{aligned} B_1[2] &\geq A_{o1} + K_{11}P_1^{CD}[1] - A_{o1}\left(1 + \mu_{31}\mu_{32}\right) - \mu_{32}A_{o2} + \mu_{32}\left(A_{o2} + \mu_{31}A_{o1}\right) \\ &= K_{11}P_1^{CD}[1]. \end{aligned}$$

From before, $P_1^{CD}[1] \geq 0$. Therefore, $B_1[2] \geq 0$, implying $P_1^{CD}[2] \geq 0$. Note that, unlike at $k = 1$, the sign of $A_{out,1}[2]$ is unknown (since $P_1^{CD}[1] \geq 0$, which may drive $A_{out,1}[2] \geq 0$). Moreover, since the sign of $B_2[1]$ may be also be positive (if the system begins in region

VII or VIII), then overjustification may arise, meaning $\Delta A_{out,1}^{OJ,rew}[2]$ may be non-zero. This has implications at the $k = 3$ step.

Now consider $B_2[2]$, which is given by $B_2[2] = A_{out,2}[2] + \mu_{31} B_1[1]$. If the system begins in region VII or VIII, then $A_{out,2}[2] = A_{o2} + K_{12} P_2^{CD}[1] \geq 0$. Since $B_1[1] \geq 0$, it follows that $B_2[2] \geq 0$. On the other hand, if the system begins in region V, then $A_{out,2}[2] < 0$, since $P_2^{CD}[1] = 0$. However, by substituting the expression for $B_1[1]$ into the above equation and using the second expression for R_o given in (5.18), the following inequality is obtained:

$$\begin{aligned} B_2[2] &\geq A_{o2} + K_{12} P_2^{CD}[1] + \mu_{31} \left(A_{o1} + \mu_{32} A_{o2} \right) - \left(A_{o2} + \mu_{31} A_{o1} \right) \left(1 + \mu_{31} \mu_{32} \right) \\ &= -\mu_{31}^2 \mu_{32} A_{o1}. \end{aligned}$$

From the initial conditions, $A_{o1} < 0$ and thus, $B_2[2] \geq 0$, which implies $P_2^{CD}[2] \geq 0$.

The final sample to consider in Part A is $k = 3$. From (5.11), $B_1[3]$ can be expressed as

$$B_1[3] = A_{out,1}[3] + \mu_{11} R_o + \mu_{32} B_2[2].$$

$A_{out,1}[3]$ can be expanded to give

$$B_1[3] = A_{out,1}[2] + K_{11} P_1^{CD}[2] + \Delta A_{out,1}^{OJ,rew}[2] + \mu_{11} R_o + \mu_{32} B_2[2].$$

From the $k = 2$ step, $P_1^{CD}[2] \geq 0$ and $B_2[2] \geq 0$. Moreover, if overjustification does arise $k = 2$, then Lemma 5.10 implies $A_{out,1}[2] + A_{out,1}^{OJ,rew}[2] \geq 0$. Thus, $B_1[3] \geq 0$. On the other hand, if overjustification does not arise at $k = 2$, then $\Delta A_{out,1}^{OJ,rew}[2] = 0$, in which case, the above expression for $B_1[3]$ must be further expanded using the expressions for $A_{out,1}[2]$ and $B_2[2]$:

$$\begin{aligned} B_1[3] = A_{o1} &+ K_{11} \left(P_1^{CD}[1] + P_1^{CD}[2] \right) + \mu_{11} R_o \\ &+ \mu_{32} \left(A_{o2} + K_{12} P_2^{CD}[1] + \mu_{31} \left(A_{o1} + \mu_{32} A_{o2} + \mu_{11} R_o \right) \right). \end{aligned}$$

Substituting the first expression for R_o from (5.18) into the two reward terms gives

$$\begin{aligned} B_1[3] &\geq A_{o1} + K_{11} \left(P_1^{CD}[1] + P_1^{CD}[2] \right) - A_{o1} \left(1 + \mu_{31} \mu_{32} \right) - \mu_{32} A_{o2} \\ &\quad + \mu_{32} \left(A_{o2} + K_{12} P_2^{CD}[1] + \mu_{31} \left(A_{o1} + \mu_{32} A_{o2} - A_{o1} \left(1 + \mu_{31} \mu_{32} \right) - \mu_{32} A_{o2} \right) \right) \\ &= K_{11} \left(P_1^{CD}[1] + P_1^{CD}[2] \right) - A_{o1} \left(\mu_{31} \mu_{32} \right) + \mu_{32} \left(K_{12} P_2^{CD}[1] - A_{o1} \mu_{31}^2 \mu_{32} \right). \end{aligned}$$

Since $A_{o1} < 0$, $P_1^{CD}[1] \geq 0$, $P_1^{CD}[2] \geq 0$ and $P_2^{CD}[1] \geq 0$, all of the terms in the above expression are non-negative and thus $B_1[3] \geq 0$. Note that this implies $P_1^{CD}[3] \geq 0$. Similar to the $k = 2$ step, the sign of $A_{out,1}[3]$ is unknown, thus, $\Delta A_{out,1}^{OJ,rew}[3]$ may be non-zero.

Finally, consider $B_2[3]$, which can be expressed as

$$B_2[3] = A_{o2} + K_{12} \left(P_2^{CD}[1] + P_2^{CD}[2] \right) + \mu_{31} B_1[2].$$

If the system begins in region VII or VIII, then $A_{o2} \geq 0$ and $P_2^{CD}[1] \geq 0$. Moreover, from the $k = 2$ step, $P_2^{CD}[2] \geq 0$ and $B_1[2] \geq 0$. Thus all the terms on the above equations are non-negative, and $B_2[3] \geq 0$. On the other hand, if the system begins in region V, then $A_{o2} < 0$ and $P_2^{CD}[1] = 0$. In this case, the expression for $B_1[2]$ is further expanded:

$$B_2[3] = A_{o2} + K_{12} P_2^{CD}[2] + \mu_{31}\left(A_{o1} + K_{11} P_1^{CD}[1] + \mu_{11} R_o + \mu_{32} A_{o2} + \mu_{31}\mu_{32} A_{o1}\right).$$

Substituting the second expression for R_o from (5.18) gives

$$\begin{aligned} B_2[3] &\geq A_{o2} + K_{12} P_2^{CD}[2] - (A_{o2} + \mu_{31} A_{o1})(1 + \mu_{31}\mu_{32}) \\ &\quad + \mu_{31}\left(A_{o1} + K_{11} P_1^{CD}[1] + \mu_{32} A_{o2} + \mu_{31}\mu_{32} A_{o1}\right) \\ &= K_{12} P_2^{CD}[2] + \mu_{31} K_{11} P_1^{CD}[1]. \end{aligned}$$

Since $P_2^{CD}[2] \geq 0$ and $P_1^{CD}[1] \geq 0$, it follows that $B_2[3] \geq 0$.

Thus, in Part A, it has been shown that for $1 \leq k \leq 3$, $B_1[k] \geq 0$ and for $2 \leq k \leq 3$, $B_2[k] \geq 0$. As a result, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$ for $1 \leq k \leq 3$ (since $P_2^{CD}[1] \geq 0$ both when the system begins in region V, VII or VIII). Moreover, the sign of $A_{out,1}[k]$ is unknown for samples of $k \geq 2$, since $A_{o1} < 0$ and $A_{out,1}[k]$ is an increasing function of k for $1 \leq k \leq 3$.

Part B

Now, induction is used to show that if $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k = 3$, then for $k \geq 3$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$. Two cases are considered. *Case 1* considers the situation when $A_{out,1}[k] < 0$ for $k \geq 0$, whereas *Case 2* considers the situation when $A_{out,1}[k] \geq 0$ at some $k = \hat{k}$. For both cases, the following expressions are used for $B_1[k]$ and $B_2[k]$:

$$\begin{aligned} B_1[k] &= A_{o1} + \sum_{i=0}^{k-1} \Delta A_{out,1}[i] + \mu_{11} R_o + \mu_{32} B_2[k-1] \\ B_2[k] &= A_{o2} + \sum_{i=0}^{k-1} K_{12} P_2^{CD}[i] + \mu_{31} B_1[k-1]. \end{aligned}$$

For *Case 1*,

$$\Delta A_{out,1}[k] = K_{11} P_1^{CD}[k] \text{ for } k \geq 0,$$

whereas, for *Case 2*,

$$\Delta A_{out,1}[k] = \begin{cases} K_{11} P_1^{CD}[k] & \text{for } 0 \leq k \leq \hat{k}, \\ K_{11} P_1^{CD}[k] + \Delta A_{out,1}^{OJ,rew}[k] & \text{for } k > \hat{k}. \end{cases}$$

Each case is examined separately below.

Case 1

Suppose, for $3 \leq k < \bar{k}$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$. Then, over this sample range, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$. Moreover, since $A_{out,1}[k] < 0$ for $k \geq 0$, it follows that $\Delta A_{out,1}^{OJ,rew}[k] = 0$ for $k \geq 0$. Using the first expression for R_o in (5.18) results in the following inequality:

$$B_1[\bar{k}] \geq \sum_{i=0}^{\bar{k}-1} K_{11} P_1^{CD}[i] - \mu_{31}\mu_{32} A_{o1} - \mu_{32} A_{o2} + \mu_{32} B_2[\bar{k}-1]$$

If the system begins in region V, then $A_{o1} < 0$ and $A_{o2} < 0$. Moreover, from the initial conditions, $k = 1, 2, 3$ steps, and the inductive hypothesis, $P_1^{CD}[k] \geq 0$ for $0 \leq k < \bar{k}$. Additionally, from the inductive hypothesis, $B_2[\bar{k}-1] \geq 0$; thus all terms in the above inequality are non-negative and $B_1[\bar{k}] \geq 0$. On the other hand, if the system begins in region VII or VIII, then, by expanding the $B_2[\bar{k}-1]$ term, the above inequality becomes

$$\begin{aligned} B_1[\bar{k}] &\geq \sum_{i=0}^{\bar{k}-1} K_{11} P_1^{CD}[i] - \mu_{32}\left(\mu_{31} A_{o1} + A_{o2}\right) + \mu_{32}\left(A_{o2} + \sum_{i=0}^{\bar{k}-2} K_{12} P_2^{CD}[i] + \mu_{31} B_1[\bar{k}-2]\right) \\ &= \sum_{i=0}^{\bar{k}-1} K_{11} P_1^{CD}[i] - \mu_{31}\mu_{32} A_{o1} + \mu_{32}\left(\sum_{i=0}^{\bar{k}-2} K_{12} P_2^{CD}[i] + \mu_{31} B_1[\bar{k}-2]\right). \end{aligned}$$

From the initial conditions, $k = 1, 2, 3$ steps, and the inductive hypothesis, $P_1^{CD}[k] \geq 0$ for $0 \leq k < \bar{k}$ and $B_1[\bar{k}-2] \geq 0$; thus all terms in the above inequality are non-negative and $B_1[\bar{k}] \geq 0$.

Now, consider $B_2[\bar{k}]$ when $A_{out,1}[k] < 0$ for $k \geq 0$. Since

$$B_2[\bar{k}] = A_{o2} + \sum_{i=0}^{\bar{k}-1} K_{12} P_2^{CD}[i] + \mu_{31} B_1[\bar{k}-1],$$

if the system begins in region VII or VIII, then all terms in the above inequality are non-negative and thus, $B_2[\bar{k}] \geq 0$. On the other hand, if the system begins in region V, then $A_{o2} < 0$ and thus, $B_1[\bar{k}-1]$ is expanded, giving

$$B_2[\bar{k}] = A_{o2} + \sum_{i=0}^{\bar{k}-1} K_{12} P_2^{CD}[i] + \mu_{31}\left(A_{o1} + \sum_{i=0}^{\bar{k}-2} K_{11} P_1^{CD}[i] + \mu_{11} R_o + \mu_{32} B_2[\bar{k}-2]\right).$$

Substituting the second expression from (5.18) for R_o in the above expression yields the following inequality for $B_2[\bar{k}]$:

$$
\begin{aligned}
B_2[\bar{k}] &\geq A_{o2} + \sum_{i=0}^{\bar{k}-1} K_{12} P_2^{CD}[i] - (A_{o2} + \mu_{31} A_{o1})(1 + \mu_{31}\mu_{32}) \\
&\quad + \mu_{31} \left(A_{o1} + \sum_{i=0}^{\bar{k}-2} K_{11} P_1^{CD}[i] + \mu_{32} B_2[\bar{k}-2] \right) \\
&= \sum_{i=0}^{\bar{k}-1} K_{12} P_2^{CD}[i] - \mu_{31}\mu_{32}(A_{o2} + \mu_{31} A_{o1}) + \mu_{31} \left(\sum_{i=0}^{\bar{k}-2} K_{11} P_1^{CD}[i] + \mu_{32} B_2[\bar{k}-2] \right).
\end{aligned}
$$

From the initial conditions, $A_{o1} < 0$, $A_{o2} < 0$, $P_1^{CD}[0] = 0$ and $P_2^{CD}[0] = 0$. Moreover, from $k = 1, 2, 3$ steps, and the inductive hypothesis, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$ for $1 \leq k < \bar{k}$, and $B_2[\bar{k}-2] \geq 0$; thus, $B_2[\bar{k}] \geq 0$.

To summarize, if $A_{out,1}[k] < 0$ for $k \geq 0$, then, by induction, $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$.

Case 2

Suppose, $A_{out,1}[k] \geq 0$ at $k = \hat{k}$. In this situation, the results of *Case 1* hold for $3 \leq k < \hat{k}$, i.e., for $3 \leq k < \hat{k}$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$ and thus, over this sample range, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$. To show $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k \geq \hat{k}$, induction is used. Before giving the inductive hypothesis, it must be shown that $B_1[\hat{k}] \geq 0$ and $B_2[\hat{k}] \geq 0$. Then the inductive hypothesis can be given and used to shown that for $k \geq \hat{k}$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$. Since $B_1[\hat{k}]$ is given by

$$B_1[\hat{k}] = A_{out,1}[\hat{k}] + \mu_{11} R_o + \mu_{32} B_2[\hat{k}-1],$$

all of the terms in the above expression are non-negative and thus, $B_1[\hat{k}] \geq 0$.

Now consider $B_2[\hat{k}]$:

$$B_2[\hat{k}] = A_{o2} + \sum_{i=0}^{\hat{k}-1} K_{12} P_2^{CD}[i] + \mu_{32} B_1[\hat{k}-1].$$

If the system begins in region VII or VIII, then $A_{o2} > 0$. Moreover, for $0 \leq k < \hat{k}$, $P_2^{CD}[k] \geq 0$ and at $k = \hat{k}$, and, from *Case 1*, $B_1[\hat{k}-1] \geq 0$; thus $B_2[\hat{k}] \geq 0$. On the other hand, if the system begins in region V, then $A_{o2} < 0$; thus, $B_1[\hat{k}-1]$ can be expanded to

$$B_2[\hat{k}] = A_{o2} + \sum_{i=0}^{\hat{k}-1} K_{12} P_2^{CD}[i] + \mu_{31} \left(A_{o1} + \sum_{i=0}^{\hat{k}-2} K_{11} P_1^{CD}[i] + \mu_{11} R_o + \mu_{32} B_2[\hat{k}-2] \right).$$

Substituting the second expression in (5.18) for R_o into the above equation gives

$$
\begin{aligned}
B_2[\hat{k}] &\geq A_{o2} + \sum_{i=0}^{\hat{k}-1} K_{12} P_2^{CD}[i] - \left(A_{o2} + \mu_{31} A_{o1}\right)\left(1 + \mu_{31}\mu_{32}\right) \\
&\quad + \mu_{31}\left(A_{o1} + \sum_{i=0}^{\hat{k}-2} K_{11} P_1^{CD}[i] + \mu_{32} B_2[\hat{k}-2]\right) \\
&\geq \sum_{i=0}^{\hat{k}-1} K_{12} P_2^{CD}[i] - \mu_{31}\mu_{32}\left(A_{o2} + \mu_{31} A_{o1}\right) + \mu_{31}\left(\sum_{i=0}^{\hat{k}-2} K_{11} P_1^{CD}[i] + \mu_{32} B_2[\hat{k}-2]\right).
\end{aligned}
$$

From Part 1, for $0 \leq k \leq 3$, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$. Additionally, for $3 \leq k < \hat{k}$, *Case 1* applies and thus $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$ for all $3 \leq k < \hat{k}$ and therefore, the summation terms in the above expression are non-negative. Finally, since $B_2[\hat{k}-2] \geq 0$ (from *Case 1*) and $A_{o1} < 0$ and $A_{o2} < 0$ (from the initial conditions of region V), all of the terms in the above expression are non-negative, and thus, $B_2[\hat{k}] \geq 0$.

More generally, suppose for $\hat{k} \leq k < \bar{k}$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$. Therefore, for this sample range, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$. Moreover, by Lemma 5.10, $A_{out,1}[k] \geq 0$ for $\hat{k} \leq k < \bar{k}$. From the system equations, $B_1[\bar{k}]$ can be given by

$$
B_1[\bar{k}] = A_{out,1}[\bar{k}-1] + K_{11} P_1^{CD}[\bar{k}-1] + \Delta A_{out,1}^{OJ}[\bar{k}-1] + \mu_{11} R_o + \mu_{32} B_2[\bar{k}-1].
$$

Even though $\Delta A_{out,1}^{OJ}[\bar{k}-1] \leq 0$, Lemma 5.10 implies that $A_{out,1}[\bar{k}-1] + \Delta A_{out,1}^{OJ,rew}[\bar{k}-1] \geq 0$. Additionally, from the inductive hypothesis, $P_1^{CD}[\bar{k}-1] \geq 0$ and $B_2[\bar{k}-1] \geq 0$. Therefore, $B_1[\bar{k}] \geq 0$.

Finally, $B_2[\bar{k}]$ can be expressed as

$$
B_2[\bar{k}] = A_{o2} + \sum_{i=0}^{\bar{k}-1} K_{12} P_2^{CD}[i] + \mu_{32}\left(A_{out,1}[\bar{k}-1] + \mu_{11} R_o + \mu_{32} B_2[\bar{k}-2]\right).
$$

If the system begins in region VII or VIII, then all terms in the above expression are non-negative and thus, $B_2[\bar{k}] \geq 0$. On the other hand, if the system begins in region V, then, using the second expression for R_o in (5.18) gives

$$\begin{aligned} B_2[\bar{k}] &\geq A_{o2} + \sum_{i=0}^{\bar{k}-1} K_{12} P_2^{CD}[i] - \left(A_{o2} + \mu_{31} A_{o1}\right)\left(1 + \mu_{31}\mu_{32}\right) \\ &\quad + \mu_{31}\left(A_{out,1}[\bar{k}-1] + \mu_{32} B_2[\bar{k}-2]\right) \\ &= \sum_{i=0}^{\bar{k}-1} K_{12} P_2^{CD}[i] - \mu_{31}\left(A_{o1}\left(1 + \mu_{31}\mu_{32}\right) + \mu_{32} A_{o2}\right) \\ &\quad + \mu_{31}\left(A_{out,1}[\bar{k}-1] + \mu_{32} B_2[\bar{k}-2]\right). \end{aligned}$$

From Part 1, for $0 \leq k \leq 3$, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$, from *Case 1*, and the inductive hypothesis, for $3 \leq k < \bar{k}$, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$, $B_2[\bar{k}] \geq 0$, and from the initial conditions of region V, $A_{o1}, A_{o2} < 0$. Finally, it follows from the inductive hypothesis that $A_{out,1}[\bar{k}-1] \geq 0$; thus, all of the terms in the above expression are non-negative, and, $B_2[\bar{k}] \geq 0$. Therefore, in the case when $A_{out1}[k] \geq 0$ at some $k = \hat{k}$, the reward given in (5.18) is sufficient to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$.

To summarize, for both *Case 1* and *Case 2*, a step-reward magnitude satisfying (5.18) ensures $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ □

E.10 Proof of Lemma 5.12

Proof. The proof is split into three parts. Part 1 considers the first few samples of the system's response. Part 2 uses induction to show that if $A_{out,1}[k] < 0$ for all $k \geq 0$, then the reward given in (5.20) is sufficient to ensure $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. Part 3 shows that if $A_{out,1}[k]$ becomes positive, then the reward is sufficient to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$.

Part 1

The equations for $B_1[k]$ and $B_2[k]$ given in (4.24) can be expanded to use the following form:

$$\begin{aligned} B_1[k] = A_{o1} + \sum_{i=1}^{k-1} K_{11} P_1^{CD}[i] + \sum_{i=1}^{k-1} \Delta A_{out,1}^{OJ,rew}[i] + \mu_{11} R_o \\ + \mu_{32}\left(A_{o2} + \sum_{i=1}^{k-2} K_{12} P_2^{CD}[i] + \mu_{31} B_1[k-2]\right), \end{aligned} \tag{E.1}$$

and

$$B_2[k] = A_{o2} + \sum_{i=1}^{k-1} K_{12} P_2^{CD}[i] + \mu_{31} \Bigg(A_{o1} + \sum_{i=1}^{k-2} K_{11} P_1^{CD}[k-2]$$
$$+ \sum_{i=1}^{k-2} \Delta A_{out,1}^{OJ,rew}[i] + \mu_{11} R_o + \mu_{32} B_2[k-2] \Bigg). \tag{E.2}$$

These forms for $B_1[k]$ and $B_2[k]$ are used for the first two parts of the proof.

First, we show that the reward in (5.20) satisfies the requirements of the strategy to drive $B_1[1] \geq 0$ and $B_2[2] \geq 0$. From the equations given in (E.1) and (E.2), $B_1[1]$ is given by $B_1[1] = A_{o1} + \mu_{32} A_{o2} + \mu_{11} R_o$. By substituting the first expression for R_o from (5.20) into the above expression for $B_1[1]$, we obtain

$$\begin{aligned} B_1[1] &\geq A_{o1} + \mu_{32} A_{o2} - A_{o1} \left(1 + \mu_{31}\mu_{32}\right) - \mu_{32} A_{o2} \\ &= -\mu_{31}\mu_{32} A_{o1}. \end{aligned}$$

From the initial conditions, $A_{o1} < 0$, and therefore, the reward in (5.20) is sufficient to drive $B_1[1] \geq 0$. Furthermore, from (4.29), $B_1[1] \geq 0$ implies $P_1^{CD}[1] \geq 0$, and since $A_{out,1}[1] = A_{o1} < 0$, $P_1^{OJ,rew}[1] = 0$.

Now consider $B_2[2]$, which from the equations given in (E.1) and (E.2), can be expressed as

$$B_2[2] = A_{o2} + K_{12} P_2^{CD}[1] + \mu_{31} \left(A_{o1} + \mu_{32} A_{o2} + \mu_{11} R_o \right).$$

To determine the sign of $P_2^{CD}[1]$, consider the relationship between A_{o1} and A_{o2} from (5.4). It has previously been shown that for the zero-input system beginning in initial region VI, $B_2[1] < 0$. Since the $R[0]$ first affects $B_2[2]$, it follows that in the case of a step reward, $B_2[1]$ is the same as the in the zero-input case, and thus,

$$P_2^{CD}[1] = -\frac{|A_{out,2}[1]|}{|A_{conf,2}[1]| + |A_{out,2}[1]|} < 0,$$

which, when using the initial conditions, simplifies to

$$P_2^{CD}[1] = -\frac{A_{o2}}{\mu_{31}|A_{o1}| + A_{o2}}.$$

Using the above expression for $P_2^{CD}[1]$ and the second expression for R_o from (5.20) in the expression for $B_2[2]$ gives

$$\begin{aligned} B_2[2] \geq A_{o2} &- \frac{K_{12} A_{o2}}{(A_{o2} + \mu_{31}|A_{o1}|)} + \mu_{31} \left(A_{o1} + \mu_{32} A_{o2} \right) \\ &+ \frac{K_{12} A_{o2}}{(A_{o2} + \mu_{31}|A_{o1}|)} - \left(A_{o2} + \mu_{31} A_{o1} \right) \left(1 + \mu_{31}\mu_{32} \right) \\ &= -\mu_{31}^2 \mu_{32} A_{o1}. \end{aligned}$$

From the initial conditions, $A_{o1} < 0$, and therefore, the reward in (5.20) is sufficient to drive $B_2[2] \geq 0$, hence $P_2^{CD}[2] \geq 0$.

Part 2

Now, induction is used to show that if person one does not experience the overjustification effect, i.e., $A_{out,1}[k] < 0$ for all $k \geq 0$, then $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. Suppose then, that $A_{out,1}[k] < 0$ for all $k \geq 0$. Before proceeding to the base step, the sign of $B_1[2]$ must be determined. From (E.1) and the initial conditions, $B_1[2]$ is given by

$$B_1[2] = A_{o1} + K_{11}P_1^{CD}[1] + \Delta A_{out,1}^{OJ,rew}[1] + \mu_{11}R_o + \mu_{32}\left(A_{o2} + \mu_{31}A_{o1}\right).$$

However, as previously shown, $P_1^{OJ,rew}[1] = 0$ and therefore,

$$B_1[2] = A_{o1} + K_{11}P_1^{CD}[1] + \mu_{11}R_o + \mu_{32}\left(A_{o2} + \mu_{31}A_{o1}\right).$$

By substituting in the first expression for R_o given in (5.20) into the above expression, the following relationship is obtained

$$\begin{aligned} B_1[2] &\geq A_{o1} + K_{11}P_1^{CD}[1] - A_{o1}\left(1 + \mu_{31}\mu_{32}\right) - \mu_{32}A_{o2} + \mu_{32}\left(A_{o2} + \mu_{31}A_{o1}\right) \\ &= K_{11}P_1^{CD}[1]. \end{aligned}$$

From Part 1, $B_1[1] \geq 0$; therefore, $P_1^{CD}[1] \geq 0$. As a result, $B_1[2] \geq 0$ and $P_1^{CD}[2] \geq 0$. Lastly, since $A_{conf,1}[2] = \mu_{31}B_2[1] < 0$, it follows that $P_1^{OJ,rew}[2] = 0$.

To proceed with the base step of the inductive proof, let $k = 3$. Since $P_1^{OJ,rew}[2] = 0$, the $\Delta A_{out,1}^{OJ,rew}[2]$ term can be omitted from (E.1), and thus, $B_1[3]$ is given by

$$B_1[3] = A_{o1} + K_{11}\left(P_1^{CD}[1] + P_1^{CD}[2]\right) + \mu_{11}R_o + \mu_{32}\left(A_{o2} + K_{12}P_2^{CD}[1] + \mu_{31}B_1[1]\right).$$

By substituting the first expression for R_o given in (5.20) into the above expression, the following relationship is obtained

$$\begin{aligned} B_1[3] &\geq A_{o1} + K_{11}\left(P_1^{CD}[1] + P_1^{CD}[2]\right) - A_{o1}\left(1 + \mu_{31}\mu_{32}\right) - \mu_{32}A_{o2} \\ &\quad + \mu_{32}\left(A_{o2} + K_{12}P_2^{CD}[1] + \mu_{31}B_1[1]\right) \\ &= K_{11}\left(P_1^{CD}[1] + P_1^{CD}[2]\right) - \mu_{31}\mu_{32}A_{o1} + \mu_{32}K_{12}P_2^{CD}[1] + \mu_{31}\mu_{32}B_1[1] \end{aligned}$$

From the initial conditions, $A_{o1} < 0$. From Part 1, $B_1[1] \geq 0$ and $P_1^{CD}[1] \geq 0$ and from the previous step, $P_1^{CD}[2] \geq 0$. However, since $P_2^{CD}[1] < 0$, $B_1[1]$ is expanded and the second expression for R_o from (5.20) is substituted into the above inequality, giving

$$
\begin{aligned}
B_1[3] \geq K_{11}\left(P_1^{CD}[1]+P_1^{CD}[2]\right) - \mu_{31}\mu_{32}A_{o1} + \mu_{32}K_{12}P_2^{CD}[1] + \frac{K_{12}A_{o2}}{A_{o2}+\mu_{31}|A_{o1}|} \\
- \left(A_{o2}+\mu_{31}A_{o1}\right)\left(1+\mu_{31}\mu_{32}\right) + \mu_{31}\mu_{32}\left(A_{o1}+\mu_{32}A_{o2}\right) \\
= K_{11}\left(P_1^{CD}[1]+P_1^{CD}[2]\right) + \mu_{32}K_{12}P_2^{CD}[1] - K_{12}P_2^{CD}[1] \\
- \left(A_{o2}+\mu_{31}A_{o1}\right)\left(1+\mu_{31}\mu_{32}\right) + \mu_{31}\mu_{32}^2A_{o2} \\
= K_{11}\left(P_1^{CD}[1]+P_1^{CD}[2]\right) - \mu_{32}K_{12}P_2^{CD}[1]\left(1-\mu_{32}\right) \\
- B_2[1]\left(1+\mu_{31}\mu_{32}\right) + \mu_{31}\mu_{32}^2A_{o2}.
\end{aligned}
$$

Since $A_{o2} \geq 0$, $B_2[1] < 0$, $P_2^{CD}[1] < 0$ and $|\mu_{32}| < 1$, it follows that all terms in the above inequality are non-negative; thus $B_1[3] \geq 0$. Therefore, the reward given in (5.20) is sufficient to drive $B_1[3] \geq 0$.

Additionally, from (E.2), $B_2[3]$ is given by

$$
B_2[3] = A_{o2} + K_{12}\left(P_2^{CD}[1]+P_2^{CD}[2]\right) + \mu_{31}\left(A_{o1}+K_{11}P_1^{CD}[1]+\mu_{11}R_o+\mu_{32}B_2[1]\right).
$$

By substituting the second expression for R_o given in (5.20) into the above equation, the following relationship is obtained

$$
\begin{aligned}
B_2[3] \geq A_{o2} + K_{12}\left(P_2^{CD}[1]+P_2^{CD}[2]\right) + \frac{K_{12}A_{o2}}{A_{o2}+\mu_{31}|A_{o1}|} \\
- A_{o2} - \mu_{31}\left(A_{o1}+\mu_{32}A_{o2}+\mu_{31}\mu_{32}A_{o1}\right) + \mu_{31}\left(A_{o1}+K_{11}P_1^{CD}[1]+\mu_{32}B_2[1]\right) \\
= K_{12}P_2^{CD}[2] - \mu_{31}\mu_{32}\left(A_{o2}+\mu_{31}A_{o1}\right) + \mu_{31}\left(K_{11}P_1^{CD}[1]+\mu_{32}B_2[1]\right).
\end{aligned}
$$

There are two negative terms in the above inequality, $-A_{o2}$ and $B_2[1]$. However, since $A_{o2} + \mu_{31}A_{o1} = B_2[1]$, the $\mu_{31}\mu_{32}B_2[1]$ term is cancelled out by the $\mu_{31}\mu_{32}\left(A_{o2}+\mu_{31}A_{o1}\right)$ term. Consequently, $B_2[3] \geq 0$ and thus, $P_2^{CD}[3] \geq 0$.

More generally, suppose $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $3 \leq k < \bar{k}$. Therefore, since $A_{out,1}[k] < 0$ for all $k \geq 0$, person one does not experience overjustification pressure for $3 \leq k < \bar{k}$, i.e., $P_1^{OJ,rew}[k] = 0$, and hence, $\Delta A_{out,1}^{OJ,rew}[k] = 0$ for $0 \leq k < \bar{k}$. Additionally, since each behaviour is positive for $3 \leq k < \bar{k}$, it follows that $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$ for $3 \leq k < \bar{k}$. Now, from (E.1), $B_1[\bar{k}]$ can be given by

$$
B_1[\bar{k}] = A_{o1} + \sum_{i=1}^{\bar{k}-1} K_{11}P_1^{CD}[i] + \mu_{11}R_o + \mu_{32}B_2[\bar{k}-1].
$$

By substituting in the first expression for R_o given in (5.20), the following relationship is obtained

$$
\begin{aligned}
B_1[\bar{k}] &\geq A_{o1} + \sum_{i=1}^{\bar{k}-1} K_{11} P_1^{CD}[i] - A_{o1}\left(1 + \mu_{31}\mu_{32}\right) - \mu_{32} A_{o2} + \mu_{32} B_2[\bar{k}-1] \\
&= \sum_{i=1}^{\bar{k}-1} K_{11} P_1^{CD}[i] - \mu_{32} B_2[1] + \mu_{32} B_2[\bar{k}-1]
\end{aligned}
$$

From the initial conditions, $B_2[1] < 0$. From Part 1, the base step, and the inductive hypothesis, $P_1^{CD}[k] \geq 0$ for $1 \leq k \leq \bar{k}$. Also, from the inductive hypothesis, $B_2[\bar{k}-1] \geq 0$, and hence, all of the terms in the above expression are non-negative. Thus $B_1[\bar{k}] \geq 0$.

Finally, from (E.2), $B_2[\bar{k}]$ is given by

$$
B_2[\bar{k}] = A_{o2} + \sum_{i=1}^{\bar{k}-1} K_{12} P_2^{CD}[i] + \mu_{31}\left(A_{o1} + \sum_{i=1}^{\bar{k}-2} K_{11} P_1^{CD}[i-2] + \mu_{11} R_o + \mu_{32} B_2[\bar{k}-2]\right).
$$

By substituting the second expression for R_o given in (5.20) into the above equations and expanding the expression for $P_2^{CD}[1]$, the following relationship is obtained

$$
\begin{aligned}
B_2[\bar{k}] &\geq A_{o2} + \sum_{i=1}^{\bar{k}-1} K_{12} P_2^{CD}[i] + \frac{K_{12} A_{o2}}{\left(A_{o2} + \mu_{31}|A_{o1}|\right)} - \left(A_{o2} + \mu_{31} A_{o1}\right)\left(1 + \mu_{31}\mu_{32}\right) \\
&\quad + \mu_{31}\left(A_{o1} + \sum_{i=1}^{\bar{k}-2} K_{11} P_1^{CD}[i-2] + \mu_{32} B_2[\bar{k}-2]\right) \\
&= \sum_{i=2}^{\bar{k}-1} K_{12} P_2^{CD}[i] - \mu_{31}\mu_{32} B_2[1] + \mu_{31}\left(\sum_{i=1}^{\bar{k}-2} K_{11} P_1^{CD}[i-2] + \mu_{32} B_2[\bar{k}-2]\right).
\end{aligned}
$$

Since $B_2[1] < 0$, it follows that $-B_2[1] \geq 0$. From Part 1, the base step, and the inductive hypothesis, $P_1^{CD}[k] \geq 0$ for $1 \leq k < \bar{k}$ and $P_2^{CD}[k] \geq 0$ for $2 \leq k < \bar{k}$. Finally, from the inductive hypothesis, $B_2[\bar{k}-2] \geq 0$; thus, all terms in the above inequality are non-negative, and $B_2[\bar{k}] \geq 0$.

Therefore, by induction, the reward given in (5.20) is sufficient to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[2] \geq 0$ for $k \geq 2$.

Part 3

Finally, the case when person one experiences overjustification pressure is considered; thus, suppose $A_{out,1}[k] < 0$ for $0 \leq k < \hat{k}$, and $A_{out,1}[k] \geq 0$ at $k = \hat{k}$. For $0 \leq k < \hat{k}$, the results of Part 2 apply. Therefore, it is enough to show that the reward given in (5.20) is sufficient to drive $B_1[k] \geq 0$ and $B_2[k] \geq 0$ for $k \geq \hat{k}$. Given that it has already been shown that $B_1[1] \geq 0$, $B_1[2] \geq 0$ and $B_2[2] \geq 0$, it is assumed that $\hat{k} > 2$.

Before proceeding with the inductive proof, once again, consider $B_2[\hat{k}]$. From (E.2), $B_2[\hat{k}]$ is given by

$$
\begin{aligned}
B_2[\hat{k}] = A_{o2} + \sum_{i=1}^{\hat{k}-1} K_{12} P_2^{CD}[i] + \mu_{31} \Bigg(A_{o1} + \sum_{i=1}^{\hat{k}-2} K_{11} P_1^{CD}[i-2] \\
+ \sum_{i=1}^{\hat{k}-2} \Delta A_{out,1}^{OJ,rew}[i] + \mu_{11} R_o + \mu_{32} B_2[\hat{k}-2] \Bigg).
\end{aligned}
$$

No overjustification pressure is experienced for $0 \leq k < \hat{k}$ because $A_{out,1}[k] < 0$ over the given range. Additionally, substituting the second expression for R_o from (5.20) into the above expression gives

$$
\begin{aligned}
B_2[\hat{k}] &\geq A_{o2} + \sum_{i=1}^{\hat{k}-1} K_{12} P_2^{CD}[i] + \frac{K_{12} A_{o2}}{(A_{o2} + \mu_{31}|A_{o1}|)} - (A_{o2} + \mu_{31} A_{o1})(1 + \mu_{31}\mu_{32}) \\
&\quad + \mu_{31} \left(A_{o1} + \sum_{i=1}^{\hat{k}-2} K_{12} P_1^{CD}[i-2] + \mu_{32} B_2[\hat{k}-2] \right) \\
&= \sum_{i=2}^{\hat{k}-1} K_{12} P_2^{CD}[i] - \mu_{31}\mu_{32} B_2[1] \\
&\quad + \mu_{31} \left(\sum_{i=1}^{\hat{k}-2} K_{11} P_1^{CD}[i-2] + \mu_{32} B_2[\hat{k}-2] \right)
\end{aligned}
$$

From Parts 1 and 2, $P_1^{CD}[k] \geq 0$ for $1 \leq k < \hat{k}$, $P_2^{CD}[k] \geq 0$ for $2 \leq k < \hat{k}$, $B_2[1] < 0$ and $B_2[\hat{k}-2] \geq 0$. Therefore, all of the terms in the above expression are non-negative, and therefore, $B_2[\hat{k}] \geq 0$.

More generally, suppose for $\hat{k} \leq k < \bar{k}$, $B_1[k] \geq 0$ and $B_2[k] \geq 0$. Consequently, $P_1^{CD}[k] \geq 0$ and $P_2^{CD}[k] \geq 0$ over this sample range. Moreover, since $A_{out,1}[\hat{k}] \geq 0$, $A_{conf,1}[\hat{k}] \geq 0$ and $A_{rew,1}[\hat{k}] \geq 0$, from Lemma 5.10, $A_{out,1}[\hat{k}+1] \geq 0$. From the inductive hypothesis, this implies that for $\hat{k} \leq k < \bar{k}$, $A_{out,1}[k] \geq 0$. Now, from (4.24), $B_1[\bar{k}]$ is given by

$$
B_1[\bar{k}] = A_{out,1}[\bar{k}-1] + K_{11} P_1^{CD}[\bar{k}-1] + \Delta A_{out,1}^{OJ,rew}[\bar{k}-1] + \mu_{11} R_o + \mu_{32} B_2[\bar{k}-1].
$$

From the inductive hypothesis, $A_{out,1}[\bar{k}-1] \geq 0$, $B_2[\bar{k}-1] \geq 0$ and $P_1^{CD}[\bar{k}-1] \geq 0$. Moreover, from Lemma 5.10, $A_{out,1}[\bar{k}-1] + \Delta A_{out,1}^{OJ,rew}[\bar{k}-1] \geq 0$. Thus, $B_1[\bar{k}] \geq 0$.

Moreover, from (4.24), $B_2[\bar{k}]$ is given by $B_2[k] = A_{out,2}[\bar{k}] + \mu_{31} B_1[\bar{k}-1]$. Expanding $A_{out,2}[\bar{k}]$ and using the second expression for R_o in (5.20) gives

$$\begin{aligned} B_2[\bar{k}] \geq A_{o2} &+ \sum_{i=1}^{\bar{k}-1} K_{12} P_2^{CD}[i] \\ &+ \frac{K_{12} A_{o2}}{A_{o2} + \mu_{31}|A_{o1}|} - (A_{o2} + \mu_{31} A_{o1})(1 + \mu_{31}\mu_{32}) + \mu_{31} B_1[\bar{k}-1] \\ &= \sum_{i=2}^{\bar{k}-1} K_{12} P_2^{CD}[i] - \mu_{31}(A_{o1} + \mu_{32} B_2[1]) + \mu_{31} B_1[\bar{k}-1]. \end{aligned}$$

From the initial conditions, $A_{o1} < 0$ and $B_2[1] < 0$. From Parts 1 and 2, $P_2^{CD}[k] \geq 0$ for $2 \leq k < \hat{k}$. From the base step and the inductive hypothesis, $P_2^{CD}[k] \geq 0$ for $\hat{k} \leq k < \bar{k}$. Finally, from the inductive hypothesis, $B_1[\bar{k}-1] \geq 0$; thus, all of the terms in the above inequality are non-negative, and $B_2[\bar{k}] \geq 0$.

Therefore, by induction, in the case when $A_{out,1}[k] \geq 0$ at some $k = \hat{k}$, the reward given in (5.20) is sufficient to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$.

Together, the results from Parts 1, 2 and 3 show that if the two-person begins in region VI, then there exists a sufficient reward to drive $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. □

E.11 Proof of Lemma 5.13

Proof. The proof is done to two parts, each of which focuses on one of the two signals.

Part 1

Part 1 examines $A_{out,1}[k]$ as k tends to infinity. Two cases must be considered: $A_{out,1}[k] \leq 0$ for all k, and $A_{out,1}[k] \leq 0$ for $0 \leq k < \hat{k}$, while $A_{out,1}[k] > 0$ at $k = \hat{k}$. The former case is discussed first.

Suppose $A_{out,1}[k] \leq 0$ for all k. Then, from (4.1), $A_{out,1}[k]$ can be expressed as

$$A_{out,1}[k] = A_{o1} + \sum_{i=0}^{k-1} K_{11} P_1^{CD}[i]. \tag{E.3}$$

Since $B_1[k] \geq 0$ for $k \geq 1$, it follows that $P_1^{CD}[k] \geq 0$ for $k \geq 1$ and hence, $A_{out,1}[k]$ is a non-decreasing function of k for $k \geq 1$. Additionally, owing to the restriction that $A_{out,1}[k]$

is upper bounded by zero, Lemma A.2 can be applied. Therefore, $A_{out,1}[k]$ converges to some $c_1 \leq 0$ as k tends to infinity.

Suppose instead, that at some $k = \hat{k}$, $A_{out,1}[k]$ becomes positive, i.e., $A_{out,1}[k] \leq 0$ for $0 \leq k < \hat{k}$, and $A_{out,1}[k] > 0$ at $k = \hat{k}$. Since $B_2[k] \geq 0$ for $k \geq 2$, it follows that $A_{conf,1}[k] \geq 0$ for $k \geq 3$. Moreover, since $R[k] \geq 0$ for $k \geq 0$, it follows that $A_{rew,1}[k] \geq 0$ for $k \geq 0$. Therefore, the conditions of Lemma 5.10 hold, and, as a result, $A_{out,1}[k]$ is a non-increasing function of k for $k \geq \hat{k}$, lower bounded by zero. As a result, from Lemma A.1, $A_{out,1}[k]$ converges to some $c_1 \geq 0$ as k tends to infinity.

Now, it is left to show that if $A_{out,1}[k]$ converges as k tends to infinity, then $A_{out,1}[k]$ necessarily tends to zero as k tends to infinity. Recall that $A_{out,1}[k]$ can be expressed as

$$A_{out,1}[k] = A_{out,1}[\hat{k}] + \sum_{i=\hat{k}}^{k-1} \Delta A_{out,1}^{OJ,rew}[i], \tag{E.4}$$

for $k > \hat{k}$ (because for $k \geq \hat{k}$, $P_1^{CD}[k] = 0$). Since the left-hand side of (E.4) is well-behaved as k tends to infinity, the right-hand side is also well-behaved. Thus, limits can be taken on both sides of (E.4) giving

$$\begin{aligned}
\lim_{k\to\infty} A_{out,1}[k] &= \lim_{k\to\infty} \left(A_{out,1}[\hat{k}] + \sum_{i=\hat{k}}^{k-1} \Delta A_{out,1}^{OJ,rew}[i] \right) \\
\Rightarrow c_1 &= \lim_{k\to\infty} \left(A_{out,1}[\hat{k}] \right) + \lim_{k\to\infty} \left(\sum_{i=\hat{k}}^{k-1} \Delta A_{out,1}^{OJ,rew}[i] \right) \\
\Rightarrow c_1 &= A_{out,1}[\hat{k}] + \lim_{k\to\infty} \sum_{i=\hat{k}}^{k-1} \Delta A_{out,1}^{OJ,rew}[i] \\
\Rightarrow c_1 - A_{out,1}[\hat{k}] &= \lim_{k\to\infty} \sum_{i=\hat{k}}^{k-1} \Delta A_{out,1}^{OJ,rew}[i].
\end{aligned}$$

Since c_1, and $A_{out,1}[\hat{k}]$ are constants, the above limit converges as k tends to infinity. Therefore,

$$\Delta A_{out,1}^{OJ,rew}[k] \to 0 \text{ as } k \to \infty. \tag{E.5}$$

It remains to show that (E.5) implies $A_{out,1}[k]$ tends to zero as k tends to infinity. Since, $A_{out,1}[\hat{k}] > 0$, $A_{rew,1}[\hat{k}] > 0$, and $A_{conf,1}[\hat{k}] \geq 0$, it follows that $B_1[\hat{k}] > 0$ and $P_1^{OJ,rew}[\hat{k}] > 0$. From (4.20), note that if $P_1^{OJ,rew}[\hat{k}] > 0$, then $P_1^{OJ,rew}[k] > 0$ for $k \geq \hat{k}$. To see this, note that Lemma 5.10 implies $A_{out,1}[\hat{k}+1] \geq 0$. From (4.18), $P_1^{OJ,rew}[k]$ is given by

$$P_1^{OJ,rew}[k] = \begin{cases} r_{51} P_1^{OJ,rew}[k-1] & \text{if } A_{out,1}[k], B_1[k], \\ \quad + (1-r_{51})\, A_{out,1}[k] A_{rew,1}[k] & A_{conf,1}[k], \text{ and } A_{rew,1}[k] > 0 \\ r_{51} P_1^{OJ,rew}[k-1] & \text{otherwise,} \end{cases}$$

for $k \geq 0$. Because $P_1^{OJ,rew}[\hat{k}] \geq 0$, $A_{out,1}[\hat{k}+1] \geq 0$, $A_{rew,1}[\hat{k}+1] \geq 0$ and $0 \leq r_{51} < 1$, it follows that $P_1^{OJ,rew}[\hat{k}+1] > 0$ for either of the two possible expressions of $P_1^{OJ,rew}[\hat{k}+1]$. Therefore, if $P_1^{OJ,rew}[\hat{k}] > 0$ then $P_1^{OJ,rew}[k] > 0$ for $k \geq \hat{k}$. In this case, from (4.22), $\Delta A_{out,1}^{OJ,rew}[k]$ is simply

$$\Delta A_{out,1}^{OJ,rew}[k] = \begin{cases} -K_{21} P_1^{OJ,rew}[k] & \text{if } P_1^{OJ,rew}[k] > 0, \\ & \text{and } K_{21} P_1^{OJ,rew}[k] \leq A_{out,1}[k], \\ -A_{out,1}[k] & \text{otherwise} \end{cases} \tag{E.6}$$

for $k \geq \hat{k}$. Hence, it remains to show that (E.5) and (E.6) together imply that $A_{out,1}[k]$ tends to zero as k tends to infinity.

From (E.6), there are only three admissible combinations for $\Delta A_{out,1}^{OJ,rew}[k]$:

- For $k \geq \hat{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -K_{21} P_1^{OJ,rew}[k]$,
- For $\hat{k} \leq k < \bar{k}$, $\Delta A_{out,1}^{OJ}[k] = -K_{21} P_1^{OJ,rew}[k]$, and for $k \geq \bar{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -A_{out,1}[k]$, and
- For $k \geq \hat{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -A_{out,1}[k]$.

Each case is considered below.

Suppose for all $k \geq \hat{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -K_{21} P_1^{OJ,rew}[k]$. Then, from (E.5)

$$\begin{aligned} &\Delta A_{out,1}^{OJ,rew}[k] \to 0 \text{ as } k \to \infty \\ \Rightarrow &P_1^{OJ,rew}[k] \to 0 \text{ as } k \to \infty \end{aligned}$$

Since for $k \geq \hat{k}$, $A_{out,1}[k] \geq 0$, $B_1[k] \geq 0$, $A_{rew,1}[k] \geq 0$ and $A_{conf,1}[k] \geq 0$, $P_1^{OJ,rew}[k]$ is given by

$$P_1^{OJ,rew}[k] = r_{51} P_1^{OJ,rew}[k-1] + (1-r_{51}) A_{out,1}[k] A_{rew,1}[k], \tag{E.7}$$

for $k \geq \hat{k}$. Since the left-hand side of (E.7) is well-behaved as k tends to infinity, the right-hand side of (E.7) is well-behaved as k tends to infinity. Therefore, limits can be taken on of both sides of (E.7), giving

$$\begin{aligned}&\lim_{k\to\infty} P_1^{OJ,rew}[k] = \lim_{k\to\infty}\Big(r_5 P_1^{OJ,rew}[k-1] + (1-r_5)\, A_{out,1}[k] A_{rew,1}[k]\Big)\\ \Rightarrow & 0 = 0 + (1-r_{51})\,\mu_{11} R_o \lim_{k\to\infty} A_{out,1}[k]\\ \Rightarrow & A_{out,1}[k] \to 0 \text{ as } k\to\infty.\end{aligned}$$

Hence, if, for $k \geq \hat{k}$, $P_1^{OJ,rew}[k] > 0$ and $K_{21}P^{OJ,rew}[k] \leq A_{out,1}[k]$, then $A_{out,1}[k]$ necessarily tends to zero as k tends to infinity.

Suppose instead that for $\hat{k} \leq k < \bar{k}$, $\Delta A_{out,1}[k] = -K_{21}P_1^{OJ,rew}[k]$ and for $k \geq \bar{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -A_{out,1}[k]$. Since we are interesting in $A_{out,1}[k]$ as k tends to infinity, $k = \bar{k}$ is used as the starting point. Thus, from (E.5),

$$\begin{aligned}&\Delta A_{out,1}^{OJ,rew}[k] \to 0 \text{ as } k\to\infty\\ \Rightarrow & A_{out,1}[k] \to 0 \text{ as } k\to\infty.\end{aligned}$$

Hence, if, for $\hat{k} \leq k < \bar{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -K_{21}P_1^{OJ,rew}[k]$, and, for $k \geq \bar{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -A_{out,1}[k]$, then $A_{out,1}[k]$ necessarily tends to zero as k tends to infinity.

Suppose instead for all $k \geq \hat{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -A_{out,1}[k]$. Then, the same steps as the previous case apply (with $\hat{k} = \bar{k}$), and therefore, $A_{out,1}[k]$ necessarily tends to zero as k tends to infinity.

Since $A_{out,1}[k] \to 0$ as $k \to \infty$ for any admissible combination of $\Delta A_{out,1}^{OJ,rew}[k]$,

$$\Delta A_{out,1}^{OJ,rew}[k] \to 0 \text{ as } k\to\infty \Rightarrow A_{out,1}[k] \to 0 \text{ as } k\to\infty.$$

In summary, $A_{out,1}[k]$ converges as k tends to infinity. Furthermore, if $A_{out,1}[k] > 0$ at $k = \hat{k}$, then $A_{out,1}[k]$ tends to zero as k tends to infinity.

Part 2

Now consider $A_{out,2}[k]$ as k tends to infinity. Similar to $A_{out,1}[k]$, two cases must be considered: $A_{out,2}[k] \leq 0$ for all $k \geq 0$, and $A_{out,2}[k] \leq 0$ for $0 \leq k < \tilde{k}$, while $A_{out,2}[k] > 0$ at $k = \tilde{k}$. The former case is discussed first.

From (4.1), $A_{out2}[k]$ can be expressed as

$$A_{out,2}[k] = A_{o2} + \sum_{i=0}^{k-1} K_{12}P_2^{CD}[i]. \tag{E.8}$$

Since $B_2[k] \geq 0$ for $k \geq 2$, it follows that $P_2^{CD}[k] \geq 0$ for $k \geq 2$ and hence, $A_{out,2}[k]$ is a non-decreasing function of k for $k \geq 2$. Additionally, owing to the restriction that $A_{out,2}[k]$

is upper bounded by zero, Lemma A.2 can be applied. Therefore, $A_{out,2}[k]$ converges to some $c_2 \leq 0$ as k tends to infinity.

Suppose instead, that at some $k = \tilde{k}$, $A_{out,2}[k]$ becomes positive. Then, for $k > \tilde{k}$, $A_{out,2}[k]$ can be expressed as

$$A_{out,2}[k] = A_{out,2}[\tilde{k}] + \sum_{i=\tilde{k}}^{k-1} K_{12} P_2^{CD}[i].$$

Moreover, since $B_2[\tilde{k}] \geq 0$ and $A_{conf,2}[\tilde{k}] \geq 0$, person two does not experience dissonance pressure at $k = \tilde{k}$, and thus, $P_2^{CD}[\tilde{k}] = 0$. This conclusion implies $A_{out,2}[\tilde{k}+1] = A_{out,2}[\tilde{k}]$. This argument can be applied for each $k \geq \tilde{k}$ and thus, for $k \geq \tilde{k}$, $A_{out,2}[k] = A_{out,2}[\tilde{k}]$. Thus, if $A_{out,2}[k]$ becomes positive at some $k = \tilde{k}$, then $A_{out,2}[k]$ converges to $A_{out,2}[\tilde{k}]$ as k tends to infinity. □

E.12 Proof of Lemma 5.14

Proof. $B_1[k]$ and $B_2[k]$ are considered in Part 1 and Part 2 respectively.

Part 1

Showing $B_1[\cdot]$ is bounded is done by examining a particular form of the expression for $B_1[k]$ and solving for $B_1[k]$ from initial conditions and parameter values. Then, the solution is used to show $B_1[\cdot]$ is bounded.

Consider the following form of the expression for $B_1[k]$ for $k \geq 2$, which is obtained by expanding the $A_{conf,1}[k]$ term:

$$B_1[k] = A_{out,1}[k] + \mu_{11} R_o + \mu_{32} A_{out,2}[k-1] + \mu_{31}\mu_{32} B_1[k-2].$$

Define

$$\alpha[k] := A_{out,1}[k] + \mu_{11} R_o + \mu_{32} A_{out,2}[k-1].$$

Then, for $k \geq 2$,

$$B_1[k] = \alpha[k] + \mu_{31}\mu_{32} B_1[k-2]. \tag{E.9}$$

From Lemma 5.13, $A_{out,1}[k]$ and $A_{out,2}[k]$ converge as k tends to infinity. Therefore, all signals in the expression for $\alpha[k]$ are bounded, and as a result, $\alpha[k]$ is bounded. Define M_L and M_U as the lower and upper bounds on $\alpha[k]$, i.e., for all $k \geq 2$,

$$M_L \leq \alpha[k] \leq M_U. \tag{E.10}$$

Furthermore, using (E.9) as the expression for $B_1[k]$, we can solve for $B_1[k]$ separately for even and odd values of k, giving

$$B_1[2n] = \sum_{i=0}^{n-1} (\mu_{31}\mu_{32})^i \alpha [2(n-i)] + (\mu_{31}\mu_{32})^n B_1[0] \tag{E.11}$$

$$B_1[2n+1] = \sum_{i=0}^{n-1} (\mu_{31}\mu_{32})^i \alpha [2(n-i)+1] + (\mu_{31}\mu_{32})^n B_1[1], \tag{E.12}$$

for $n \geq 1$.

Now we examine the expressions for $B_1[k]$ for even and odd values of k separately. First consider even values of k (i.e., $k = 2n$ for $n \geq 1$). From (E.10) and (E.11),

$$\begin{aligned} B_1[2n] &\leq M_U \sum_{i=0}^{n-1} (\mu_{31}\mu_{32})^i + (\mu_{31}\mu_{32})^n B_1[0] \\ \Rightarrow B_1[2n] &\leq \frac{M_U (1 - (\mu_{31}\mu_{32})^n)}{1 - \mu_{31}\mu_{32}} + (\mu_{31}\mu_{32})^n B_1[0]. \end{aligned}$$

Since $|\mu_{31}\mu_{32}| < 1$, $B_1[2n]$ is bounded from above.

Additionally, from (E.10) and (E.11),

$$\begin{aligned} B_1[2n] &\geq M_L \sum_{i=0}^{n-1} (\mu_{31}\mu_{32})^i + (\mu_{31}\mu_{32})^n B_1[0] \\ \Rightarrow B_1[2n] &\geq \frac{M_L (1 - (\mu_{31}\mu_{32})^n)}{1 - \mu_{31}\mu_{32}} + (\mu_{31}\mu_{32})^n B_1[0]. \end{aligned}$$

Again, since $|\mu_{31}\mu_{32}| < 1$, $B_1[2n]$ is bounded from below. Therefore, for even values of k, $B_1[k]$ is bounded. Showing $B_1[k]$ is bounded for odd values of k uses the same steps as in the case of even values. Hence, $B_1[\cdot]$ is bounded.

Part 2

Now consider $B_2[k]$, which is given by $B_2[k] = A_{out,2}[k] + \mu_{31} B_1[k-1]$, for $k \geq 1$. From Lemma 5.13, $A_{out,2}[k]$ converges as k tends to infinity. Moreover, Part 1 showed that $B_1[\cdot]$ is bounded. Therefore, both terms in the expression for $B_2[k]$ are bounded; thus, $B_2[\cdot]$ is also bounded. □

E.13 Proof of Lemma 5.15

Proof. From Lemma 5.13, $A_{out,1}[k]$ does indeed converge as k tends to infinity and, if $A_{out,1}[k]$ becomes positive, then it converges to zero as k tends to infinity. Thus, it only remains to show that if $A_{out,1}[k] \leq 0$ for all $k \geq 0$, then $A_{out,1}[k]$ tends to zero as k tends to infinity.

Suppose $A_{out,1}[k] \leq 0$ for all $k \geq 0$. Then, from Lemma 5.13, $A_{out,1}[k]$ converges to some $c_1 \leq 0$ as k tends to infinity. Using (E.3) as the expression for $A_{out,1}[k]$, it is clear that the left-hand side is well-behaved as k tends to infinity and thus, so is the right-hand side. Therefore, limits can be taken on both sides, giving

$$\lim_{k\to\infty} A_{out,1}[k] = \lim_{k\to\infty} \left(A_{o1} + \sum_{i=0}^{k-1} K_{11} P_1^{CD}[i] \right)$$
$$\Rightarrow c_1 - A_{o1} = \lim_{k\to\infty} \sum_{i=0}^{k-1} K_{11} P_1^{CD}[i].$$

Since c_1 and A_{o1} are constants, the right-hand side of the equation converges to some constant. Consequently, $P_1^{CD}[k]$ tends to zero as k tends to infinity.

Since $B_1[k] \geq 0$, $A_{out,1}[k] < 0$ and $A_{conf,1}[k] \geq 0$ for $k \geq 3$, the expression for $P_1^{CD}[k]$ for all $k \geq 3$ is given by

$$P_1^{CD}[k] = \frac{|A_{out,1}[k]|}{|A_{out,1}[k]| + A_{rew,1}[k] + A_{conf,1}[k]}$$

Taking the limit of both sides gives

$$\lim_{k\to\infty} P_1^{CD}[k] = \lim_{k\to\infty} \frac{|A_{out,1}[k]|}{|A_{out,1}[k]| + \mu_{11} R_o + \mu_{32} B_2[k-1]}$$
$$\Rightarrow 0 = \lim_{k\to\infty} \frac{1}{1 + \frac{\mu_{11} R_o + \mu_{32} B_2[k-1]}{|A_{out,1}[k]|}}.$$

Thus, as k tends to infinity, either $B_2[k] \to \infty$, $A_{out,1}[k] \to 0$, or both. However, Lemma 5.14 states that $B_2[\cdot]$ is bounded and, hence, the first of the two situations is impossible. Therefore, if $A_{out,1}[k] < 0$ for all $k \geq 0$, then $A_{out,1}[k]$ tends to zero as k tends to infinity. □

E.14 Proof of Lemma 5.16

Proof. From Lemma 5.13, $A_{out,2}[k]$ converges as k tends to infinity. If $A_{out,2}[k] > 0$ at some $k = \tilde{k}$, then $A_{out,2}[k]$ converges to $A_{out,2}[\tilde{k}]$ as k tends to infinity. Thus, it only remains to show that if $A_{out,2}[k] \leq 0$ for all $k \geq 2$, then $A_{out,2}[k]$ tends to zero as k tends to infinity.

Suppose then, that $A_{out,2}[k] \leq 0$ for all $k \geq 2$ (chosen because if the system begins in region VI, then it is possible for $A_{out,2}[k] \geq 0$ at $k = 0, 1$ but $A_{out,2}[k] < 0$ for $k \geq 2$).

Since $B_2[k] \geq 0$ for $k \geq 2$, it follows that $P_2^{CD}[k] \geq 0$ for $k \geq 2$ and, thus, $A_{out,2}[k]$ is a non-decreasing function of k for $k \geq 2$ with an upper bound of zero. Consequently, by Lemma A.2, $A_{out,2}[k]$ converges to some $c_2 \leq 0$ as k tends to infinity. Thus, the left-hand side of (E.8) is well-behaved as k tends to infinity. Therefore, the right-hand side of (E.8) is also well-behaved as k tends to infinity and thus, limits can be taken on both sides, giving

$$\begin{aligned}\lim_{k\to\infty} A_{out,2}[k] &= \lim_{k\to\infty}\left(A_{out,2}[2] + \sum_{i=2}^{k-1} K_{12}P_2^{CD}[i]\right)\\ \Rightarrow c_2 - A_{out,2}[2] &= \lim_{k\to\infty}\sum_{i=2}^{k-1} K_{12}P_2^{CD}[i].\end{aligned}$$

Since c_2 and $A_{out,2}[2]$ are constants, the right-hand side of the equation converges to some constant. Therefore, $P_2^{CD}[k]$ tends to zero as k tends to infinity.

Since $B_2[k] \geq 0$, $A_{out,2}[k] \leq 0$ and $A_{conf,2}[k] = \mu_{31}B_1[k-1] \geq 0$ for $k \geq 2$, the expression for $P_2^{CD}[k]$ for all $k \geq 2$ is given by

$$P_2^{CD}[k] = \frac{|A_{out,2}[k]|}{|A_{out,2}[k]| + A_{conf,2}[k]}.$$

Taking the limit of both sides gives

$$\begin{aligned}\lim_{k\to\infty} P_2^{CD}[k] &= \lim_{k\to\infty}\frac{|A_{out,2}[k]|}{|A_{out,2}[k]| + A_{conf,2}[k]}\\ \Rightarrow 0 &= \lim_{k\to\infty}\frac{1}{1 + \frac{A_{conf,2}[k]}{|A_{out,2}[k]|}}.\end{aligned}$$

Thus, as k tends to infinity, either $A_{conf,2}[k] \to \infty$, $A_{out,2}[k] \to 0$, or both. However, from Lemma 5.14, $B_1[\cdot]$ is bounded and, thus, so is $A_{conf,2}[\cdot]$. Therefore, the first situation is impossible and consequently, if $A_{out,2}[k] < 0$ for all $k \geq 0$, then $A_{out,2}[k]$ tends to zero as k tends to infinity. □

E.15 Proof of Lemma 5.17

Proof. First, $B_1[k]$ is shown to converge as k tends to infinity. Then, the expression for $B_1[\infty]$ is determined. To shown $B_1[k]$ converges to a steady-state value, recall that, for $k \geq 2$, $B_1[k]$ can be expressed as $B_1[k] = \alpha[k] + \mu_{31}\mu_{32}B_1[k-2]$, where $\alpha[k] = A_{out,1}[k] + \mu_{11}R_o + \mu_{32}A_{out,2}[k-1]$. Lemma 5.14 states $B_1[\cdot]$ has an upper bound. If there exists some $\bar{k}$ such that $\alpha[k] \geq 0$ for $k \geq \bar{k}$, then $B_1[k]$ is an increasing function for even values of k

and for odd values of k for $k \geq \bar{k}$. As a result, $B_1[k]$ converges for even values of k and for odd values of k. Thus, to show $B_1[k]$ converges as k tends to infinity, we show that there exists a $\bar{k}$ such that for all $k \geq \bar{k}$, $\alpha[k] \geq 0$.

The expression for R_o in $\alpha[k]$ depends on the initial conditions and, accordingly, uses the specific expressions for R_o given in Lemmas 5.11 and 5.12. Moreover, the signs of $A_{out,1}[k]$ and $A_{out,2}[k]$ depend on the initial conditions of the system. Therefore, the sign of $\alpha[k]$ is determined for each initial region V, VI and VII and VIII.

Suppose the system begins in region V, VII or VIII. Then, from Lemma 5.11, the reward sufficient to ensure $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ is given by (5.18). Note that $A_{out,1}[k]$ and $A_{out,2}[k]$ in $\alpha[k]$ can be expanded, and thus, for $k \geq 2$,

$$\alpha[k] = A_{o1} + K_{11} \sum_{i=0}^{k-1} P_1^{CD}[i] + \mu_{11} R_o + \mu_{32} \left(A_{o2} + K_{12} \sum_{i=0}^{k-2} P_2^{CD}[k] \right).$$

Using the first expression for R_o from (5.18) in the above equation gives

$$\begin{aligned}
\alpha[k] &\geq A_{o1} + K_{11} \sum_{i=0}^{k-1} P_1^{CD}[i] - A_{o1}\left(1 + \mu_{31}\mu_{32}\right) - \mu_{32} A_{o2} + \mu_{32} \left(A_{o2} + K_{12} \sum_{i=0}^{k-2} P_2^{CD}[k] \right) \\
&= \sum_{i=0}^{k-1} K_{11} P_1^{CD}[i] - \mu_{31}\mu_{32} A_{o1} + \mu_{32} \left(K_{12} \sum_{i=0}^{k-2} P_2^{CD}[k] \right).
\end{aligned}$$

From the initial conditions, $P_1^{CD}[0] = P_2^{CD}[0] = 0$ and $A_{o1} < 0$. By Lemma 5.11, $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, implying $P_1^{CD}[k] \geq 0$ for $k \geq 1$ and $P_2^{CD}[k] \geq 0$ for $k \geq 2$. It has previously been shown that if the system begins in region V, VII or VIII, $P_2^{CD}[1] \geq 0$. Hence, for $k \geq 1$, $\alpha[k] \geq 0$.

Suppose instead that the system begins in region VI. Then, from Lemma 5.12, the reward sufficient to ensure $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ is given by (5.20). Recall that in initial region VI, $B_2[1] = A_{o2} + \mu_{31} A_{o1} < 0$. Expanding the expressions for $A_{out,1}[k]$ and $A_{out,2}[k]$ in $\alpha[k]$, and using the second expression in (5.20) for R_o in the expression for $\alpha[k]$, gives

$$\begin{aligned}
\alpha[k] &\geq \sum_{i=0}^{k-1} K_{11} P_1^{CD}[i] + K_{12} P_2^{CD}[1] \left(\mu_{32} - \frac{1}{\mu_{31}} \right) + A_{o2} \left(\mu_{32} - \frac{1}{\mu_{31}} \right) \\
&\quad - \mu_{32} B_2[1] + \mu_{32} \sum_{i=2}^{k-2} K_{12} P_2^{CD}[i] \\
&= \sum_{i=0}^{k-1} K_{11} P_1^{CD}[i] + A_{out,2}[2] \left(\mu_{32} - \frac{1}{\mu_{31}} \right) - \mu_{32} B_2[1] + \mu_{32} \sum_{i=2}^{k-2} K_{12} P_2^{CD}[i].
\end{aligned} \tag{E.13}$$

Given that the sign of $A_{out,2}[2]$ is unknown, both positive and negative values must be checked. First, suppose $A_{out,2}[2] < 0$. Then, given that $|\mu_{31}| < 1$ and $|\mu_{32}| < 1$, it follows that $\left(\mu_{32} - \frac{1}{\mu_{31}}\right) < 0$. Therefore, $A_{out,2}[2]\left(\mu_{32} - \frac{1}{\mu_{31}}\right) > 0$, and, since all the other terms in the expression are non-negative, for $k \geq 1$, $\alpha[k] \geq 0$.

Suppose instead that $A_{out,2}[2] \geq 0$. Since, from (4.1), $A_{out,2}[2] = A_{o2} + K_{12}P_2^{CD}[1] \geq 0$, it follows that $A_{o2} \geq -K_{12}P_2^{CD}[1]$. Substituting this inequality into A_{o2} in (E.13), gives

$$\begin{aligned}\alpha[k] \geq& \sum_{i=0}^{k-1} K_{11}P_1^{CD}[i] + K_{12}P_2^{CD}[1]\left(\mu_{32} - \frac{1}{\mu_{31}}\right) - K_{12}P_2^{CD}[1]\left(\mu_{32} - \frac{1}{\mu_{31}}\right)\\ &- \mu_{32}B_2[1] + \mu_{32}\sum_{i=2}^{k-2} K_{12}P_2^{CD}[i]\\ =& \sum_{i=0}^{k-1} K_{11}P_1^{CD}[i] - \mu_{32}B_2[1] + \mu_{32}\sum_{i=2}^{k-2} K_{12}P_2^{CD}[i].\end{aligned}$$

Since all the other terms in the expression are non-negative, for $k \geq 1$, $\alpha[k] \geq 0$.

Therefore, it has been shown that if the system begins in region V, VI, VII or VIII and a step reward satisfying one of (5.18) or (5.20), depending on the initial region, is applied to the two-person system, then for $k \geq 1$, $\alpha[k] \geq 0$. Therefore, $B_1[k]$ converges as k tends to infinity for odd values of k and for even values of k as previously argued. Now, it only remains to find the steady-state expression for $B_1[k]$.

By expanding $\alpha[k]$, the following expression is obtained for $B_1[k]$:

$$B_1[k] = A_{out,1}[k] + \mu_{11}R_o + \mu_{32}A_{out,2}[k-1] + \mu_{31}\mu_{32}B_1[k-2]. \tag{E.14}$$

Assume k is even, i.e., $k = 2n$. Since $B_1[k]$ converges for even values of k as k tends to infinity, the left-hand side of (E.14) is well-behaved for $k = 2n$ as n tends to infinity and thus, so is the right-hand side. Therefore limits can be taken on both sides giving

$$\begin{aligned}\lim_{n\to\infty} B_1[2n] &= \lim_{n\to\infty}\left(A_{out,1}[2n] + \mu_{11}R_o + \mu_{32}A_{out,2}[2n-1] + \mu_{31}\mu_{32}B_1[2n-2]\right)\\ \Rightarrow \lim_{n\to\infty} B_1[2n] &= 0 + \mu_{11}R_o + \mu_{32}c_2 + \lim_{n\to\infty}\mu_{32}\mu_{31}B_1[2n-2].\end{aligned}$$

Since, it has previously been argued that $B_1[2n]$ converges as n tends to infinity, it follows that

$$\begin{aligned}\lim_{n\to\infty} B_1[2n](1 - \mu_{31}\mu_{32}) &= \mu_{11}R_o + \mu_{32}c_2\\ \Rightarrow \lim_{n\to\infty} B_1[2n] &= \frac{\mu_{11}R_o + \mu_{32}c_2}{1 - \mu_{31}\mu_{32}}.\end{aligned}$$

Suppose instead, k is odd, i.e., $k = 2n + 1$. The same argument can be applied for odd values of k and hence, the expression for $B_1[\infty]$ for odd values of k is the same as the even case, i.e.,

$$\lim_{n\to\infty} B_1[2n+1] = \frac{\mu_{11}R_o + \mu_{32}c_2}{1-\mu_{31}\mu_{32}}.$$

Therefore $B_1[k]$ converges as k tends to infinity. Moreover, the closed-form expression for $B_1[\infty]$ is given by (5.22). □

E.16 Proof of Theorem 5.1

Proof. If one of (5.1), (5.7) or (5.8) hold and $R_o \geq R_o^1$, then by Lemma 5.11, $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. As a result, (5.25) ensures $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ if one of (5.1), (5.7) or (5.8) hold; thus, by Lemma 5.17, $B_1[k]$ converges as k tends to infinity. Specifically,

$$\lim_{k\to\infty} B_1[k] = \frac{\mu_{11}R_o + \mu_{32}c_2}{1-\mu_{31}\mu_{32}},$$

where c_2 is the value to which $A_{out,2}[k]$ converges as k tends to infinity. From Lemma 5.16, c_2 has two possible values: 0 and $A_{out,2}[\tilde{k}] > 0$ (where $\tilde{k}$ is the sample at which $A_{out,2}[k] > 0$). Each value is considered separately below.

Suppose $c_2 = 0$. Then

$$\lim_{k\to\infty} B_1[k] = \frac{\mu_{11}R_o}{1-\mu_{31}\mu_{32}}.$$

To determine whether or not $B_1[\infty] \geq B_d$, two cases must be considered, $R_o^1 < R_o^3$ and $R_o^1 \geq R_o^3$. First, suppose, $R_o^1 < R_o^3$. Then, from (5.25), $R_o \geq R_o^3$ is applied to the two-person system. Since $R_o \geq R_o^3$, it follows that

$$\lim_{k\to\infty} B_1[k] \geq \frac{\mu_{11}R_o^3}{1-\mu_{31}\mu_{32}}.$$

Using the expression for R_o^3 from the right-hand side of (5.24) implies

$$\lim_{k\to\infty} B_1[k] \geq \frac{B_d\,(1-\mu_{31}\mu_{32})}{1-\mu_{31}\mu_{32}} = B_d.$$

In other words, $B_1[\infty] \geq B_d$. On the other hand, suppose $R_o^1 \geq R_o^3$. Then, from (5.25), $R_o \geq R_o^1$ is applied to the two-person system. Since $R_o \geq R_o^1$, it follows that

$$\lim_{k\to\infty} B_1[k] \geq \frac{\mu_{11}R_o^1}{1-\mu_{31}\mu_{32}}.$$

However, since $R_o^1 \geq R_o^3$, the above expression implies

$$\lim_{k\to\infty} B_1[k] \geq \frac{\mu_{11}R_o^3}{1-\mu_{31}\mu_{32}},$$

which, from before, implies $B_1[\infty] \geq B_d$. As a result, for the case when $c_2 = 0$, when the system begins in regions V, VII or VIII and $R_o \geq \max\{R_o^1, R_o^3\}$, the control objective is met, i.e., $B_1[k] \geq B_d$ as k tends to infinity.

Now, suppose $c_2 = A_{out,2}[\tilde{k}]$. Then

$$\lim_{k\to\infty} B_1[k] = \frac{\mu_{11}R_o + \mu_{32}A_{out,2}[\tilde{k}]}{1-\mu_{31}\mu_{32}}.$$

Since, however, $A_{out,2}[\tilde{k}] > 0$, it follows that

$$\frac{\mu_{11}R_o + \mu_{32}A_{out,2}[\tilde{k}]}{1-\mu_{31}\mu_{32}} > \frac{\mu_{11}R_o}{1-\mu_{31}\mu_{32}}.$$

Therefore, any R_o ensuring the right-hand side of the above inequality is greater than B_d also guarantees the left-hand side is greater than B_d. From the case when $c_2 = 0$, if the reward magnitude satisfies $R_o \geq \max\{R_o^1, R_o^3\}$, then the right-hand side of the above inequality is greater than B_d. In other words, when the system begins in regions V, VII or VIII and $R_o \geq \max\{R_o^1, R_o^3\}$, the control objective is met, i.e.,

$$\lim_{k\to\infty} B_1[k] \geq B_d,$$

for either value of c_2.

Now, suppose the system begins in region VI. Suppose $R_o \geq R_o^2$; then by Lemma 5.12, $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$. Thus, by Lemma 5.17, $B_1[k]$ converges as k tends to infinity. Specifically,

$$\lim_{k\to\infty} B_1[k] = \frac{\mu_{11}R_o + \mu_{32}c_2}{1-\mu_{31}\mu_{32}},$$

where c_2 is the value to which $A_{out,2}[k]$ converges as k tends to infinity. From Lemma 5.16, c_2 has two possible values: 0 and $A_{out,2}[\tilde{k}] > 0$ (where $\tilde{k}$ is the sample at which $A_{out,2}[k] > 0$). The analysis to show $B_1[\infty] \geq B_d$ for initial region VI is identical to that of initial regions V, VII and VIII (by replacing R_o^3 with R_o^2). Therefore, a controller of the form (5.9) can meet the control objective of driving $B_1[k] \geq B_d$ as k tends to infinity, provided the reward magnitude satisfies (5.25). □

E.17 Proof of Theorem 5.2

Proof. Suppose for some $\check{k} \geq 0$, $R[\check{k}] = \frac{1}{\mu_{11}} \left(B_d - A_{out,1}[\check{k}] - \Delta A_{out,1}[\check{k}] - \mu_{32} B_2[\check{k}]\right)$. Since

$$B_1[\check{k}+1] = A_{out,1}[\check{k}] + \Delta A_{out,1}[\check{k}] + \mu_{11} R[\check{k}] + \mu_{32} B_2[\check{k}],$$

it follows that

$$\begin{aligned} B_1[\check{k}+1] &= A_{out,1}[\check{k}] + \Delta A_{out,1}[\check{k}] \\ &\quad + \mu_{11} \frac{1}{\mu_{11}} \left(B_d - A_{out,1}[\check{k}] - \Delta A_{out,1}[\check{k}] - \mu_{32} B_2[\check{k}]\right) + \mu_{32} B_2[\check{k}] \\ &= B_d. \end{aligned}$$

Therefore $B_1[\check{k}+1] = B_d$. On the other hand, suppose $R[\check{k}] = 0$. This occurs when $B_d - A_{out,1}[\check{k}] - \Delta A_{out,1}[\check{k}] - \mu_{32} B_2[\check{k}] \leq 0$. Therefore

$$\begin{aligned} B_1[\check{k}+1] &= A_{out,1}[\check{k}] + \Delta A_{out,1}[\check{k}] + \mu_{11} R[\check{k}] + \mu_{32} B_2[\check{k}] \\ &= A_{out,1}[\check{k}] + \Delta A_{out,1}[\check{k}] + \mu_{32} B_2[\check{k}] \\ &\geq B_d. \end{aligned}$$

Hence, $B_1[\check{k}+1] \geq B_d$. Since $\check{k} \geq 0$ is arbitrary, and, for both possible values for $R[\check{k}]$, $B_1[\check{k}+1] \geq B_d$, it follows that for $k \geq 1$, $B_1[k] \geq B_d$. □

E.18 Proof of Theorem 5.3

Proof. First, from Theorem 5.2, the state-feedback portion of the controller ensures $B_1[k] \geq B_d$ for $k \geq 2$. The remainder of the proof is split into two parts. The first part shows that $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, which is used to show each signal related to person one and person two's attitudes converges as k tends to infinity. These results are used in the last part to show that $R[\cdot]$, $B_1[\cdot]$ and $B_2[\cdot]$ are bounded signals.

Part 1

Showing $B_1[k] \geq 0$ for $k \geq 1$ is straight forward. Note that $R[0]$ depends on the initial region. If the system begins in regions V, VII or VIII, then $R[0] = R_o^1$, and it can be inferred from Lemma 5.11 that $B_1[1] \geq 0$. On the other hand, if the system begins in region VI, then $R[0] = R_o^2$, and it can be inferred from Lemma 5.12 that $B_1[1] \geq 0$. Therefore, for each of initial regions V, VI, VII and VIII, (5.27) ensures $B_1[1] \geq 0$. Moreover, from Theorem 5.2, $B_1[k] \geq B_d$ for $k \geq 2$; thus, for $k \geq 1$, $B_1[k] \geq 0$.

The proof to show $B_2[k] \geq 0$ for $k \geq 2$ depends on the initial region, and uses induction (with $k = 3$ as the base step). First, note that Lemmas 5.11 and 5.12 imply that the controller given by (5.27) ensures $B_2[2] \geq 0$ (using an argument similar to showing $B_1[1] \geq 0$).

Suppose that the system begins in region VII or VIII. It has previously been shown that $B_2[1] \geq 0$. To show $B_2[2] \geq 0$, the expression for $B_2[2]$ is examined:

$$B_2[2] = A_{out,2}[2] + \mu_{31} B_1[1]. \tag{E.15}$$

Since $B_2[1] \geq 0$, it follows that $P_2^{CD}[1] \geq 0$ and, therefore, $A_{out,2}[2] \geq 0$ (because $A_{o2} \geq 0$). Note that this conclusion ensures $A_{out,2}[2] + \mu_{31} B_d \geq 0$. Moreover, $B_1[1] \geq 0$. Therefore, $B_2[2] \geq 0$.

Suppose instead, that the system begins in region VI. Then, since it has been previously shown that $B_2[1] < 0$, it follows that $P_2^{CD}[1] \leq 0$. However, since there is a maximum amount of dissonance pressure that can be experienced, $P_2^{CD}[1] \geq \frac{-1}{2}$. Therefore,

$$A_{out,2}[2] = A_{out,2}[1] + K_{12} P_2^{CD}[1] \geq A_{o2} - \frac{K_{12}}{2}.$$

Therefore, all of the terms in (E.15) may not all be non-negative, since A_{o2} (which is positive for initial region VI) may not be high enough to ensure $A_{out,2}[2] \geq 0$. However, since $B_1[2] \geq B_d$ and $B_d > \frac{1}{\mu_{31}} \frac{K_{12}}{2}$, it follows that

$$\begin{aligned} B_2[3] &\geq A_{out,2}[1] + K_{12} P_2^{CD}[1] + K_{12} P_2^{CD}[2] + \mu_{31} B_d \\ &> A_{o2} - \frac{K_{12}}{2} + K_{12} P_2^{CD}[2] + \frac{K_{12}}{2} \\ &= A_{o2} + K_{12} P_2^{CD}[2]. \end{aligned}$$

Since $B_2[2] \geq 0$, it follows that $P_2^{CD}[2] \geq 0$. Furthermore, since the system begins in region VI, $A_{o2} > 0$; thus, $B_2[3] \geq 0$.

Finally, suppose that the system begins in region V. Then,

$$A_{out,2}[2] = A_{out,2}[1] + K_{12} P_2^{CD}[1] = A_{o2} < 0.$$

Therefore, all of the terms in (E.15) are not non-negative. Since, however, $B_1[2] \geq B_d$ and $B_d > \frac{1}{\mu_{31}} |A_{o2}|$, it follows that

$$\begin{aligned} B_2[3] &\geq A_{out,2}[2] + K_{12} P_2^{CD}[2] + \mu_{31} B_d \\ &\geq A_{o2} + K_{12} P_2^{CD}[2] + \mu_{31} B_d \\ &> A_{o2} + K_{12} P_2^{CD}[2] + |A_{o2}| \\ &= K_{12} P_2^{CD}[2]. \end{aligned}$$

Since $B_2[2] \geq 0$, it follows that $P_2^{CD}[2] \geq 0$ and thus, $B_2[3] \geq 0$.

At this point, induction is used with $k = 3$ as the base step to show that $B_2[k] \geq 0$ for $k \geq 2$. Thus, suppose, for $3 \leq k < \bar{k}$, $B_2[k] \geq 0$. Then $P_2^{CD}[k] \geq 0$ over this sample range. Furthermore, from the base step,

$$A_{out,2}[2] + \mu_{31} B_d \geq 0 \tag{E.16}$$

for initial regions V, VI, VII and VIII. Since

$$B_2[\bar{k}] = A_{out,2}[2] + K_{12} \sum_{i=2}^{\bar{k}-1} P_2^{CD}[i] + \mu_{31} B_1[\bar{k} - 1],$$

and since $B_1[\bar{k}] \geq B_d$, it follows that

$$B_2[\bar{k}] \geq A_{out,2}[2] + K_{12} \sum_{i=2}^{\bar{k}-1} P_2^{CD}[i] + \mu_{31} B_d.$$

From (E.16), the above inequality becomes

$$B_2[\bar{k}] \geq K_{12} \sum_{i=2}^{\bar{k}-1} P_2^{CD}[i].$$

From the base steps and the inductive hypothesis, $P_2^{CD}[k] \geq 0$ for $k \geq 2$; thus $B_2[\bar{k}] \geq 0$. Therefore, by induction, $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$.

Now, to show each signal related to the internal attitudes converges as k tends to infinity, first recall Lemma 5.13, which states that if a reward is applied to the two-person system guaranteeing $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$, then $A_{out,1}[k]$ converges to c_1 and $A_{out,2}[k]$ converges to c_2 as k tends to infinity. This conclusion is now used to show $\Delta A_{out,1}[k]$, $\Delta A_{out,2}[k]$, $\Delta A_{out,1}^{CD}[k]$, $\Delta A_{out,1}^{OJ}[k]$, $P_1^{CD}[k]$ and $P_2^{CD}[k]$ converge as k tends to infinity. The attitude signals of person two are considered first.

Since $A_{out,2}[k]$ converges to c_2 as k tends to infinity, it follows that $\Delta A_{out,2}[k]$ must tend to zero as k tends to infinity. To see why, recall that

$$A_{out,2}[k] = A_{o2} + \sum_{i=0}^{k-1} \Delta A_{out,2}[i].$$

Since $A_{out,2}[k] \rightarrow c_2$ as $k \rightarrow \infty$, it follows that the term being summed, i.e., $\Delta A_{out,2}[k]$, necessarily tends to zero as k tends to infinity. This has implications on $P^{CD}[k]$. Recall that person two is not offered a reward; thus $P_2^{OJ,rew}[k] = 0$ for $k \geq 0$. As a result,

$\Delta A^{OJ,rew}_{out,2}[k] = 0$ for $k \geq 0$ and, along with (4.25), this conclusion implies $\Delta A_{out,2}[k] = \Delta A^{CD}_{out,2}[k]$ for $k \geq 0$. Furthermore, (4.17) implies $\Delta A_{out,2}[k] = K_{21} P^{CD}_2[k]$ and therefore, the conclusion that $\Delta A_{out,2}[k] \to 0$ as $k \to \infty$ implies $P^{CD}_2[k] \to 0$ as $k \to \infty$. Therefore, $A_{out,2}[k] \to c_2$, $P^{CD}_2[k] \to 0$ and $\Delta A_{out,2}[k] \to 0$ as $k \to \infty$.

Now, the attitude signals of person one are considered. Similar to person two, since $A_{out,1}[k]$ converges as k tends to infinity, it follows that $\Delta A_{out,1}[k]$ tends to zero as k tends to infinity. However, unlike person two, it is possible for $P^{OJ}_1[k] > 0$ at some $k > 0$. As a result

$$\Delta A_{out,1}[k] = \Delta A^{CD}_{out,1}[k] + \Delta A^{OJ,rew}_{out,1}[k]. \tag{E.17}$$

Each of the two terms in the expression for $\Delta A_{out,1}[k]$ are now shown to converge as k tends to infinity. Suppose that $A_{out,1}[k] < 0$ for $k \geq 0$; then, $P^{OJ,rew}_1[k] = 0$ for $k \geq 0$ and therefore, $\Delta A_{out,1}[k] = \Delta A^{CD}_{out,1}[k]$ for $k \geq 0$. As a result, similar to person two, $\Delta A^{CD}_{out,1}[k]$ and $P^{CD}_1[k]$ tend to zero as k tends to infinity. Therefore, if $A_{out,1}[k] < 0$ for $k \geq 0$, then each attitude signal related to person one converges as k tends to infinity. On the other hand, if there exists a $\hat{k} > 2$ such that $A_{out,1}[\hat{k}] > 0$, then since, for $k \geq \hat{k}$, $A_{rew,1}[k] \geq 0$ and $A_{conf,1}[k] = \mu_{32} B_2[k-1] \geq 0$, Lemma 5.10 applies, meaning $A_{out,1}[k] \geq 0$ for $k \geq \hat{k}$. Therefore, for $k \geq \hat{k}$,

$$P^{CD}_1[k] = r_{41} P^{CD}_1[k-1].$$

Since $B_1[k] \geq 0$ for $k \geq 1$ and $\hat{k} > 2$, it follows that for $k \geq \hat{k}$, $P^{CD}_1[k] \geq 0$. Furthermore, since $|r_{41}| < 1$, $P^{CD}_1[k]$ is a decaying function of k for $k \geq \hat{k}$ and therefore, $P^{CD}_1[k]$ converges to zero as k tends to infinity. As a result, $\Delta A^{CD}_{out,1}[k]$ tends to zero as k tends to infinity. By (E.17), this conclusion implies that $\Delta A^{OJ}_{out,1}[k]$ necessarily tends to zero as k tends to infinity (since both $\Delta A_{out,1}[k]$ and $\Delta A^{CD}_{out,1}[k]$ have now been shown to converge to zero as k tends to infinity). To show this conclusion implies $P^{OJ,rew}_1[k]$ converges as k tends to infinity, the expression for $\Delta A^{OJ,rew}_{out,1}[k]$ is considered. From (4.22),

$$\Delta A^{OJ,rew}_{out,1}[k] = \begin{cases} -K_{21} P^{OJ,rew}_1[k] & \text{if } P^{OJ,rew}_1[k] > 0, K_{21} P^{OJ,rew}_1[k] \leq A_{out,1}[k], \\ -A_{out,1}[k] & \text{if } P^{OJ,rew}_1[k] > 0, K_{2i} P^{OJ,rew}_1[k] > A_{out,1}[k], \\ 0 & \text{otherwise.} \end{cases}$$

It has been previously argued that there are three possible sequences of expressions for $\Delta A^{OJ,rew}_{out,1}[k]$:

- For $k \geq \hat{k}$, $\Delta A^{OJ,rew}_{out,1}[k] = -K_{21} P^{OJ,rew}_1[k]$,
- For $\hat{k} \leq k < \bar{k}$, $\Delta A^{OJ}_{out,1}[k] = -K_{21} P^{OJ,rew}_1[k]$, and for $k \geq \bar{k}$, $\Delta A^{OJ,rew}_{out,1}[k] = -A_{out,1}[k]$, and
- For $k \geq \hat{k}$, $\Delta A^{OJ,rew}_{out,1}[k] = -A_{out,1}[k]$.

Each case is considered below.

If, for $k \geq \hat{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -K_{21} P_1^{OJ,rew}[k]$, then since $\Delta A_{out,1}^{OJ,rew}[k]$ tends to zero as k tends to infinity, it follows that $P_1^{OJ,rew}[k]$ also tends to zero as k tends to infinity. On the other hand, if one of the other two combinations above occurs, then, at least for $k \geq \bar{k}$, $\Delta A_{out,1}^{OJ,rew}[k] = -A_{out,1}[k]$. Therefore, $A_{out,1}[k] = 0$ for $k > \bar{k}$, which, from (4.18), this implies that $P_{raw,1}^{OJ,rew}[k] = 0$ for $k > \bar{k}$. Since

$$P_1^{OJ,rew}[k] = r_{51} P_1^{OJ,rew}[k-1] + (1 - r_{51})\, P_{raw,1}^{OJ,rew}[k],$$

the conclusion that $P_{raw,1}^{OJ,rew}[k] = 0$ for $k > \bar{k}$ means $P_1^{OJ,rew}[k] = r_{51} P_1^{OJ,rew}[k-1]$, for $k \geq \bar{k}$. Since $0 \leq r_{51} < 1$, it follows that $P_1^{OJ,rew}[k]$ tends to zero as k tends to infinity, as required. Therefore, if $A_{out,1}[\hat{k}] > 0$, then all signals related to person one's attitude converge as k tends to infinity.

To summarize, (5.27) ensures $B_1[k] \geq 0$ for $k \geq 1$ and $B_2[k] \geq 0$ for $k \geq 2$ for initial regions V, VI, VII and VIII. This implies that $A_{out,1}[k]$, $A_{out,2}[k]$, $\Delta A_{out,1}^{CD}[k]$, $\Delta A_{out,2}^{CD}[k]$, $\Delta A_{out,1}^{OJ}[k]$, $\Delta A_{out,2}^{OJ}[k]$, $P_1^{CD}[k]$, $P_2^{CD}[k]$, $P_1^{OJ}[k]$ and $P_2^{OJ}[k]$ each converge as k tends to infinity.

Part 2

Since $\Delta A_{out,1}[k]$ and $\Delta A_{out,2}[k]$ each converge as k tends to infinity, it follows that the only possible term in the controller expression for $k \geq 1$, that may possibly be unbounded is $B_2[k]$. Note that $B_2[k] = A_{out,2}[k] + \mu_{31} B_1[k-1]$. Therefore, the controller expression for $k \geq 1$, can be expanded to

$$R[k] = \max\left\{0, B_d - A_{out,1}[k] - \Delta A_{out,1}[k] - \mu_{32} A_{out,2}[k] - \mu_{31}\mu_{32} B_1[k-1]\right\rangle.$$

$R[k]$ is bounded below by zero. Moreover, $A_{out,1}[k]$, $A_{out,2}[k]$ and $\Delta A_{out,1}[k]$ converge as k tends to infinity and therefore are bounded. Additionally, since $B_1[k] \geq B_d$ for $k \geq 2$, it follows that $-B_1[k] \leq -B_d$ for $k \geq 2$, i.e., $-B_1[k]$ is bounded above for $k \geq 2$. Finally, since B_d is a constant, the second expression in the controller is bounded above. Therefore, $R[\cdot]$ is bounded.

Since $R[\cdot]$ is bounded, it follows that $B_1[\cdot]$ is bounded. Finally, because $B_2[k] = A_{out,2}[k] + \mu_{31} B_1[k-1]$, it follows that $B_2[\cdot]$ is bounded. □

Bibliography

[1] Robert A. Adams and Christopher Essex. *Calculus: A Complete Course, Seventh Edition*. Pearson Education Canada, 2009. 140

[2] Icek Ajzen. *Attitudes, Personality, and Behavior*. Open University Press, Milton Keynes, 1988. xi, 8, 12

[3] Icek Ajzen and Martin Fishbein. *Understanding Attitudes and Predicting Social Behavior*. Prentice Hall, 1980. 8

[4] Jean-Léon Beauvois and Robert-Vincent Joule. *A Radical Dissonance Theory*. Taylor and Francis, London, 1996. 13

[5] Donald Broadbent. *Perception and Communication*. Oxford University Press, 1958. 1

[6] Noam Chomsky. A Review of B. F. Skinner's Verbal Behavior. *Language*, 35, 1959. 1

[7] Edward Deci. Effects of Externally Mediated Rewards on Intrinsic Motivation. *Journal of Personality and Social Psychology*, 18(1):105–115, 1971. 13, 17

[8] Edward L. Deci, Richard Koestner, and Richard M. Ryan. A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, 125(6):627–668, 1999. 17

[9] Leon Festinger. *A Theory of Cognitive Dissonance*. Stanford University Press, Stanford, California, 1957. 1, 13, 16, 26, 79

[10] Leon Festinger and James M. Carlsmith. Cognitive Consequences of Forced Compliance. *Journal of Abnormal and Social Psychology*, 58:203–210, 1959. 15

[11] Martin Fishbein and Icek Ajzen. *Predicting and Changing Behaviour*. Psychology Press, New York, 2010. 7, 8, 9, 76, 78, 79, 94

[12] Jay W. Forrester. *Industrial Dynamics*. Productivity Press, 1961. 2

[13] Pushkin Kachroo. *Pedestrian Dynamics: Mathematical Theory and Evacuation Control*. CRC Press, 2009. 2

[14] David G. Myers and Steven J. Spencer. *Social Psychology*. McGraw-Hill Ryerson, Toronto, 2002. 3, 8, 11, 78, 79, 95

[15] Ulric Neisser. *Cognitive Psychology*. Appleton-Century-Crofts, 1965. 1

[16] William T. Powers. *Behavior: The Perception of Control*. Aldine Publishing Company, 1973. 2

[17] J.D. Sterman. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, 2000. 2, 18

[18] Abraham Tesser and John Achee. *Dynamical Systems in Social Psychology*, chapter Aggression, Love, Conformity and Other Social Psychological Catastrophes, pages 95–109. Academic Press, 1994. 95