



## UvA-DARE (Digital Academic Repository)

### Understanding and optimising electronic audit and feedback to improve quality of care

Gude, W.T.

**Publication date**

2019

**Document Version**

Other version

**License**

Other

[Link to publication](#)

**Citation for published version (APA):**

Gude, W. T. (2019). *Understanding and optimising electronic audit and feedback to improve quality of care*. [Thesis, fully internal, Universiteit van Amsterdam].

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

## **Chapter 2**

# Control Theory to design and evaluate audit and feedback interventions

Wouter T. Gude  
Niels Peek

In: Scott P, de Keizer NF, Georgiou A, eds. Applied interdisciplinary theory in health informatics: a knowledge base for practitioners. Amsterdam: IOS Press; 2019.

### **Abstract**

Control Theory is about the processes underlying the behaviour of self-regulating agents. It proposes that behaviour is regulated by a negative feedback loop, in which the agent compares the perception of its current state against a goal state and will strive to reduce perceived discrepancies by modifying its behaviour. Although studies in health informatics often do not report the use of this theory, the principle of a negative feedback loop underlies many applications in the field. This chapter describes how Control Theory fits within health informatics, discussing its role in the development and assessment of audit and feedback interventions in healthcare. Control Theory has been used to synthesise evidence of audit and feedback, and to design and evaluate interventions to improve the quality of blood transfusion practice, cardiac rehabilitation, and intensive care. This has driven progress in our understanding of the underlying mechanisms of audit and feedback for improving health care, and has helped to design better interventions.

**Learning objectives** After reading this chapter the reader will:

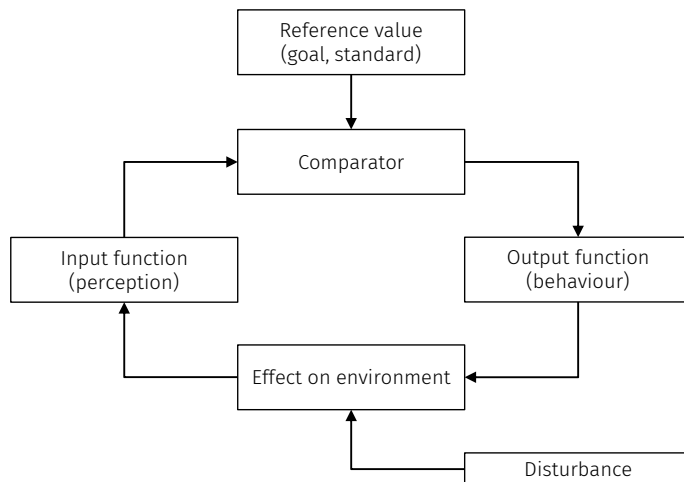
1. Have a broad understanding of Control Theory and its applications in behavioural sciences.
2. Understand the application of Control Theory in the design and evaluation of health informatics interventions, such as audit and feedback; decision support; health behaviour change apps; and supervised machine learning.
3. Understand how Control Theory has been used in scientific studies to evaluate and improve the design of audit and feedback interventions.

## Introduction of Control Theory

Control Theory (CT) as espoused by Carver and Scheier [100] is a general approach to understanding the behaviour of self-regulating agents, which could be humans or artefacts. Its conception is usually traced to the publication of Wiener's seminal book on cybernetics – the science of feedback processes involving control or regulation of certain values within living or artificial systems [101]. Since then, CT has influenced a diverse range of fields including engineering, applied mathematics, economics, medicine, and cognitive and behavioural science. Ammons [102] used feedback processes in the context of human learning, and stated that a person's knowledge of their own performance, obtained through feedback on that performance, will affect the rate of learning and the competence level ultimately reached by that person.

### Discrepancy-reducing feedback loop

The core component of CT is a negative feedback loop (Figure 2.1), termed negative because its function is to negate, or reduce, discrepancies between a perceived present state and a *reference value* (such as a goal state or standard). An agent perceives its current condition via an *input function*, and compares that perception against the reference value through a mechanism termed a *comparator*. If the agent observes a difference between the two values, it will attempt to reduce the discrepancy by performing a behaviour (termed the *output function*). The behaviour usually does not counter the discrepancy directly but has an impact on the *agent's environment*. This should lead to a different present condition, which in turn is perceived by the input function and compared to the reference value. This arrangement thus constitutes a closed loop of control, the overall purpose of which is to minimise deviations from the standard of comparison.



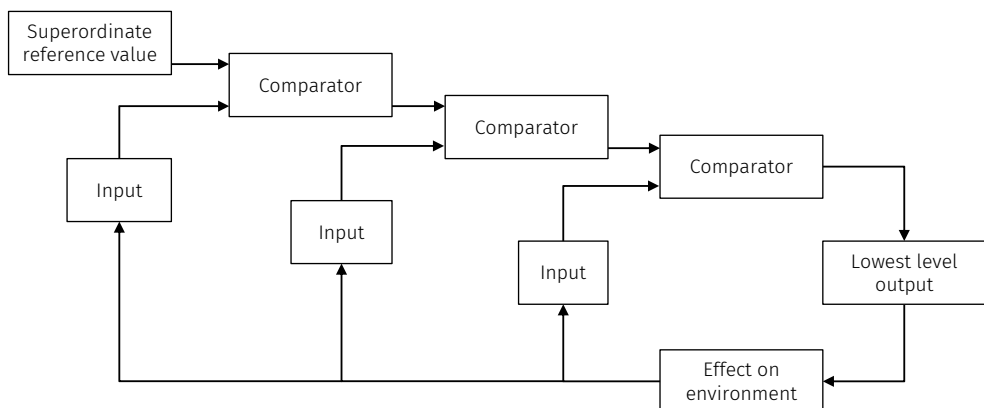
**Figure 2.1:** Negative feedback loop (source: Carver & Scheier 1982 [100]).

Feedback processes like the one described above can occur in diverse physical systems; the best-known example of which is the thermostat. In this system, the input function contin-

uously samples current air temperature from a particular environment such as a room. This input information goes to the device that compares the sensed value to the thermostat's settings. As long as there is no notable difference between values, nothing will happen. If the comparator does detect a difference between values, it sends a message that turns on the heater (output function) which begins to bring warm air into the room. The thermostat will continue to request activity from the heater until the room has warmed up enough so that it can no longer sense a discrepancy between the current air temperature and the thermostat's setting. In the next section we will provide examples of feedback processes in Health Informatics (HI) interventions. As [Figure 2.1](#) shows, effects on a system's environment do not only depend on operations by the output function. Disturbance, originating outside the loop, does not affect the components of feedback loop directly, but it can modify the agent's perceptions via the input value and lead to changes in the discrepancy from the standard. These changes can be adverse (creating or increasing the discrepancy) or favourable (closing the discrepancy). In the example of the thermostat, an open window on a cold day might allow cold wind to enter the room and reduce room temperature; increasing the gap with the thermostat's target temperature range. Alternatively, a warm sun shining through the window or a large group of people standing in the room radiating excess body heat might establish the opposite effect. In that case, the result of the disturbance is that there is no need for an output adjustment because the system observes no discrepancy. Hence the main purpose of the feedback loop is not to undertake an action, but to create and maintain the perception of a specific desired condition i.e. no discrepancy between the input and reference value.

### Hierarchical systems and reference values

Feedback loops are often organised in a hierarchical fashion such that there are superordinate and subordinate systems ([Figure 2.2](#)) [103]. Each system relates to superordinate (at the higher end of the hierarchy) or subordinate (at the lower end of the hierarchy) goals, where achievement of subordinate goals is a requisite for attaining superordinate goals. Superordinate systems act by changing the reference value of the subordinate system at the lower level in the hierarchy. That is, the output of the superordinate level sets the standards for the next lower level. In turn, the subordinate system changes the reference value for the next lower level, and so on.



**Figure 2.2:** Three-level hierarchy of interconnected feedback loops (source: Carver & Scheier 1982 [100]).

Building on the thermostat example, the thermostat receives its reference value from a superordinate system which may be a person in the room. This person also has a reference value e.g. 'be comfortably warm'. Rather than operate directly on the environment to produce heat, for example, by building a fire, the person operates by providing a new reference value to the subordinate system—resetting the thermostat to a warmer temperature setting. Due to the change in reference value, the thermostat activates the furnace and the room temperature rises.

Hierarchically-organised systems each act to create their desired condition and monitor their own input which exists at their own level of abstraction. In the example of the two-level hierarchy of the person and their thermostat, the thermostat assesses air temperature whereas the person assesses their comfort level. Superordinate goals at the higher end of the hierarchy tend to be more abstract whereas subordinate goals at the lower end are more tangible and concrete. As the subordinate system executes, both systems progress towards discrepancy reduction. Higher level (i.e. abstract) goals may be achieved more gradually over time than lower level (i.e. concrete) goals.

Especially amongst human behaviour theorists it is thought that hierarchies may have many levels of control [103], where behaviour outputs at the highest level are to live up to one's self-image (e.g. to be a responsible and thoughtful person) and go all the way down to muscle movements at the lowest level. So, while the basic negative feedback loop from the thermostat example is relatively simple and has limited applicability in the real world, the hierarchical approach enables modelling of arbitrarily complex systems, both in the mechanical, biological, and behavioural world.

### **Alternative strategies of reducing discrepancies**

Feedback loops involving human behaviour are more complex than in relatively simple systems such as the thermostat. Whereas ideally potential discrepancies are resolved through behaviour, an alternative strategy is to change the reference value so that it better matches the input value. Both responses would effectively reduce perceived discrepancies, however with a different effect on the environment. For example, people being confronted by the fact that they are not achieving a lifestyle goal (e.g. walking 10,000 steps a day) sometimes respond by changing the goal rather than changing their behaviour. Other options are to reject the reference value, or to abandon the situation (physically or mentally) that signals the discrepancy, e.g. deeming the reference value unachievable or discounting the presented data or source. For instance, professionals may question the quality of underlying data when feedback on their performance indicates that it is below par. It is difficult to predict which strategy will be used in response to perceptions of discrepancies, but it is likely that alternative strategies are used when attempts by the output function to change the input seem to repeatedly fail or if the discrepancy is perceived to have a low likelihood of reduction through actions [104].

### **Similarities and differences with related theories**

Several other theories use the concept of feedback as a central component. Key examples are Goal-setting Theory (Locke & Latham [105]), Feedback Intervention Theory (Kluger & DeNisi [104]), and Social Cognitive Theory (Bandura [106]).

Goal-setting Theory [105] posits that people are motivated to achieve a goal rather than reduce discrepancies. In particular, it describes the mechanisms through which goals (com-

parable to CT's reference value but set by people themselves) influence behaviour and the relationship between goal characteristics and subsequent performance i.e. goal attainment. The theory proposes that specific goals are more effective than general ones (e.g. 'do your best' goals); and that challenging yet achievable goals lead to better performance than both trivial goals and overambitious goals.

Feedback Intervention Theory [104] considers feedback as the provision of information regarding some aspects of someone's performance on a certain task. According to this theory, people's behaviour is regulated by goals and standards which are organised, as also posited by CT, in a hierarchical fashion. Attention is limited and usually directed at an intermediate level within the hierarchy; only gaps that receive attention have the potential for change. Feedback works by providing people with new information which allows a shift of attention either toward the task or away from it. An attention shift towards the task tends to strengthen the feedback's effect on task performance whereas a shift away from it weakens the effect. The theory proposes that feedback characteristics, the nature of the task performed, and situational and personality variables determine how effectively this shift occurs.

Social Cognitive Theory [106] aims to guide the study of human behaviour, thought and motivation. It proposes that environment, behaviour, and personal and cognitive factors all interact as determinants of each other. The theory argues that self-efficacy, the beliefs regarding one's capabilities of successfully completing tasks, determine what challenges people choose to undertake, how much effort to expend in the endeavour, how long to persevere in the face of obstacles and failures, and whether failures are motivating or demoralising [106]. The relationship between those beliefs and behaviour is described, similar to CT, as a reciprocal learning process in which people select, react to, and learn from experiences.

## **Usage of Control Theory in health informatics**

There are numerous examples of HI interventions that are based on the principle of a negative feedback loop, although few would explicitly reference CT. So, most references to CT in HI are implicit, and developers of interventions would often not consciously use the negative control loop when they design their tool or software – but the control loop would tacitly play a role in their intervention. The same holds for evaluation studies of interventions that build on CT: such evaluation studies would often assume a feedback loop around which the study is designed, without explicitly referencing CT. It is therefore challenging to assess how broadly CT is used in HI. However, we believe that there is a profound influence of CT on HI, and we will illustrate that by describing various broad areas that involve a feedback loop as a key component. The principal area that we will use to describe the usage of CT in HI is audit and feedback, but we also highlight several other areas.

### **Audit and feedback**

audit and feedback (A&F) interventions [21] aim to improve the quality of care by comparing observed quality parameters (quality indicators) with predefined quality targets or benchmark values. Typically, feedback on quality indicators is delivered to healthcare professionals on a regular basis, thus enabling multiple cycles through the control loop. For instance, in the Netherlands all 32 teaching intensive care units (ICUs) and 51 non-teaching ICUs participate in the National Intensive Care Evaluation (NICE) [77], and receive biannual feedback reports on standardised mortality ratios, readmission rates, length of stay, and other quality indicators.

A&F is one of the most widely-used interventions in quality improvement and implementation research. It is generally used when the patient is not present, thereby making it distinctly different from clinical decision support.

We can map the components of the feedback loop (Figure 2.1) to elements of A&F interventions as follows. The input function (perception) consists of the feedback on clinical performance that sits at the heart of each A&F intervention, and would typically materialise through recurring, paper-based or electronic reports issued by a national audit or governing body. Feedback reports summarise the performance of individual clinicians or clinical units over a set period of time (e.g. the last 3 months) using pre-defined indicators of clinical quality, typically using a combination of graphical and numerical information (scores). Reference values may be either explicitly provided or left implicit, and different for each quality indicator. Feedback on clinical processes is often determined by reference values provided by national guidelines, while feedback on outcomes would often be determined through benchmarking between care providers. For instance, in the NICE feedback report, outcome statistics are compared to the national average and the average of a group similar sized ICUs. In A&F interventions it is commonly left to the recipient of the report (i.e. the clinician or clinical unit) to interpret the information and translate it into behaviour. For instance, a unit might decide to start a quality improvement initiative based on poor performance scores in feedback reports. If that initiative is effective, improved performance should transpire in subsequent reports and can inform the decision about whether or not to continue the programme. However, there may also be disturbances (e.g. organisational barriers) that impede actual improvements to care quality, despite the efforts of the quality improvement initiative.

Over the last decade, A&F interventions have increasingly moved from using static, paper-based (or PDF) documents to interactive electronic tools. The interactive computer interface of an e-A&F intervention may allow users to filter, drill down and further explore their performance summaries. For example, NICE participants can also view these data, updated after each monthly data upload, on a website called NICE Online and perform subgroup analyses [107]. In general, if an A&F system is linked to an electronic health record database, performance summaries may be generated on demand at each point in time, thus creating more flexibility for users. A recent review of e-A&F evaluation studies [108] classified them using the theoretical domains framework [109], an integrated theoretical framework synthesised from 128 theoretical constructs from 33 theories judged most relevant to implementation questions. The review found that the domains of knowledge; motivation and goals; and 'social influences' were most commonly targeted by these interventions. In contrast, professional identity and emotion were never targeted.

Despite the clear roots of A&F in CT, it is uncommon that CT is explicitly mentioned in studies concerning the design or evaluation of A&F interventions – but it does happen that related theories are mentioned. Colquhoun and colleagues [31] conducted a systematic review of the use of theory in randomised controlled trials of A&F interventions. They found that only 20 out of 140 studies (14%) reported use of theory in any aspect of the study design, measurement, implementation or interpretation. Only 13 studies (9%) reported the use of a theory to inform development of the intervention. A total of 18 different theories across educational, psychological, organisational and diffusion of innovation perspectives were mentioned. Arguably, many of these resonate with elements of CT. For instance, Social Cognitive Theory also proposes that feedback processes drive behaviour; and the Theory of Planned Behaviour postulates that attitudes, social norms, self-efficacy, and controllability provide reference values for behaviour.



## Other uses of Control Theory in health informatics

There are other HI areas that draw upon the negative feedback loop depicted in [Figure 2.1](#) and thus have roots in CT. As with the A&F literature, explicit references to CT are rare – but we would nevertheless argue that there is a clear relationship. We provide three examples here.

First, a large number of decision support systems for the management of long-term conditions such as diabetes and cardiovascular diseases have been developed that deploy a negative feedback loop for controlling important clinical parameters. For instance, Athena [110] is a clinical decision support for the management of hypertension that issues an alert to its clinician users whenever a patient's latest blood pressure measurement is too high. Similarly, Pandit [111] is a web-based diabetes management system for patients that asks them to measure and enter their blood glucose level. Whenever a glucose value is outside the normoglycaemic range, the system responds by suggesting adjustments to the patient's insulin dose. Similar mechanisms have been used in expert systems for critical care [112].

Second, many smartphone apps that assist in health behaviour change provide users with feedback on their achievements against pre-set goals, such as the number of steps taken, or time spent on physical activity per day. The feedback aims to incentivise users to increase their level of healthy behaviour when it is below target, and maintain it when it is on par.

Third, CT plays a central role in supervised machine learning methods such as Hebb's learning rule [113]; the Newton-Raphson algorithm [114]; gradient boosting [115]; and deep learning [116]. Essentially, each of these methods utilise the negative feedback loop to derive a model of an input-output function from training data. Initially, a default or random model is chosen that bears no relationship to the data, and that model is subsequently 'trained' to better fit the data. The feedback is always derived from discrepancies between observed outputs (in the data; typically called training labels) and predicted outputs (predicted by the model). At each iteration of the feedback loop the classifier will better approximate the input-output function that produced the data, and the discrepancies will disappear after which the process is terminated.

## Explanation of success or failure of audit and feedback

Despite being commonly applied as a healthcare quality improvement strategy, A&F interventions yield variable and often only marginal effects [21]. Moreover, over four decades of research in the field seems to have failed to enable A&F researchers to successfully enhance intervention designs and achieve larger effects consistently. It has been only recently that the use of extant theory has been recognised as an essential component of both the design and evaluation of A&F interventions [28]. In response there have been various studies making explicit use of CT to enhance understanding the A&F's underlying mechanisms in improving healthcare and quality interventions. We have selected four recent studies that jointly illustrated the breadth of activities that can be supported with CT. These activities include interpreting published literature; designing new interventions; secondary analysis of clinical trial data; and the design of new scientific experiments.

## Synthesising evidence from A&F interventions

In an illustration of how theory can be used to synthesise published evidence from behaviour change interventions, Gardner *et al.* [38] used CT to organise, understand and synthesise evidence relating to behaviour change techniques within A&F. Using CT as conceptual frame-

work, the authors hypothesise that A&F may be enhanced through the use of specific performance targets to permit comparison between current and target performance, and action plans to inform behavioural adjustment to reduce discrepancy [38]. The authors conducted a re-analysis of the 2006 Cochrane review [117], recoding each study included in the review, to test target-setting and action plans as effect-modifiers of A&F. The results however were inconclusive because very few studies explicitly described their use of targets or action plans. When Ivers *et al.* updated the Cochrane review in 2012 [21], and repeated Gardner's analysis, explicit targets and action plans were found to be significant effect modifiers of A&F.

### **Improving the design of A&F to increase uptake of evidence-based blood transfusion practice**

The second illustration is a study by Gould *et al.* [118] that used CT to enhance the content of a feedback intervention for improving blood transfusion practice. The authors describe the feedback loop as a dynamic, iterative process of control in which “individuals manage their behaviour by knowing what they want to do or achieve (i.e. setting a goal or standard), trying to do it (i.e. action), monitoring the behaviour (i.e. audit), assessing whether they are making progress towards the goal (i.e. feedback, which informs the nature and extent of any discrepancy between behaviour and goals), and adapting what they do in light of the feedback (i.e. action planning)” [118]. They also used the taxonomy of behavioural change techniques [119] for identifying and describing intervention components that are consistent with CT and that may enhance practice. A number of the techniques included in the taxonomy encompass strategies proposed in CT, such as ‘goal setting’; ‘feedback on behaviour’; ‘discrepancy between behaviour and goal’; and ‘action planning’. The authors then aimed to enhance content in feedback documents by incorporating behavioural change techniques consistent with CT that were previously absent. For example, to incorporate goal setting as a change technique, the authors added an introductory statement in documents that proposes an evidence-based goal, e.g. “XX% of patients with [XX clinical attributes] are likely to require transfusion and so we suggest that, within your clinical team, you make this your explicit goal”. The authors propose that such enhanced feedback has the potential to facilitate the enactment of CT's feedback processes and lead to larger improvements.

### **Quantitative process evaluation of an A&F intervention in cardiac rehabilitation**

The third study illustrates how CT can be used to enrich quantitative process evaluations of A&F interventions. In this case CT was instrumental to understand the outcomes observed in a cluster randomised trial of e-A&F to improve cardiac rehabilitation [54]. The intervention involved local quality improvement teams receiving quarterly web-based feedback in combination with outreach visits. Feedback was given on 18 quality indicators and included benchmark comparisons. During the visits teams reviewed their feedback and selected indicators they wished to improve upon into their action plan—within the same web-based system. For each indicator that teams targeted for improvement, they were prompted to describe the problem, goal, and concrete actions on how to achieve the goal. During each outreach visit (corresponding to feedback cycles), teams reviewed the new performance scores and updated their action plan accordingly. The associated trial did not show any significant changes in either care processes or patient outcomes [120]. Following CT, the intervention's ineffectiveness must have been either the result of the possibility that feedback indicating sub-benchmark

performance still failed to convince recipients to change, actions were not completed, or completed actions were ineffective. We designed a two-part study to investigate the first gap. Part 1 was a laboratory experiment involving 41 individual cardiac rehabilitation professionals who were given two feedback reports in an adjusted version of the web-based feedback system. These professionals were asked to select indicators for improvement, based on the feedback. If their response was at odds with CT's hypothesis (i.e. indicator's performance score was below the benchmark but not selected for improvement; or the score was above benchmark and still selected for improvement), they were asked to explain their choice. Part 2 was a field study concerning a secondary analysis of the trial data, in which multidisciplinary teams selected indicators for improvement across multiple cycles of feedback. Regression techniques were applied to assess determinants of cardiac professionals' intentions to improve practice. The principal findings were that performance scores and benchmark comparisons influenced intentions, but between one third and half of the time intentions were at odds with CT because professionals either disagreed with benchmarks; deemed improvement unfeasible; or did not consider the indicator an essential aspect of care quality. In addition, it revealed that intentions remained similar in subsequent feedback cycles (because actions were not completed) and that professionals prioritised improving data quality rather than care quality. This study contributed to the understanding of A&F in both the current intervention and in general, in the sense that it quantified how often the feedback loop stagnates and provided insight into the determinants and reasons for not following feedback.

### **Understanding the influence of A&F in pain management in intensive care**

In the previous example CT was used as a conceptual framework in a post-hoc analysis of decision processes. The final illustration also concerns work of our own research team and builds on the findings of the previous study, but in this case CT was used in the very design of the experiment. This study involved an e-A&F dashboard providing intensive care teams with periodic feedback on four pain management indicators [55, 121]. Inspired by the cardiac rehabilitation study we recognised that healthcare professionals often already have beliefs about their clinical performance and feedback may fail to change those beliefs. We studied the extent to which those beliefs correspond to actual practice; how they are influenced by feedback; and, ultimately, how feedback changes intentions to improve practice. To that end we designed an online two-step experiment, driven by CT, to elicit these beliefs and intentions before and after receiving first-time feedback. The experiment took place upon first login into the dashboard; 83 intensive care professionals from 21 units participated. In step 1, professionals were presented with the indicator descriptions whilst withholding all performance feedback (that is, no performance scores or benchmark comparisons were displayed). Professionals were asked to estimate for each indicator their own unit's performance score, the national average score, the minimum score they would consider "good performance" (target), and whether or not they would perform actions to improve. The study found that half of the time professionals overestimated their own performance and rarely underestimated it. Targets were set very high. In step 2 professionals received feedback on their performance. Feedback included the unit's own performance, median and top 10% peer performance, and improvement recommendations based on peer comparisons (good performance; room for improvement; or improvement recommended). Professionals were asked again, but now given the performance information at hand, what their performance target was and whether they intended to improve practice. If improvement intentions were at odds with CT in step 1 (score < target and no intention; or score  $\geq$  target and still intention) or in

step 2 (e.g. room for improvement but no intention to improve) we asked professionals to explain their choice. Also, if there were discrepancies between intentions in the first and second step, professionals were asked what feedback elements drove them to change (e.g. measured score or benchmark was higher/lower than expected). Even before receiving any feedback some 68% of professionals' intentions corresponded with the feedback recommendations. In other words, while professionals were not very good at estimating absolute performance, they had good intuitions about whether it was on target or not—without seeing any numerical information. After receiving the feedback, this number increased to 79%. In more than half of the cases in which units were already top performers, professionals still wanted to improve. In 8% of cases professionals lacked improvement intentions because they did not consider the indicators important; did not trust the data; or deemed benchmarks unrealistic. This research concluded that audit and feedback does indeed help healthcare professionals to work on those aspects for which improvement is recommended because it increases the accuracy of their clinical performance perceptions. However, given the abundance of professionals' prior good improvement intentions, efforts to optimise A&F interventions should focus on translating those intentions into (effective) actual change in clinical practice.

## Discussion

Control Theory (CT) provides a conceptual framework for self-regulation and human behaviour and has already demonstrated its usefulness for the field of HI and in particular A&F interventions. In the A&F literature CT has been used to synthesise evidence of interventions, enhance their design, explain why interventions were or were not successful, and generate hypotheses about how feedback mechanisms work in practice. Nevertheless, the majority of studies have not explicitly reported the use of CT (or other relevant theories) for such purposes.

The simplicity of CT's negative feedback loop makes for an elegant framework that is widely applicable, but it also has limitations. Individuals may compare feedback to multiple internal standards or goals at the same time; based on beliefs about past performance, expectations, norms, or an ideal goal [104]. Further, HI interventions like A&F interventions are typically complex and placed into a social and organisational context. This context is not in the scope of CT; taking it into account would require the use of different theories such as Social Cognitive Theory. Finally, in contrast to for example Feedback Intervention Theory or Goal-setting Theory, CT provides no guidance as to which factors related to the context, recipients, or feedback itself may influence success of the feedback loop. Nevertheless, it is fair to say that CT has been very influential in our thinking about information systems and behaviour and will undoubtedly continue to do so.