The truth about balancing loops - and how to avoid unintended consequences!

Dennis Sherwood

Managing Director, The Silver Bullet Manufacturing Company Limited

Introduction

Balancing loops? An article about balancing loops?? And with a crazy title about 'truth'? What truth? Everyone knows about balancing loops – balancing loops are taught in the class before Systems Thinking 101!

Well, yes; I thought I knew everything about balancing loops too. But I've recently come to the view that what I knew, though not 'wrong', certainly hasn't been 'right' – or rather, I think there are some assumptions that I had habitually made about balancing loops without even recognising that I was making them: assumptions that certainly are false in many real situations; assumptions that had the unintended consequence that a lot of the balancing loops I'd drawn were just plain wrong. Maybe I'm dumb, and particularly so at drawing causal loop diagrams; but maybe the assumptions I hadn't noticed I'd been relying on are so deep-seated that perhaps some other people haven't noticed them either – hence my taking the liberty of writing this…

Beware unintended consequences

To set the scene, here is a quotation from an article from one of the UK's 'quality' newspapers, *The Sunday Telegraph*, published on 30th June 2013: on the front page, under the banner headline *A&E is 'grinding to a halt'* we read:

> *Patients have been forced to wait up to eight hours in ambulances queuing outside accident & emergency units amid a crisis in the system. … Last year* [28th October 2012] *this newspaper revealed that Monitor, the regulator of NHS foundation trusts* [the administers of state-run National Health Service hospitals in England], *had warned hospitals not to keep patients in ambulances in order to comply with the government target of treating all patients admitted to A&E* [accident and emergency departments] *within four hours. Monitor said that hospitals who used the technique – known as "stacking" – risked "serious implications" for their patients and tied down ambulances needed to respond to emergencies. Although the practice began under Labour* [the UK's previous government], *the extent of this has risen sharply under the Coalition* [the current UK government formed jointly by the Conservative and the Liberal Democrat parties]. *Local newspapers have found instances of as many as 14 ambulances queuing up outside a hospital A&E department. In a letter to trusts last year, Monitor specifically warned hospitals against "gaming to meet health care targets". "We would encourage all trusts to ensure that such practices are not taking place at your hospitals," the regulator's letter read. "Evidence of foundation trusts carrying out these practices would be taken very seriously by Monitor."*
> (*The Sunday Telegraph*, 2013.)

Here is the background to this story: in 2000, the UK government, at that time led by Labour Prime Minister Tony Blair, announced a new target for hospital accident and emergency departments - by 2004, all patients arriving at A&E must be seen within four hours of arrival (Letham and Gray, 2012). This, of course, was very well-intentioned - the idea is clearly to incentivise hospitals to treat A&E patients within a reasonable time. All very sensible. But as this article describes, one way to ensure that this performance measure is fulfilled is to keep patients waiting in ambulances outside A&E, for that four-hour clock doesn't start running until they patient's name is registered...

This is just one, real - and rather startling - example of an 'unintended consequence': something that no one could possibly have expected.

Or could they? The UK Health Service, and the government civil service, is full of lots of very clever people. How could they design and implement a system with such an inherent flaw? Or rather many flaws, for "stacking" in ambulances is just one of many ways in which hospitals have met their performance measures successfully whilst failing to deliver the service that really matters.

Performance measures are of course embedded in balancing loops, and so in this article I'm taking a deeper look at balancing loops, for I've come to the conclusion that a key reason why 'unintended consequences' happen is attributable to some assumptions that are deeply embedded within the balancing loops we know so well...

The balancing loops we know well

So, let's take a look at that oh-so-familiar balancing loop: the standard example that is to be found in every text-book (see, for example, Morecroft, 2007, p.42; Senge, 1990, p.84; Sterman, 2000, p.111) – including mine (Sherwood, 2002, p.56) too! Here it is
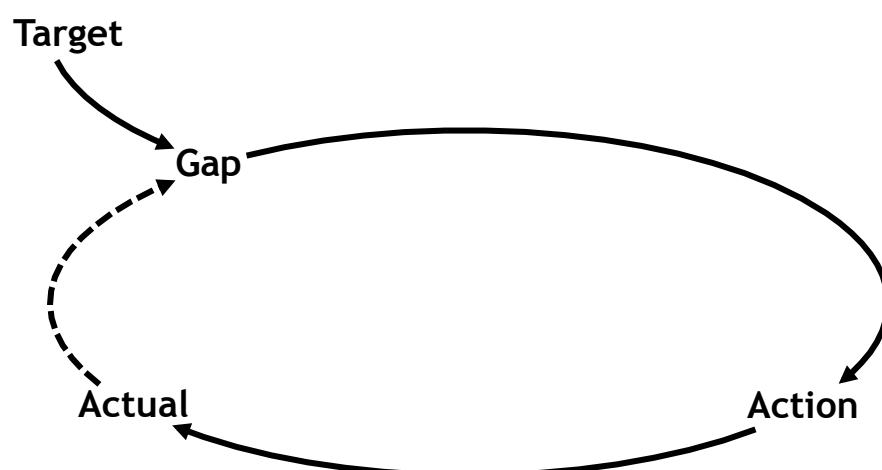
What could be more straight-forward? Any *gap* between the *target* and the *actual* triggers an *action* to change the *actual* to become closer to the *target*, until such time that the *actual* becomes equal to the *target*, at which point the action of the loop stops, and the system remains stable at the *target* level – just like a domestic thermostat regulating the temperature of a room.

Indeed. And since all budgeting systems, and all performance measures, are about setting and meeting *targets,* balancing loops of this form underpin much of organisational life.

Ah! That's the trap – the hidden assumption: that innocent-looking phrase 'balancing loops of this form'. Balancing loops, yes. But 'of this form', no. For the 'canonical form' of the standard balancing loop makes two implicit assumptions:

- firstly, that a *gap* between any *target* and the corresponding *actual* will actually trigger *action*, and
- secondly, that any such *action* is indeed the 'right' *action* to close the *gap*.

Business school professors may well be puzzled at this point. Of course a *gap*, a *variance*, triggers *action*! What else would it do? And of course all managers want to meet their *targets* in the best possible way, so of course they'll take the 'right' *action*! What else would they do?

Well, I've encountered many situations in which a *gap* doesn't trigger any *action* at all, and where that *action,* when it happens, isn't the right *action* either.

Balancing loops in (my) real world

So, let me explain. The big difference between a thermostat, which does indeed take *action* when there is a *gap*, and a human being is that the human being feels things, has emotions. My starting point is therefore to suggest that, in a human system, the result of any *gap* between a *target* and an *actual* isn't *action* – it's a feeling of a lower, or higher, level of *anxiety*, which itself depends on the individual's degree of *apathy*:
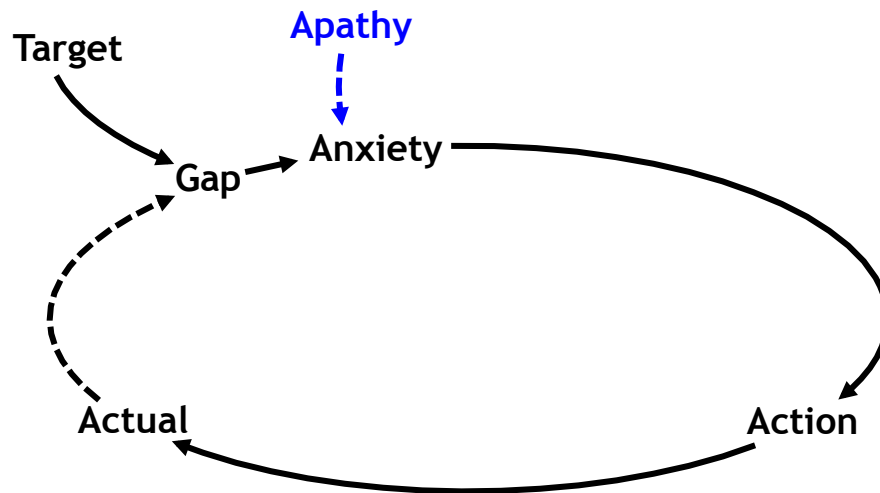
My mental model here is that we should bear in mind the possibility that the individual ostensibly required to take the appropriate *action* may, in fact, not be remotely concerned as to whether the *actual* and the *target* are close together, or miles apart – indeed, the individual might not regard the *target* with anything other than disdain, especially if he individual who has to take the *action* is different from the individual who set the *target*, and had no part in the target-setting process.

This is probably more a feature of the professional world than the commercial one: a doctor, for example, might feel that her main objective is to treat patients, not meet financial targets; a 'rain maker' lawyer might be motivated by personal billings rather than to share 'my' client with a fellow partner, despite the firm's 'ethos of collegiality'; a school teacher might wish to devote time to the development of the pupil, rather than just do the minimum required to get the pupil through some test.

So, if the individual just doesn't care about any *gap*, or indeed about the *target*, if the level of *apathy* is high, then this causes the degree of *anxiety* about any *gap* to be correspondingly low, and so no *action* is taken – the *gap* just gets bigger and bigger. But not indefinitely: sooner or later, there will be a big row between those setting the *target* and those who are ignoring it:
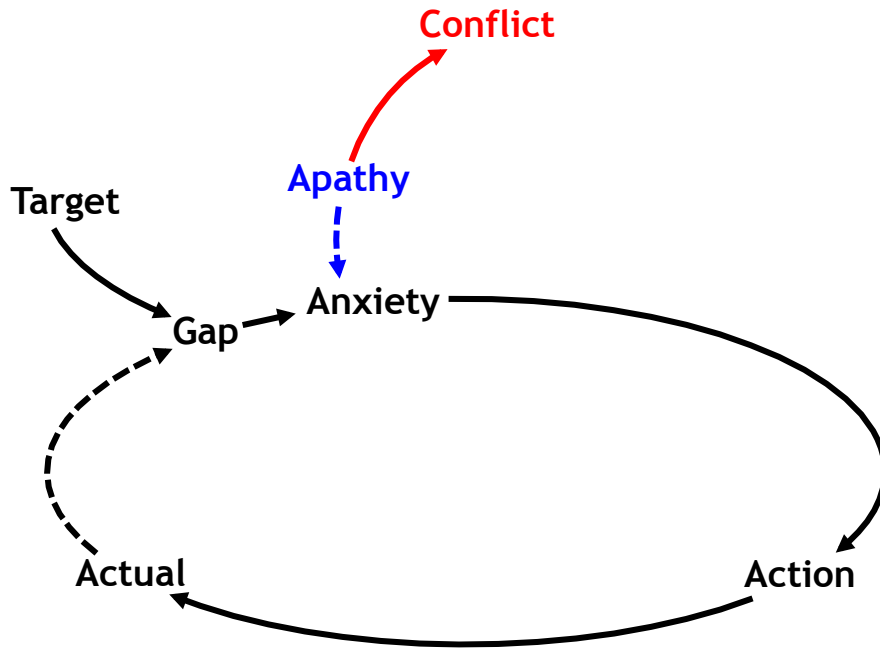
If, however, the individual is conscientious, the level of *apathy* is very low, and even a small *gap* generates *anxiety*. But does this then cause *action*? Perhaps; but I think that something else happens first – as the individual's *anxiety* increases, this increases not *action* but the *pressure to act*:
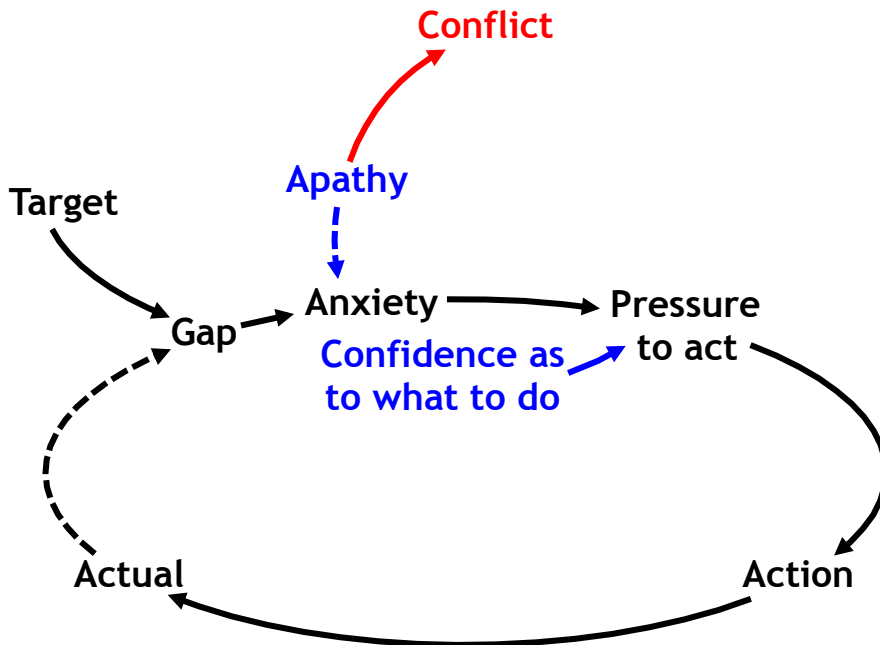
An individual who is conscientious, who knows what to do, who has all the necessary means, and who is *confident*, will respond to any *gap* between the *target* and the *actual* by taking *action*, and speedily too. But if, like Hamlet, I'm not sure, if I worry, if I dither, if I'm indecisive, if I lack *confidence*, then the *gap* between what I *ought to do* and what I *actually do* can be a yawning chasm before I *actually do* anything at all. As all great drama vividly illustrates.

The *pressure to act* is also influenced by two other important factors. Suppose, for example, that I am conscientious, that I am concerned about the *gap*, and that I therefore sense a strong *pressure to act*. But suppose further that I just *don't accept that the level of the target is right* – yes, I agree that there should be a target, but I genuinely believe that the level of the *target* is wrong. If I have no power, then I will be forced to acquiesce, and accept the level of the *target* as it is. But If I do have power, then I will resist. And I will resist even more strongly if I *don't believe in the target at all*:
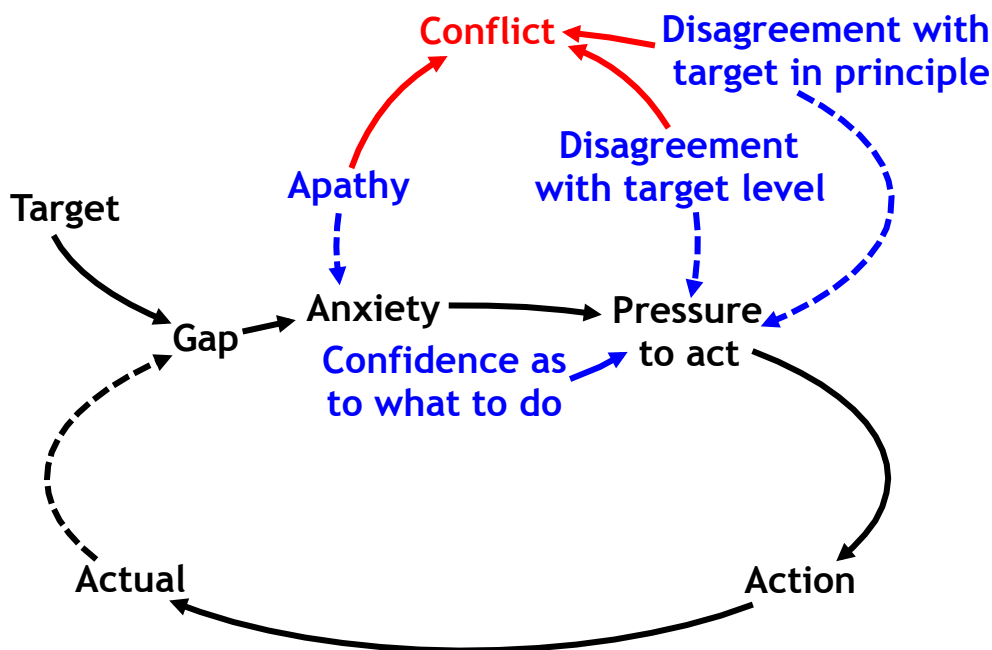
Both of which result in *conflict*: the more the individual (or the community) under *pressure to take action* is in *opposition to the target level*, or to the *target in principle* – or indeed in opposition to those setting the *target* - the greater the likelihood of *conflict*, from relatively low-level grumbling, to more active non-compliance, to strikes, to political action, to revolution.

Before the 'revolution' breaks out, however, it's likely that those responsible for setting the *target* (or more likely those close to them) realise that there is an alternative to resolving the *conflict* other than by 'shouting louder' – maybe the *target* was unreasonable, or just not achievable ("how come that *gap* is still there, even though more resources have been piled in, and we changed the Project Manager?").
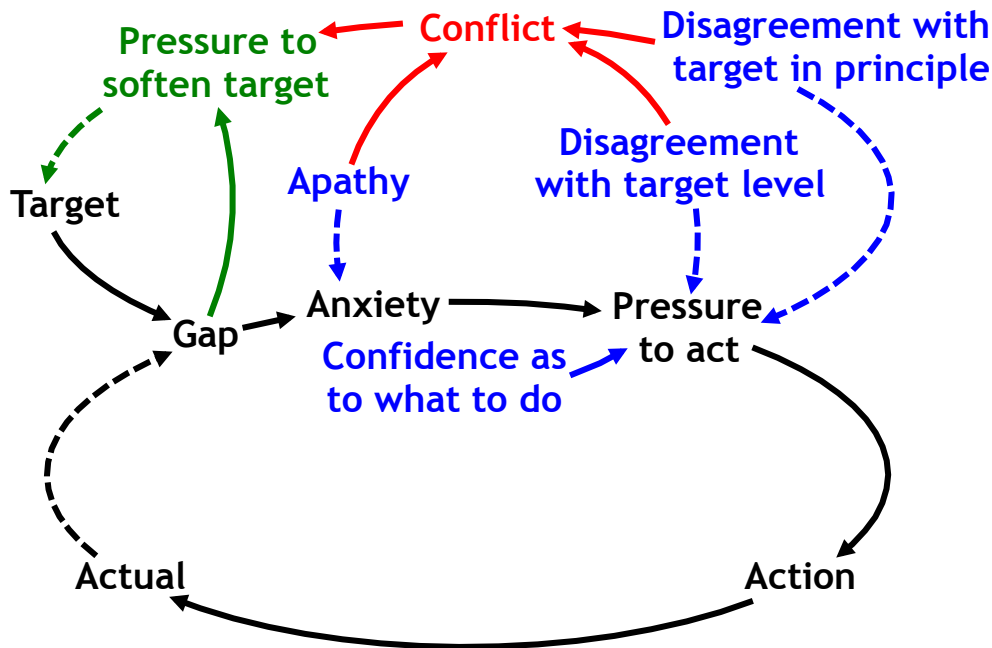
Whether or not any *pressure to soften the target* has any effect – before it's too late – depends very much on whether the *target* setter is, or is not, willing to change his or her mind, to listen to sound advice. Which is a very difficult for many people to do, especially those who see themselves as strong, decisive, determined leaders. Surely, changing one's mind is a sign of weakness and indecisiveness, isn't it? And what about losing face? So the stubborn 'leader' presses on regardless, failing to distinguish between pig-headedness and wisdom. Yes, human beings are indeed more complex than thermostats.

Let's move on: suppose that *action* is indeed taken. What happens? In some situations, there is little choice; but in many situations, there are choices – choices between 'easy' ways of doing things and 'hard' ways:
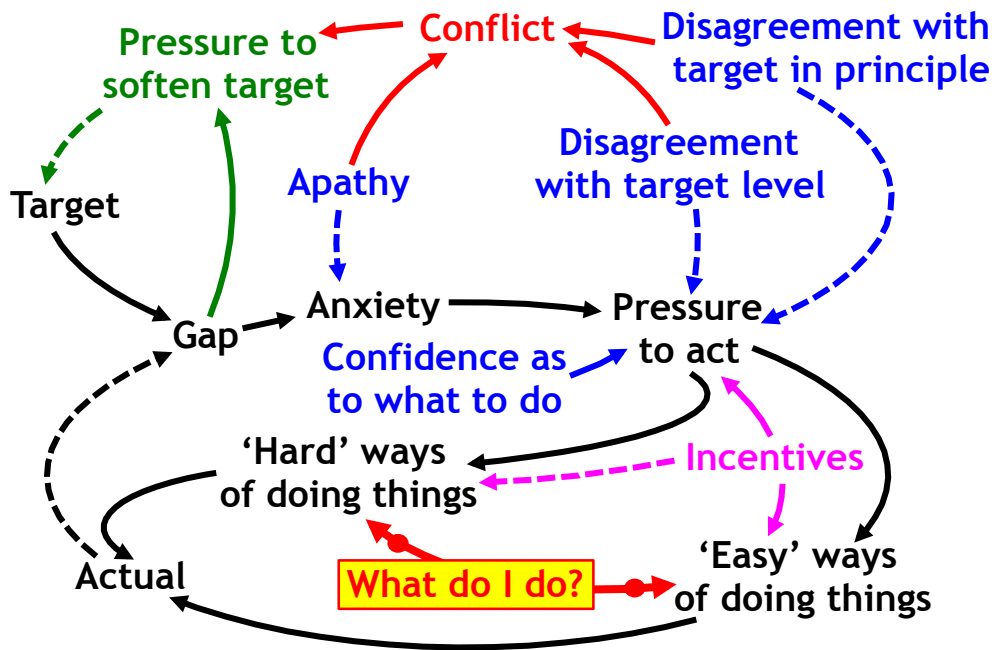
What choice is actually made? That's a no-brainer: the *easy way* - why would anyone voluntarily do something the *hard way* if there is an *easy* alternative? But there's a catch: in reality, the *easy ways* are often in some sense 'wrong', and the *hard ways* 'right' – but the natural, and oh-so-human, tendency to opt for the *easy way* tempts people to take the *easy way*, and hope they will get away with it. This is the territory of many so-called 'unintended consequences' – situations in which a system has been designed, with good intent, but in practice fails because real people have discovered ways to operate within the system which the designer just never thought of. Like "stacking" A&E patients in ambulances.

Are these truly 'unintended consequences'? Or are they evidence of poor systems design? Poor because the system designer made an unnoticed assumption that people would always do the 'right thing', and didn't even consider the possibility that there might be *easy ways* of making the system 'work'? Anticipating that people might choose, or indeed discover, 'easy' ways of 'gaming the system', what can a wise designer do to encourage people to take the 'right' actions, even though they might, in some ways, be 'harder'?

A real example – school examinations in England

So, let's make all that real. And to do that, let me cite another of the UK's 'quality' newspapers, *The Times*. Here is an extract from an article that appeared on Monday 14th October 2013, under the headline ***Gove to overhaul GCSE league tables to stop schools 'gaming' the system***:

School league tables are to be overhauled amid fears that bright and struggling pupils are being sidelined. Michael Gove, the Education Secretary, wants to stop schools from improving their ranking by simply maximising the number of students achieving five GCSEs with a C grade or above. Schools will be barred from boosting their league position by entering pupils for subjects in which they believe it is easier to obtain a crucial C grade.
(*The Times*, 2013)

The context here concerns the English school examinations known as 'GCSE' – the General Certificate of Secondary Education. These examinations are taken by students at around the age of 16, and they are important – a candidate's results are the direct entry qualification to post-16 education, they are significant as regards admission to subsequent higher education at colleges and universities, and they are material as regards employment. To give a feel for scale, and referring only to the GCSE examinations taken in England in the Summer of 2013, the official statistics (http://www.jcq.org.uk/examination-results/gcses) state that some 1.5 million candidates sat around 5 million examinations in subjects such as English, French, mathematics, biology, drama and economics. The subject curricula are defined by the UK government's Department for Education (of which Michael Gove, referred to in the article from *The Times,* is the senior minister – see https://www.gov.uk/government/organisations/department-for-education), but the GCSE examinations in England are designed, set, administered, marked and graded by (at the time of writing) three competing commercial enterprises known as 'awarding organisations'.

For the last 20 years, GCSEs have been graded as A*, A, B, C, D, E, F and G, with A* being the highest grade. Each candidate is awarded a grade in each subject taken, and, since 1994, an important performance measure for every school – and hence of each school's Head Teacher - has been the percentage of their students who achieve at least five GCSEs with grades of A*, A, B or C (see, for example, http://www.education.gov.uk/schools/performance/archive/index.shtml).

Over the years since 1994, various details have changed, but the overall measures are more-or-less the same: a Head Teacher will 'win' if a candidate gets a C, but 'lose' if the grade is a D. Hence the content of the quoted article in *The Times*, which is about the incentives on a school, for example, not to bother much with brighter students (if the student is very likely to get a B, why 'waste' time and energy in encouraging her, and in motivating and teaching her, at a level which will achieve an A?) and to dump less able students (if the student is struggling to get a D, maybe the student shouldn't be entered for the examination at all…).

Given the social and national importance of schools in general, and examinations in particular, school education in England is supervised by two regulators, known by the acronyms Ofsted (http://www.ofsted.gov.uk) and Ofqual (http://ofqual.gov.uk), both of which are independent of the government of the time, but are (separately) accountable to parliament. Ofsted's remit includes the inspection of schools, and Ofqual regulates the standards of school examinations, including GCSEs, primarily by exerting influence on the awarding organisations.

The stakeholder context within which all this happens is bewilderingly involved, and highly political too. Everyone in the land wants schools to be good, and all the candidates, and their parents, wish for high examination grades. So do the Head Teachers, and of course the individual subject teachers, of the schools, for high marks reflect well on them. Examination grades, and their implied quality, are important to potential 'users' such as employers and institutions of further education, for they rely on the grades to select candidates. The three commercially competing awarding organisations have a vested interest too, for each wishes to increase its market share. And let's also bear in mind that over the years, each successive government has interpreted the number of higher grades awarded as evidence of the success of their education policies, so the politicians like lots of high grades too. Everyone in the country has an interest in schools, and school examination systems; everyone has an axe to grind.

Ofqual's role as the regulator of examination standards in the middle of all this is complex indeed. If the standards are too tough, then fewer candidates get higher grades, and a lot of people get upset; if the standards are too low, there are lots of happy candidates with high grades, but maybe employers start to say "that qualification isn't worth the paper it's printed on!". And, when you think about it, defining 'the standard' isn't so obvious – how, exactly, do you define the 'standard' of, say, an English essay?

A community central to all of this are the Head Teachers (or equivalent) of the schools, colleges and other institutions that teach students who will ultimately sit the appropriate examinations, and become beneficiaries of the corresponding qualifications. As I've mentioned, in England, there are currently three commercial awarding organisations who administer those critical GCSE examinations, each offering its own (quite) different examination, set in accordance with a (slightly) different syllabus: for although the UK government sets the syllabus in general, some differences are allowed. So, for example, in GCSE English Literature, the government might specify that one Shakespeare play must be studied in some depth. One awarding organisation might have *Romeo and Juliet* as a set text, whilst another might choose *Twelfth Night*. An important choice for the Head Teachers and their colleagues is therefore "which awarding organisation?", for that choice will then determine whether the students spend a year studying *Romeo and Juliet*, or *Twelfth Night*. How is this choice made? Is it

a question as to whether the Head of the English Department prefers *Romeo and Juliet* to *Twelfth Night*? Or are some other factors at work?

So let's look at the world through the eyes of a hypothetical, but I trust reasonably representative, Head Teacher.

As the manager of a large institution – as almost every school and college indeed is – the Head Teacher will be subject to a set of performance measures, which determine, at the end of the year, whether or not that Head Teacher has done 'a good job' – or indeed not. Those performance measures imply the existence of some corresponding targets – targets which, if met, bring the Head Teacher appropriate glory; if missed, the corresponding opprobrium. A rational Head Teacher will therefore make choices that are more likely to contribute to the successful attainment of those targets rather than to failure, all of which implies that the Head Teacher lives within a balancing loop, which, if expressed in standard form, would look like this:

Target

Gap

Actual

Action
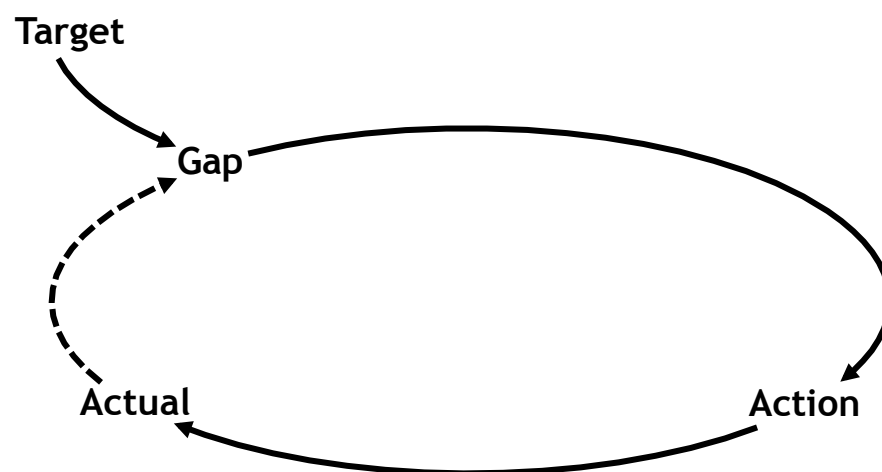
But I believe that in the real world of the Head Teacher, things are much more complex. Here is my version of the upper part of the diagram:
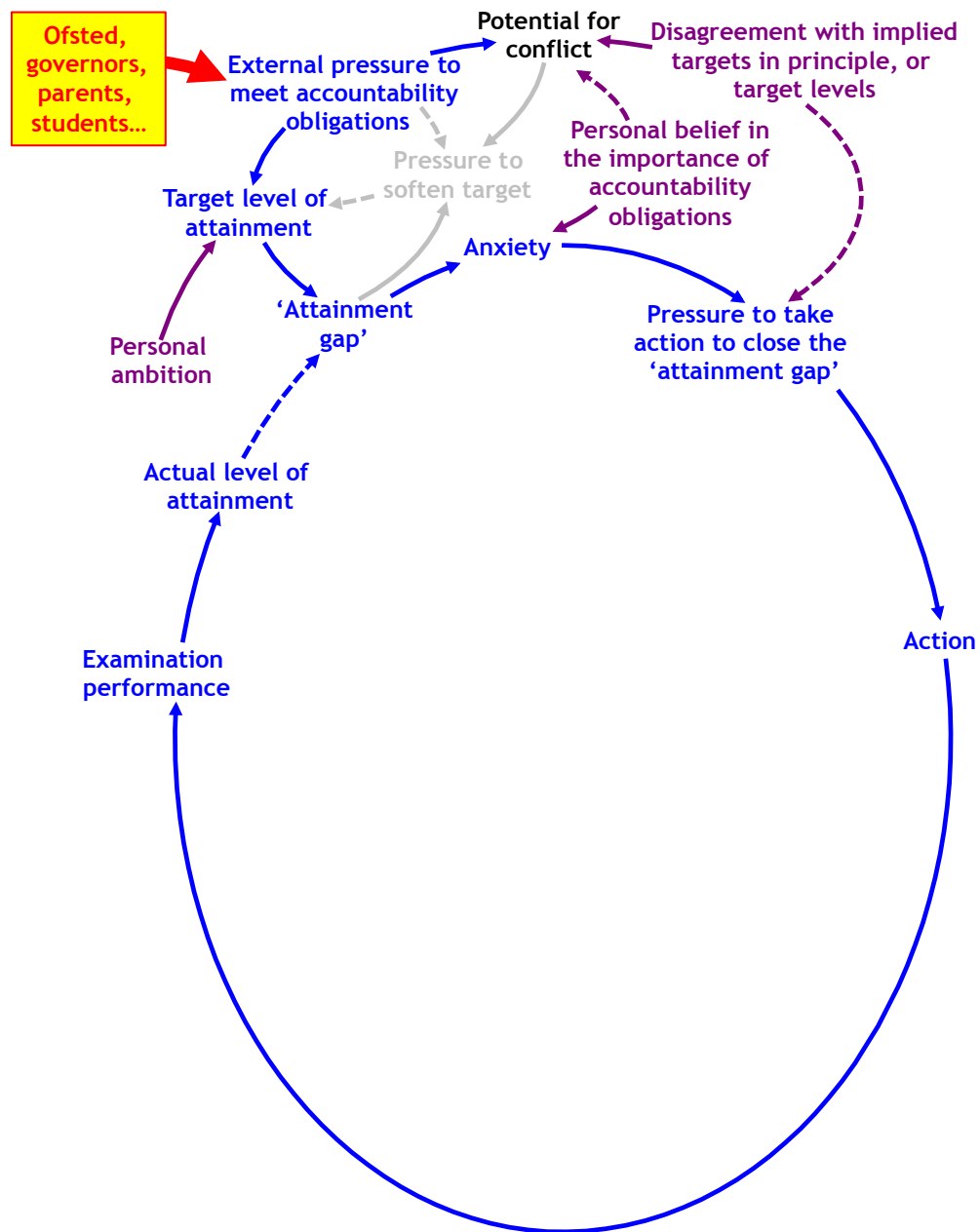
The Head Teacher is under *pressure to meet a variety of accountability obligations*, which themselves are attributable to the requirements of a variety of *stakeholders* who are in a stronger, or weaker, position to exert that *pressure*. Take, for example, the school governors, to whom the Head Teacher is accountable in a managerial sense – for it is the governors who hired the Head, and can fire the Head too. The governors have a sense of what they want the school to be, and look to the Head to deliver that vision. The students are stakeholders too, but weak; somewhat stronger is the community of parents.

And there is another major player too – Ofsted, which has significant power indeed: Ofsted carries out regular, and generally feared, school inspections, and the resulting report can influence the reputation of the school, and – more importantly from the Head's point-of-view – whether or not the Head has a career. And – despite the similarity of the acronyms – let me stress that Ofsted is different, and quite separate, from Ofqual: Ofsted regulates the schools; Ofqual regulates the awarding organisations who set and grade the examinations. Yes, it is indeed a highly complex stakeholder environment.

Back to the Head Teacher, who is under *pressure to meet a variety of accountability obligations*, obligations which in turn determine the *target level of attainment* which the Head, ostensibly, tries to achieve, with those Heads of high *personal ambition* perhaps setting an even higher *target level of attainment* than as required by the external stakeholders.

Any difference between the *target level of attainment* and the current *actual level of attainment* opens a *gap* - but what happens next depends on a number of important factors. Let's suppose, for example, that the *accountability obligations* comprise a number of different factors, of which one is the Head Teacher's personal sense of accountability to parents to develop each student to the best of their individual abilities – stretching the best students, and supporting the less able. At the same time, a particular governor might be a sports fanatic, dominating every governors' meeting with a (long!) speech about how important it is that the school wins a particular national sports trophy; meanwhile, an Ofsted inspection is scheduled for a month's time, and the Head knows that the criticism about science teaching, made after the last inspection, needs to be fixed. Given all these competing demands, the Head takes the view that failing to meet the sport-fanatic governor's demands will cause less pain than another bad Ofsted report, and although the Head knows that he faces a row at the next governors' meeting, he takes a decision that the senior chemistry teacher needs to be seen to be delivering a high-quality lesson by the Ofsted inspector, rather than taking that afternoon coaching the soccer team. That may be a trivial example, but you get the point.

Less trivial is what happens next. Let's suppose that there is very strong pressure on the school to rank as highly as possible in the published league table of school performance that rank schools according to the percentage of students who achieve grades of A*, A, B or C in five or more GCSE exams. This is a real, and very strong, pressure indeed – and the important point about this measure is the threshold of grade C: any student who achieves a grade C or higher contributes to 'good' performance, whilst students who achieve a D don't. Furthermore, under the current government rules, a grade of an A* carries no more weight in the league table score than a C: grades of A*, A, B and C are of equal 'value' as regards the league table – but not, of course, to the student.

The reality of this is that, from the point of view of the league table, scoring a B, an A or an A* gives no benefit – a C will do. So Head Teachers are under pressure to push as many students over the D-C boundary as they can, which, as we will shortly see in more detail, leads to many possible dysfunctional behaviours, such as not stretching more able students towards an A or an A*. If the Head strongly believes in the development of the 'whole student', this can cause a very strong reaction: the Head could well disagree fundamentally with the paramountcy of league tables in general, and how the league tables are constructed and weighted in particular, and this can cause significant *conflict*. But those *external pressures* are very strong, and the Head Teacher is most unlikely to be able to exert any effective *pressure to change the targets*, and is obliged to 'play the game'…

The roles of examination grades, and the league tables, are captured in the next diagram:
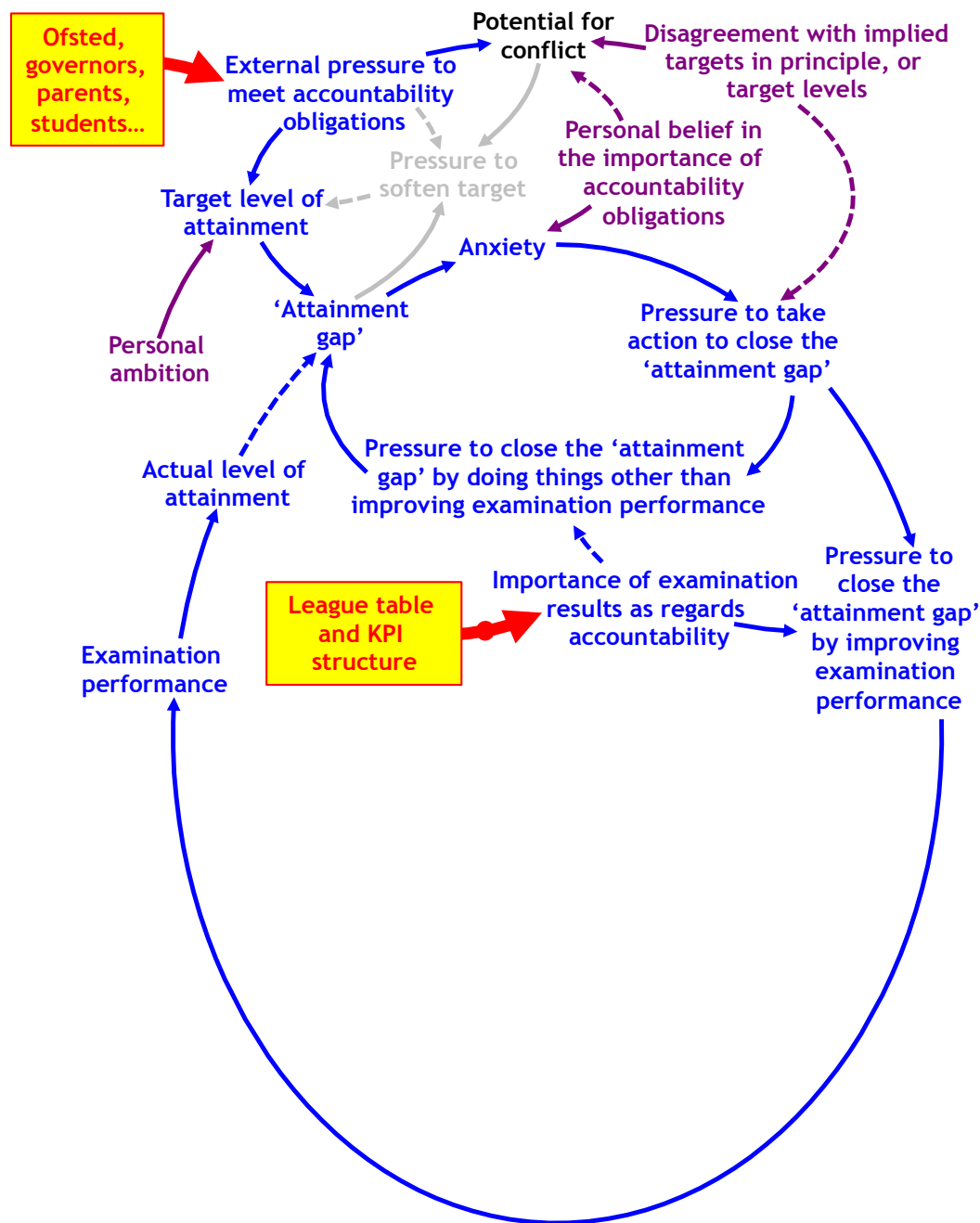
Note: In this, and subsequent, diagrams, the symbol ➡ signifies influence, not directionality. It is used only with input dangles, and so has no bearing on loop polarity.

Those aspects of *attainment* which are not related to examinations – such as prowess in sport, music, drama, or just simply the development of a well-brought-up 'whole person' - can be addressed in any number of appropriate ways, so closing the corresponding *gap*. But if *attainment* is highly dependent on *examination results*, then resources of time and energy will be directed there. And the truth of the matter in England right now is that examination results – especially the percentage of students achieving five or more A*, A, B or C grades in their CGSEs - is very important indeed.

So, given this reality, what might a Head Teacher actually do to improve examination performance? Here are two possibilities:

Ofsted, governors, parents, students…

Potential for conflict

Disagreement with implied targets in principle, or target levels

External pressure to meet accountability obligations

Pressure to soften target

Personal belief in the importance of accountability obligations

Target level of attainment

Anxiety

'Attainment gap'

Personal ambition

Pressure to take action to close the 'attainment gap'

Actual level of attainment

Pressure to close the 'attainment gap' by doing things other than improving examination performance

League table and KPI structure

Importance of examination results as regards accountability

Pressure to close the 'attainment gap' by improving examination performance

How do I do this?

Examination performance

delay

delay

Quality of learning

Pressure to build a stronger 'learning culture'

Quality of teaching

Pressure to employ 'better' teachers

The first possibility is to *build a stronger 'learning culture'*, so improving the *quality of learning*, which will, as a consequence, achieve the desired result of improved *examination performance*. And everyone will agree that this a very 'good thing' to do – and indeed all those inspirational conference speeches from the educational equivalents of Jack Welch and Tom Peters are about exactly that. But building a new culture takes time, and I need those improved exam results this year… and, even worse, I'm not Jack Welch or Tom Peters. I'm just plain old me. And I can sit in thos inspirational speeches till the cows come home, and be truly inspired. *BUT WHAT DO I ACTUALLY DO*?

A second possibility is to *employ better teachers*, so enhancing the *quality of teaching*. That's very likely to work too – but this solution also takes time. And there are some rather difficult knock-on effects too: what happens to all the teachers we currently have? Nasty…

What else might the Head do? For there are some other 'interesting' possibilities, possibilities made available by the way the system of examination performance measures has been designed, for example…
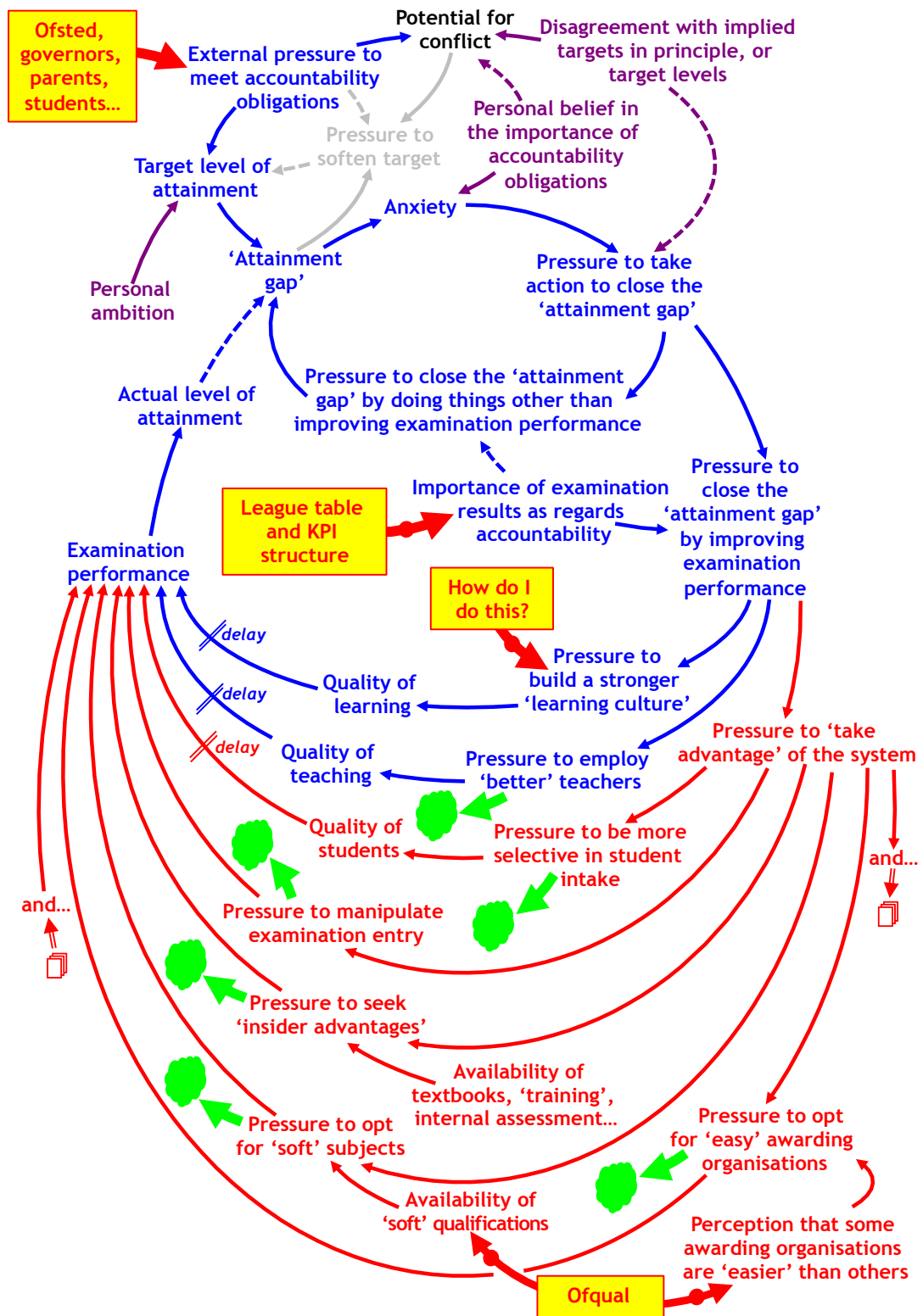
So, instead of employing *better teachers*, why not try to enrol smarter students? What ways might there be of 'bending' our annual *student intake*, without breaking the law?

Rather easier, and perhaps more direct, might be to *manipulate the examination entry*. So, if a teacher feels that a particular student might not make a C grade, perhaps the student might be 'encouraged' not to sit the examination this year... that way, the student isn't counted... Yes, there is the possibility that the student might, in fact, have surpassed all expectations and scored a B. But scoring a D would be bad news, and if this were to happen, the student would have to retake the examination the following year - so why not relieve all the anguish and wait a year anyway?

Another possibility is to try to *get on the 'inside'*. So, for example, some of the awarding organisations publish textbooks, so it's bound to be a safer bet to take their exams, and buy their books, for surely the books are written with the exams in mind – and (even better!) the exams with the books in mind? Furthermore, some awarding organisations offer rather more of their examination in a form which is known as 'controlled assessment' – this being something of a euphemism for the fact that the teacher can have rather more influence over the grade given to a student for at least part of the curriculum, such as work done during the course itself and graded by the teacher, rather than under examination conditions and graded by an unknown 'outsider'...

There may also the option of selecting *softer subjects* – physics, of course, is notoriously 'hard'; maybe sociology, information technology or business studies aren't so demanding - and a C in media studies counts as a C, just like a C in statistics...

And let's not ignore the possibility that perhaps *some awarding organisations might be 'softer' than others*. Some awarding organisations, for example, have examination structures which are modular, and come is smaller 'bite-sizes' – which must be easier. And let's also remember that the three awarding organisations are mutually competitive, all seeking to increase their market share, and all wanting to keep their customers happy – what better way could there be of making Head Teachers happy than to be just that little bit 'softer' on grading? No-one, of course, would ever acknowledge this in public, but I did hear a rumour going around at that conference last month...

Let me stress that I'm not making any of this up. And these are just some of the ways in which schools in England have been 'gaming' the system to optimise their league table positions – there are others too, as indicated on the diagram by the '*and...*'.

In this diagram, the actions shown in blue – *building a learning culture* and improving the *quality of teaching* (and not necessarily by recruiting 'better' teachers) - are, in general, 'good' things to do. But they are also hard things to do, and they take time.

In contrast, the actions shown in red – *manipulating the exam entry* and so on - are 'bad' things to do, but they are relatively easy, and quick. These actions are generally referred to as 'unintended consequences' – a term used as a defence mechanism by the system designer to deflect criticism for bad system design, as if to say "well of course unexpected things will happen – how could those possibly have been anticipated?".

By thinking harder, that's how.

And by recognising that all balancing loops, in human systems, will almost always have two, parallel, paths from *action* to *actual* – the slow, hard path and the quick, easy one. I would argue that it is absolutely essential for all designers of human systems to recognise this, to anticipate the 'unintended consequences', and so design the system from the outset to eliminate them. And just to rub the point home, let me refer once again to the Accident and Emergency example with which I started this paper, for "stacking" patients in the ambulances isn't the only ruse that some hospitals get up to – another is to 'admit' patients within the four-hour deadline by having a nurse say 'hello', take a few details, and then inviting the patient to sit in a further waiting area which is not within the performance measure. No, I didn't make that up.

Anyway, back to schools. The *pressure to meet examination targets* by *'taking advantage' of the system* can cause real *personal stress* in the mind of a conscientious teacher:
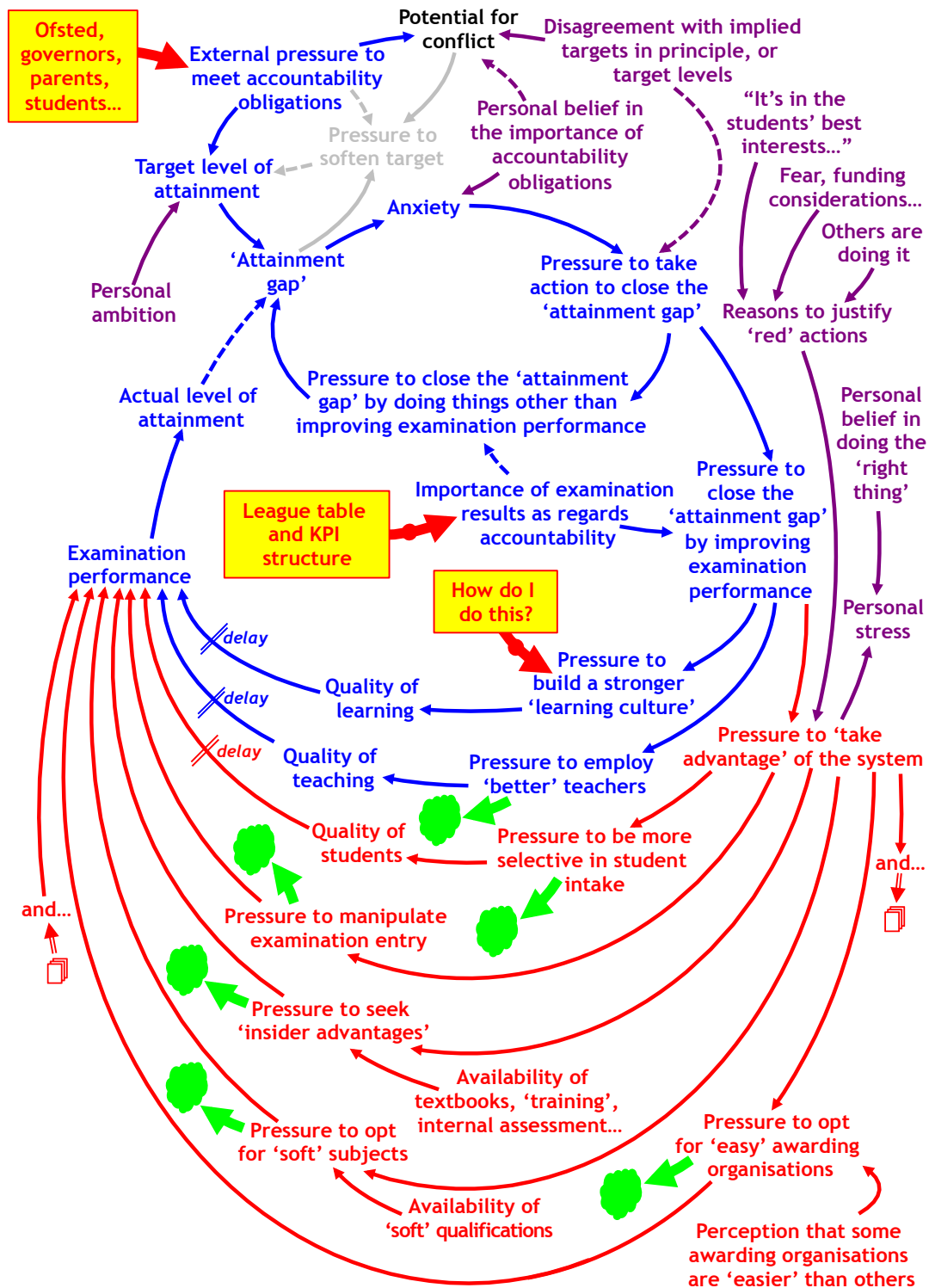
Yes, *other schools may also be gaming* the system, maybe there is an argument that can be made that *it's all in the students' best interests*, or perhaps I'm just plain *scared* of failing to meet my performance measures, all of which help *justify my actions* – but that *stress* is still there…

But where is Ofqual, the examinations regulator, in all this? Ofqual is indeed present, but only 'in the shadows'. Ofqual's regulatory remit relates solely to the awarding organisations, and Ofqual has no direct contact schools at all – school regulation is through Ofsted, not Ofqual (with due apologies to everyone reading this who does not live in England, and to whom these acronyms must be most bewildering!). This is how Ofqual actually does fit in:

Ofsted, governors, parents, students…

External pressure to meet accountability obligations

Potential for conflict

Disagreement with implied targets in principle, or target levels

Pressure to soften target

Personal belief in the importance of accountability obligations

"It's in the students' best interests…"

Fear, funding considerations…

Others are doing it

Target level of attainment

Anxiety

'Attainment gap'

Personal ambition

Pressure to take action to close the 'attainment gap'

Reasons to justify 'red' actions

Actual level of attainment

Pressure to close the 'attainment gap' by doing things other than improving examination performance

Personal belief in doing the 'right thing'

League table and KPI structure

Importance of examination results as regards accountability

Pressure to close the 'attainment gap' by improving examination performance

How do I do this?

Personal stress

Examination performance

delay

Quality of learning

Pressure to build a stronger 'learning culture'

Pressure to 'take advantage' of the system

delay

Quality of teaching

Pressure to employ 'better' teachers

delay

Quality of students

Pressure to be more selective in student intake

and…

Pressure to manipulate examination entry

Pressure to seek 'insider advantages'

Availability of textbooks, 'training', internal assessment…

Pressure to opt for 'easy' awarding organisations

and…

Pressure to opt for 'soft' subjects

Availability of 'soft' qualifications

Perception that some awarding organisations are 'easier' than others

Ofqual

What Ofqual can, and does, do is to influence the awarding organisations to ensure as far as possible that all examinations are of a common standard, and that there is very little, if any, opportunity for an awarding organisation to be selected as the supplier of examinations to a school on the basis of its perceived 'softness'; similarly, Ofqual does its utmost to ensure a parity of standards across different subjects, so reducing the opportunities for choosing supposedly 'softer' subjects.

What Ofqual *can't* do, however, is to interact directly with schools, and so veto, for example, manipulating examination entries. That's Ofsted's job. And what this causal loop diagram obviously portrays is the fundamental importance of Ofqual and Ofsted 'being on the same page', so that the very direct prods given to schools by Ofsted inspectors are totally in harmony with the objectives of Ofqual in maintaining the standards and integrity of examinations. Ofqual and Ofsted, together, can also form a powerful lobby on the government, who determine the structure of the league tables and the weighting of the components within them, as well as the performance measures applied to schools and their Head Teachers. And indeed, over recent months, much has been happening, and continues to happen, to address the dysfunctionalities of the examination systems, as the quotation from *The Times* indicates, and as some recent announcements from Ofqual make very clear – see, for example, http://ofqual.gov.uk/news/survey-launched-gather-teachers-view-experiences-exam-strategies/. *But why was a system with so many 'unintended consequences' designed and implemented in the first place? And allowed to remain operational for so many years?* I don't know. But it surely wasn't well thought-through.

So, back to the truth about balancing loops. In the past, when I was designing human systems, my starting point was the 'standard' form of balancing loop:
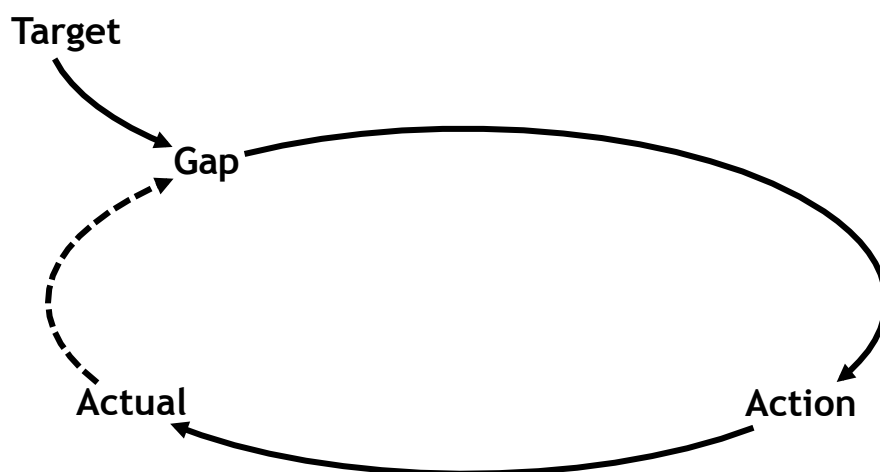
Now, I'm rather wiser – I start here:

Pressure to soften target **Conflict** **Disagreement with target in principle**

Apathy **Disagreement with target level**

Target

Gap Anxiety Pressure to act

Confidence as to what to do

'Hard' ways of doing things Incentives

Actual **What do I do?** 'Easy' ways of doing things

And I work very hard indeed to think through all the *'easy' ways*, and seek the advice of those who have experience within the current systems. If we all think rather harder, using the 'human' version of the standard balancing loop rather than the 'thermostat' version as our starting point, we'll all design better systems, with far fewer 'unintended consequences'. And the world will surely be a better place as a result.

References

Letham K and Gray A. 2012. *Emergencias 24: 69-72* and
   http://www.semes.org/revista/vol24_1/15_ing.pdf

Morecroft J. 2007. *Strategic Modelling and Business Dynamics.* John Wiley & Sons Ltd. Chichester.

Senge PM. 1993. *The Fifth Discipline*. Century Business. London.

Sherwood D. 2002. *Seeing the Forest for the Trees – A manager's guide to applying systems thinking*. Nicholas Brealey Publishing, London.

Sterman JD. 2000. *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Irwin/McGraw-Hill, Boston.

The Sunday Telegraph, 2013
http://www.thepaperboy.com/uk/2013/06/30/front-pages-headlines.cfm, http://www.pressdisplay.com/pressdisplay/viewer.aspx, http://www.telegraph.co.uk/health/healthnews/10150635/Patients-facing-eight-hour-waits-in-ambulances-outside-AandE-departments.html and http://www.telegraph.co.uk/news/9637865/Dont-leave-patients-in-ambulances-to-hit-AandE-targets-hospitals-told.html.

The Times, 2013
http://www.thetimes.co.uk/tto/news/politics/article3893857.ece and http://schoolsimprovement.net/gove-to-overhaul-gcse-league-tables-to-stop-schools-gaming-the-system/