

# Evaluating the Wider Outcomes of Schools: Complex Systems Modelling for Leadership Decisioning

Educational Management  
Administration & Leadership  
1–25

© The Author(s) 2016

Reprints and permission:

[sagepub.co.uk/journalsPermissions.nav](http://sagepub.co.uk/journalsPermissions.nav)

DOI: 10.1177/1741143215597233

[emal.sagepub.com](http://emal.sagepub.com)



**Ruth Deakin Crick, Steven Barr, Howard Green  
and David Pedder**

## Abstract

A continuing challenge for the education system is how to evaluate the wider outcomes of schools. Wider measures of success – such as citizenship or lifelong learning – influence each other and emerge over time from complex interactions between students, teachers and leaders, and the wider community. Unless methods are found to evaluate these broader outcomes, which are able to do justice to learning and achievement as emergent properties of the learner's engagement with his or her world the education system will continue to focus on narrow measures of school effectiveness which do not properly account for complexity. In this article we describe the rationale and methodology underpinning a pilot research project that applied hierarchical process modelling to a group of schools as complex living systems, using software developed by engineers at the University of Bristol, called Perimeta. The aim was to generate a stakeholder owned systems design which was better able to account for the full range of outcomes valued by each school, and for the complex processes which facilitate or inhibit them, thus providing a more nuanced leadership decision-making analytic. The project involved three academies in the UK.

## Keywords

complexity, hierarchical process modelling, leadership, leadership decision-making, management, wider outcomes

## Introduction

The purpose of this article is to describe the rationale and methodology of a pilot improvement research project that applied hierarchical process modelling to schools as complex systems, using software developed by engineers at the University of Bristol, called Perimeta. The aims of the

---

### Corresponding author:

Ruth Deakin Crick, Graduate School of Education and Systems Centre, University of Bristol, 35 Berkeley Square, Bristol BS8, UK.

Email: [ruthdeakinrick@waitrose.com](mailto:ruthdeakinrick@waitrose.com)

project were (1) to create a systems design that accounted for the full range of key outcomes valued by each school (2) to collect evidence of success of the key processes identified in the design – in the form of quantitative, qualitative and narrative data – and (3) to model this using the Perimeta software, which returns visual analytic feedback for leaders against each outcome in the form of the Italian flag. Red represents what is not working, green represents what is successful and white represents what is not known, and is therefore an area for organizational learning and development. The project involved three Oasis Academies in England. The systems design which was developed for this study began with the overall purpose of the academies to facilitate transformative learning and achievement for all students. The stakeholders identified the core processes essential for fulfilling this purpose as: leadership learning, teacher learning and student learning. In the first section of the article we present the rationale for the project, then we explore why complex systems thinking is a useful approach for school improvement and discuss the three core processes that provided the focus for data collection and modelling. We then describe hierarchical process modelling, the Perimeta software, data collection and feedback, before finally discussing critical issues and summarizing next steps.

## **Background and Rationale**

Macbeath and McGlynn (2002) reviewed 30 years of school effectiveness research and described the shift that had occurred from evaluating schools as whole units to a more specific focus on what is happening in individual classrooms. They (2002: 6) argued that this shift should be complemented by a wider focus on school culture:

It is not a matter of either/or: school or classroom, management or teachers, teaching or learning. Measuring effectiveness means sharpening our thinking as to where we should give most attention and invest our energies at any given time and in the light of the priorities we pursue. And as we get better at it we recognize that in good schools the boundaries between different levels become so blurred that they defy even the most inventive of statistical techniques.

They then go on to describe a model of evaluation that puts student learning at the centre but set in the context of a school culture that sustains staff learning, leadership that creates and maintains the culture and an outward-facing dimension involving home and community. They (2002: 7) continue:

In deciding what to evaluate there is an irresistible temptation to measure what is easiest and most accessible to measurement. Measurement of pupil attainment is unambiguously concrete and appealing because over a century and more we have honed the instruments for assessing attainment (and used them) for monitoring and comparing teacher effectiveness.

The distinction between assessment and evaluation is that the latter demands taking a step back from a specific piece of work or programme of study and asking questions that lead to actionable insights and organizational improvement strategies for example: Was the experience worthwhile? What did I learn from the process? What might I do next time to improve? How am I developing as an effective learner? Both assessment and evaluation can be formative or summative – the focus of this study was on evaluation for organizational leadership decision making – hereafter referred to as ‘leadership decisioning’.

Complexity is a core feature of learning communities. A complex system is characterized by uncertainty, that is by randomness, incompleteness and fuzziness (Blockley 2010). This means that we cannot easily predict particular outcomes, since the system contains so many variables and processes which influence each other through feedback loops, and these operate at different levels and through different processes within the system. Some sub-processes of schools, such as time-tabling, or presentation of data, may be predictable and complicated rather than complex when taken in isolation, but they are processes that are nested within, and dependent upon wider human, technical, physical and social processes which constitute the system as a whole (Blockley and Godfrey, 2000). It is the interdependencies between these processes that make the system complex.

An additional factor for schools is that their purpose is *learning* and learning itself is a complex, dynamic process that is inherently unpredictable (Deakin Crick, 2012). In order to learn something a person has, by definition, to not know it already. If we accept that the development of learning to learn is an important student outcome – then the process of handing responsibility for learning to students themselves (rather than simply teaching them to pass the test) inevitably requires that leaders of learning engage with uncertainty. This project focused on the transformative learning of students as the system's purpose – that is student learning which led to significant change in both achievement and aspiration. The project design identified the learning of teachers, leaders and parents/carers as core processes that were critical for the success of student learning. So built into the design was the intention for the academies to become self-organizing learning systems, rather than systems which were defined only by compliance with external regulation. Hence the academy leaders' desire to explore richer ways of evaluating the outcomes that most appropriately achieved their unique purpose.

## Reframing Schools as Complex Living Systems

A complex living system is one which is self-organizing, purposeful, layered, interdependent and operating 'far from equilibrium' (Checkland and Scholes, 1999; Davis and Sumara, 2006). There can be no single blue print for school improvement because each school operates in a unique context – what works in one may literally not work in another and what works in one school at one time may not be repeatable at another time. Leaders therefore need to be able to respond appropriately to their context as 'designers of learning' rather than 'deliverers of pre-determined curricula'. This means they need to be able to be responsive to what emerges at a particular time and place and manage that emergence dependably.

In terms of developing evaluation models for schools, systems thinking and complexity thinking demonstrate that a reductionist focus on the measurement and improvement of a single variable (for example a test result) distorts both the process and the outcome of the system (Assessment Reform Group, 1999; James and Gipps, 1998; James et al., 2007; Reay, 1999). As Mason (2008: 41) argues

trying to isolate and quantify the salience of any particular factor is not only impossible, but also wrongheaded. Isolate, even hypothetically, any one factor and not only is the whole complex web of connections among the constituent factors altered – so is the influence of (probably) every other factor too.

We know for example, that an over focus on high stakes summative testing and assessment not only distorts how teachers teach, but it also distorts student learning and creates an 'own goal'

since it depresses student motivation for learning (Harlen and Deakin Crick, 2003a, 2003b). It does more than simply relegating wider outcomes to second class goals – it actually scores an own goal.

Not only is it important to focus on a range of processes in schools that are deemed to fulfil a particular purpose, but it is also important from a systems perspective to understand the whole, the parts and how they interact. These are described by Goldstone (2006) as ‘contextualized’ and ‘decontextualized’ aspects of a system. In developing contextualized accounts learners and their environments (students, teachers, leaders and organizations) are seen as parts of a single whole. How someone learns depends in part on the larger system in which they learn. Elements of the system cannot be understood independently. Rather, the interactions of the elements give rise to emergent behaviours that would not arise through their independence.

## **Technology, Measurement Models and What Might Be Possible**

The early application of information technology as a tool to support data-gathering and analysis for the evaluation of school performance, combined with a target-driven culture of school improvement, has resulted in an ever more single-minded focus on quantifiable measures of success in schools. Fuelled by political imperatives, and in keeping with the dominant ‘zeitgeist’, this has led to a reductionist focus on test performance at the expense of a wider, more balanced range of outcomes. Davies (2011) raises similar concerns about the current restricted view of ‘success’ from the perspective of strategic leadership. He (2011: 10) suggests that, by focusing on too narrow a range of school performance measures, the upward trajectory of pupil attainment might stall:

Success can be seen in how children achieve academically, socially, spiritually, physically and emotionally; it is enabling children to be all they can be. The difficult question is how do you know that you have been successful? Standardized test scores, even when adjusted for value-added dimensions, tell only part of the story.

However as technology has continued to develop there are fresh opportunities for the representation of complex data and the development of learning analytics – the use of computers and computation to enhance learning. These offer new ways of responding to complexity in learning communities (Buckingham Shum and Deakin Crick, 2012). Combined with insights into systems thinking, systems design and systems modelling these new technologies offer a potentially richer approach to evaluating the wider purposes of schools as complex systems.

While we believe there is an appropriate place for political accountability for schools and, crucially, a common framework of entitlement and equity for all children, we seek a more productive approach for the 21st century based on the collaboration, participation and trust among stakeholders. Our aim is to develop richer ways of measuring and assessing complex processes in learning communities which encourage powerful learning and feedback at all levels of the learning system.

A promising model of school improvement, which draws on systems designing, is Improvement Science which has been developed significantly by the Carnegie Foundation for the Advancement of Teaching (Bryk, 2009; Bryk and Gomez, 2008; Bryk et al., 2010). Sometimes called a ‘Networked Improvement Community’ a school or a community of schools identifies a shared, complex problem which it then analyses and models into a ‘Driver Diagram’ which is almost identical to a hierarchical process model (HPM). Once a measurement model has been established the Networked Improvement Community proceeds through a process of rapid prototyping enquiries, led by teams of teachers in 24-week cycles. While teachers undertake their own evaluation and

analysis of their interventions which are designed to address the chosen complex issue, the whole system commits to a shared disciplined evaluation framework. Meanwhile the collective intelligence which emerges from these prototypes is harnessed by school leadership – and academic critical friends – and shared across the network and beyond.

This approach calls for new ways of using data in practice. Perhaps two of the most powerful affordances of technology today that enable us to address this challenge are its ability to:

1. provide rapid (sometimes instantaneous) feedback for learning to individuals, groups and organizations on measures which are pre-designed to stimulate strategic change;
2. represent complex data visually and more simply, encouraging a holistic approach to ‘decisioning’ for individuals, teams or organizations.

## Measuring What Is not Easily Measured

All this is a particular challenge for schools and groups of schools that have a broader view of education and a desire to extend the measures of school performance beyond the easily quantifiable. These include, but are not limited to, schools underpinned by alternative philosophies such as cooperative schools or human scale schools as well as many faith-based schools. Bryk et al. (1993) point to the need for more studies of the inspirational ideology that animates many Catholic schools. They (1993: 303–4) admit the scepticism that this idea produces, especially in a research culture strongly influenced by secular and positivistic assumptions, but make a powerful case for such research:

Some may question our claim of a causal role for this inspirational ideology . . . unlike the effects of academic organization or school structure, which can be largely captured in regression analysis and effect sizes, estimating the influence of ideology is a more complex and less certain endeavour. Ironically, these effects are harder to study and yet also more pervasive . . . To ignore the importance of ideology because it cannot be easily captured in statistical analysis or summarized with numbers would be a serious mistake. Statistical analysis can help us to see some things but they can also blind us to the influence of factors that are beyond their current horizons.

This project was designed as a proof of concept – an initial test of whether the systems designing, modelling and feedback enabled by the Perimeta tool held promise in schooling systems. The focus was on the value and the application of the concepts and ideas rather than their usability or capacity for scaling up. Because of limitations of space, this paper is restricted to this focus, while identifying issues of usability, reliability, validity and trustworthiness for subsequent papers.

The fundamental questions behind this project were these: if we acknowledge the importance of student attainment but also have the ambition to educate our students for a set of broader outcomes, how can we know how well we are doing and what we might need to do to improve? How can we do this in systematic, sustainable and convincing ways?

## The ECHO Project Case Study

For this case study we worked with Oasis Community Learning (OCL), the charitable trust responsible for the Oasis group of academies in England. The Oasis Education Charter is the

shared stakeholder derived statement about the values, vision and purposes which describe what the academies will be like and the expectations of students' experience and achievement. From the opening of its first academies in 2007, the trust was committed to evaluate the wider outcomes of education, taking into account more easily measurable and important aspects, such as attendance and exam results, but with the addition of qualitative evidence about students' achievement and their personal experiences of learning at an academy. The trust also has commitments to care for OCL staff, helping them to develop and flourish, and to engage effectively with the wider community, particularly parents and carers. It believes that one of the most important ways in which it can prepare students for life after the academy is by nurturing confident, resilient and caring life-long learners. To do this, teachers and leaders need to model such learning and parents and carers must be fully involved with the education of their children and feel welcome at an Oasis Academy.

The 'Evaluating CHarter Outcomes' (ECHO) Project, which began in 2011/12, was the vehicle for this proof of concept study, focusing initially on one section of the charter – *transforming learning*, which, it was presumed, would then have an impact on the other two sections of the charter, *Transforming Lives and Transforming Communities*. The project worked with three secondary, co-educational academies located in areas of disadvantage in the south of England, each with approximately 1100 students and 60 staff.

## Generating a Systems Design for the Three Academies

The following section describes the process of generating the design for the case study, proceeding as it did from an analysis of the whole system to a process design to inform an improvement protocol. The first phase of the project was to build a 'systems architecture' (Sillitto, 2014) for the group of academies. A 'systems architecture' is the fundamental conception of a system which sets out what the parts of the system are, what they do and how they fit and work together' (Sillitto, 2014: 4). It can be produced as a visual representation which describes the key processes of the system. A system is defined by its purpose (Blockley and Godfrey, 2010) and the espoused purpose for these Oasis Academies was to provide education which was 'transforming learning for all members of the community including parents and carers'.

One common form of 'architecture' is a HPM, which identifies the core processes in a system that contribute to its purpose in the form of a hierarchy of processes and sub-processes. Generating a HPM involves stakeholders determining what are the critical organizational processes necessary for the achievement of the system's purpose. Stakeholders include all participants in the system – users and providers. Based on the Oasis Education Charter, the team identified the following three core processes:

1. leaders learning and facilitating learning in school and community;
2. teacher professional learning;
3. students taking responsibility for their own learning and achievement.

These core processes enabled the stakeholder team (school leaders and researchers) to construct a HPM (Figure 1). The leadership process was decomposed into two one of which focused on leaders learning and the other focused on leaders engaging parents and carers in the learning of the academy.

<b>1. Outcome</b>	<b>1.1 Establishing and sustaining a group of high achieving learning communities that enables everyone to realise their full potential and refuses to put limits on achievement</b>			
<b>2. Processes</b>	2.1 Developing the learning of students so that they realise their full potential	2.2 Developing the learning of teachers so that they realise their full potential	2.3 Developing the learning of leaders so that they lead the learning of teachers and students effectively	2.4 Engaging parents/carers effectively in the learning activities of the Academy and in supporting the learning of their children
<b>3. Outputs</b>	<p>3.1 Feeling valued and included in our community</p> <p>3.2 Feeling part of a community that focuses its activities on learning and achievement</p> <p>3.3 Knowing that I am becoming an effective learner</p> <p>3.4 Knowing that we are being helped to reach our full potential</p> <p>3.5 Maintaining a good rate of progress at key transitions, for example, primary to secondary school</p> <p>3.6 Achieving results that meet my aspirations and expectations</p> <p>3.7 Feeling that my successes are recognised</p> <p>3.8 Contributing actively to a community which focuses on service to others</p> <p>3.9 Contributing actively to a community where shared leadership is promoted</p> <p>3.10 Providing evidence that training and development opportunities have helped me to improve my classroom</p> <p>3.11 Contributing actively to a community that learns together, shares what works best and knows what it</p> <p>3.12 Feeling included and involved with my children's education at the Academy and knowing that there is an open door</p> <p>3.13 Being effective role models as leaders of learning</p>			
<b>Sources of evidence</b>	Students	Teachers	Leaders	Parents/carers

Figure 1. Hierarchical process model design.

### Defining the Measurement Model

In order to develop a precise measurement model the four core processes were decomposed to a third level in the hierarchy in 13 sub-processes or outputs. These described the sorts of measurable experiences, behaviours and values that students, teachers, leaders or parents/carers might manifest if the academy was being successful at level two.

Next the team identified nine strands of data to be collected as evidence of performance that mapped onto these 13 level three processes. The types of data included research-validated questionnaires, new questionnaires, narrative interviews and existing school data on performance and attendance. The questionnaires had response values as four-point Likert-type scales. The narrative interviews were conducted according to a pro-forma by senior teachers with 20 randomly selected students. They focused on students' stories of significant change in the academy and were analysed by two researchers and rated according to the degree to which the students reported personal transformation.

These nine data collection points were operationalized in each academy over a period of one year. Questionnaire data was captured online using the academies' intranet systems, while performance data was exported from the academies' management information systems. These were integrated and entered into the Perimeta software. The research ratings of student interviews were entered manually. A summary of the data strands and sources of evidence can be viewed in Appendix A.

### Learning as a Critical Viewpoints at Different Levels of System Design

In this section we explore how the process of learning is multi-layered, how timely and useful feedback is critical for learning, how learning is a dynamic process and how, by definition, this makes managing uncertainty and emergence a critical leadership skill for schools as complex systems.

### *Viewpoint 1: Leaders Learning*

In their conclusions to an extensive international survey of educational leadership and management, Davies and West-Burnham (2003) argue for new models of leadership that are fundamentally concerned with strategy, values and learning. Beare (2001), looking to the future, suggested that schools would need a focus on learning as their prime mission and professional leadership where the leader and leadership teams give highest priority to the professional purpose of the school, personally and frequently monitoring learning programmes, putting time and energy into organizational learning and improvement, giving support to the staff involved with learning programmes and putting tangible emphasis on instructional leadership.

The relationship between school leadership and learning has been explored in several studies (Day et al., 2010; National College for School Leadership, 2004), which all concluded that the most successful systems, based on measures of student engagement and attainment, prioritized staff motivation and commitment, teaching and learning practices and developing teachers' capacities for leadership. In these systems, the headteacher was the 'leader of learning' and, by having a key role in developing the school's culture and climate, was the second most important influence on student outcomes after quality of teaching.

In a comprehensive study of leadership effects on student and organizational learning, Silins and Mulford (2001) found that student outcomes are more likely to improve when leadership is distributed throughout the school community and when teachers are empowered in their spheres of interest and expertise. The emphasis on empowerment brings the concept of uncertainty to the fore: to empower someone is to allow them to pursue their professional purpose and this may take unexpected directions. When a teacher or a student embarks on a learning journey it is not possible to completely predict what is going to be learned, or achieved, or even how it will be learned. This is particularly the case when the focus is on 'transformative' learning – learning that is authentic and meaningful to the learner and results in personal and professional change.

In Bryk et al.'s (2010) research, the most effective school leaders were catalytic agents for systemic improvement, synchronously and tenaciously focusing on new relationships with parents and community; building teachers' professional capacity; creating a student-centred learning environment and providing guidance about pedagogy and supports for teaching and learning. These leaders focused on the processes of learning itself at all levels, rather than controlling outcomes. They were comfortable with uncertainty while tenaciously focusing on learning. Goldspink's research identified that the leadership qualities required for complexity are not among the typical selection criteria for principals. They include: a level of modesty and circumspection and a capacity to question one's own deepest assumptions while inviting others to participate in critical enquiry (Goldspink and Kay, 2007).

What is clear from these studies are the interdependencies between the 'layers' of leadership learning, teacher learning, student learning and sustained organizational improvement. These interdependencies operate through feedback loops – energy and information that circulates throughout the system. What emerges from these interdependent processes may be able to be managed responsively by wise leadership but cannot be completely predicted or controlled.

### *Viewpoint 2: Teachers as Learners*

Collaborative, classroom-based, research-informed professional learning and enquiry in schools represents an important facet of the 'deep structures' of organizational learning. Professional



learning is a vital pre-condition for school improvement through its positive influence on teachers' classroom practices and their students' learning. Schools' promotion of teachers' participation in school-to-school and other networking activity can support improvement-related activity in schools and classrooms through enabling teachers to engage with an expanded pool of practice ideas, resources, sources of support, increased opportunities for mutual problem solving, knowledge creation and transfer (Jackson and Temperley, 2007; Little and Veugelers, 2005; Stoll and Louis, 2007).

Misunderstanding the nature of teacher learning by underplaying its complexity leads to a focus on the micro-context (individual teachers or individual activities or programmes) to the exclusion of influences from meso (institutional) and macro (school system) contexts (Bore and Wright, 2009). Adopting a complexity thinking perspective, we assume that teacher learning does not emerge as a series of isolated events but simultaneously in the activity of autonomous entities (teachers), collectives (school phase and subject groups) and subsystems within grander unities (schools within school systems within socio-political educational contexts). These nested systems and subsystems associated with teacher learning are interdependent and reciprocally influential. As a result, to explain teacher professional learning, one must consider what sort of local knowledge, problems, routines, and aspirations shape and are shaped by individual practices and beliefs. How are these then framed by the other systems involved? Thus we construe teacher learning as a complex process representing recursive interactions between systems and elements that coalesce in ways that are unpredictable and uncertain (Clarke and Collins, 2007).

Complex systems need to be off balance in order to move forward. Wheatley (1999) notes that organizational equilibrium is 'a sure path to institutional death'. Helping schools and groups of teachers become aware of the full range of dissonance between their values and practices in relation to teachers' and school learning and leadership was an effective intervention for promoting change, growth and deep learning as part of the Learning how to Learn project (Pedder, 2006, 2007; Pedder et al., 2005). Further mapping of patterns of dissonance in schools nationwide was an important feature of the State of the Nation CPD study (Opfer and Pedder, 2011). Dissonance between what teachers consider important for enhancing the quality of their students' learning opportunities and perceptions of current practice may result in what Woolfolk et al., (2009) refer to as 'change-provoking disequilibrium', further underlining the practical significance of attending to relationships between values and practices. Argyris and Schön's (1996) work on 'theories of action' illustrate that dissonance serves as a catalyst for schools to attempt to change their environment in ways that better support learning. Dissonance and disequilibrium are forms of uncertainty – requiring the management of emergence and an accounting for interdependencies within the system. Learning proceeds through uncertainty and what is not known – and school evaluation models need to be able to account for this.

### *Viewpoint 3: Students as Learners – Deep Learning as A Complex Psycho-social System*

Deep learning occurs when students choose to invest in processes of learning that are authentic, personally owned and lead to the construction of new knowledge in pursuit of a chosen purpose (Bereiter and Scardamalia, 1989; Bateson 1972; De Jaegher and Di Paolo 2007; Dweck 2000). The Learning Futures research (Deakin Crick et al., 2010, 2011) identified 'authenticity, agency and identity' as key elements of pedagogy which lead to engagement and depth in learning, rather than superficial recall or performance orientation. Where these occurred, students described their

learning as transformative for them as individuals – they were authors of their own learning journey, in a process of ‘becoming’ (Seely Brown and Thomas, 2009).

In a major re-modelling of 15 years of data on learning dispositions, Deakin Crick et al. (2015) demonstrated the complex nature of learning in terms of several relationships: that of the learning agent with the self; the interrelationships between the learner’s internal learning processes (‘intra-personal’); the relationships between learners (‘inter’ personal); and the relationship between learners and their contexts (inter-contextual). It has enabled the development of an agency-based concept of learning in a complex social ecology, where resilience in learning is developed and achieved through mindful agency (Deakin Crick et al. 2015).

Research in interpersonal neuro-biology has highlighted the connections between the brain, the mind and inter-personal relationships and offers insight into the complex nature of feedback loops between and within people in organizations. Siegel (2012: 3) argues that ‘a core aspect of the human mind is an embodied and relational process that regulates the flow of energy and information within the brain and between brains’ and that ‘the mind is an emergent property of the body and relationships . . . created within internal neurophysiological processes and relational experiences’.

Energy and information flow is what is shared between people within organizations. It is also the ‘subject matter’ of learning. How a person-in-relation regulates that flow of information and energy – how they select what information matters, make sense out of it and apply it to their purpose – is a process of learning in any domain that is inherently uncertain and complex. Measuring such learning at different levels in the system, in order to improve, requires more than simply reductionist measures of student test scores.

Learning – mindfully regulating the flow of energy and information over time in the service of a purpose of value (Deakin Crick et al. 2015) is core process or property of all stakeholders in a learning community – students, teachers, leaders and parents/carers. Organizational learning emerges from the alignment of purpose of all stakeholders (Tracy, 2014). Making learning and the alignment of purpose visible in school improvement processes is a sort of ‘double loop’ learning that is necessary for improvement and change regardless of its focus. Developing a systems architecture and a measurement model to reflect and empower this process was the challenge for this proof of concept project.

## **Introducing Perimeta as a Tool to Visualize and Evaluate a Hierarchical Process Model**

The Perimeta software was developed by Davis et al. (2010) at the University of Bristol as a tool for systems architecting. It supports the collaborative development of solutions to complex problems by providing a highly visual interface for understanding complex cause-and-effect relationships and complex evidence. The visualization provided by Perimeta is structured as a HPM. Perimeta can be described as being:

- a learning analytic designed to model diverse and complex processes;
- driven by stakeholder purpose;
- capable of dealing with hard, soft and narrative data in evidence of success, failure and ‘what we don’t know’;
- a visual environment for sense-making in complexity;
- a framework for self-evaluation and dialogue.

## Summary of the Perimeta Methodology

The approach to the research design and methodology for this pilot project followed the process of systems designing developed in the Systems Centre at the University of Bristol (Blockley, 2010, Blockley and Godfrey, 2000). First, the system boundaries of the academies were defined by first identifying their purpose – all stakeholders in that purpose were included in the system. Next a rich picture was elicited about the system, which informed a systems analysis and the identification of the critical processes for achieving the system's purpose. These informed the HPM design and level three sub-processes provided outputs for the measurement parameters – the ways in which the outputs could be evidenced. Data were then systematically collected and entered into the Perimeta software for modelling. The software accepted a range of types of data and its juniper algorithms based on interval probability theory (Davis et al., 2000) provided evidence about what was working well – what was working negatively and what was not known. This was returned to the stakeholders as a rich, visual analytic dashboard to be used for decision making, celebration and improvement.

## Modelling Uncertainty

The use of mean scores often masks the contextual variations and differences within particular settings. The degree of variation around the mean indicates a level of uncertainty. In contextualized and complex situations it is this uncertainty that is of interest since it indicates where improvement strategies can be directed (Bryk et al., 2015). With any data set the degree of certainty upon which decisions can be made will vary and the skill of deriving actionable insights from complex data is a key leadership attribute, particularly in high stakes, complex settings such as schools. Recognizing the uncertainty inherent data the Perimeta model adopts an 'Italian flag' graph to represent the quality of all of the evidence. The example below, focusing on students' perception of their learning power, shows how this consists of:

- 'green' representing the strength of evidence of positive self-perception of learning;
- 'red' representing the strength of evidence of negative self-perception of learning; and
- 'white' representing lack of evidence, or uncertainty in self-perception of learning. (In this black and white version the colours are represented by G, R and W.)

Figure 2 below gives an example of this in responses to a self-report questionnaire about students' learning power, based on a four-point Likert-type scale. Rather than return a simple mean score for a group, the visualization indicates the degree of uncertainty there is in the data, as well as what can be interpreted with more certainty – either negatively or positively according to the purpose.

## From Likert-type Scales to Perimeta Modelling

Uncertainty is introduced in the evidence from questionnaires and interviews by answers that are not strongly positive or strongly negative. Given a scale from 'strongly disagree' to 'strongly agree', the intermediate points of 'disagree' and 'agree' entail some contrary evidence and some uncertainty because by definition the respondent is less sure of their response. Another factor in making sense of the evidence is the degree of confidence we have in the respondents and in the context in which they answer the questionnaire. A number of considerations are relevant, for example the respondent's understanding of the question and experience to answer it, the size of the sample and so on.

Categories of evidence in Italian Flag	Evidence of success	Lack of evidence	Evidence of failure
Application to learning community	Evidence of positive self-perception of learning power	Uncertainty in self-perception of learning power	Evidence of negative self-perception of learning power
Example: High self-perception of learning from high proportion of 'strongly agree' responses			
Example: Medium self-perception of learning from high proportion of 'agree' responses			
Example: Medium self-perception of learning from high proportion of 'disagree' responses			
Example: Low self-perception of learning indicated by high proportion of 'strongly disagree' responses			

Figure 2. Illustration of Italian flag graphs in the context of a learning community.

In this study the questionnaire data were collected with Likert-type responses. These were translated into the Italian flag figures through a two-stage linguistic technique developed by Hall et al. (2004). This technique required two judgements to be made on each outcome measure – the actual performance measured and the degree of confidence in that performance. For example, with a small sample of students whose mean score on a scale was 60% we may form a view that we actually have low confidence in that performance measure, for a variety of reasons such as those described above. By using ‘performance’ and ‘confidence’ as two axes it is possible to position such a score on a map that accounts for these two judgements and this can then be converted into an Italian flag figure which models the uncertainty about the judgement.

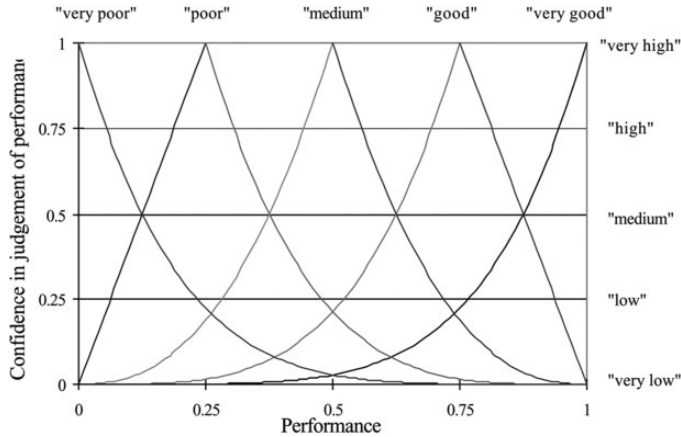
Figure 3 reproduces Hall et al.’s (2004) illustration of the mapping and Figure 4 presents an interpretation of their mapping for a set of 25 discrete combinations of performance and confidence scales in which the best possible (100%) performance was full green (respondent strongly agrees, and very high confidence in the respondent) and the worst possible performance was mostly red (respondent strongly disagrees, and very high confidence in the respondent).

### Modelling Process Performance in the Case Study Example

Process performance functions were created by combining the Likert conversion of Figure 4 with the process performance measurement scales of Figure 5. At the input level, respondents’ raw scores (from 1 to 4) for each question were converted into Italian flag figures of merit as described in the process above. For output processes (collated by question, by gender, by academy and overall) the definitions of best and worst performance and all points in between were judged on a similar scale.

### Modelling the Significance of Inter-process Relationships

This step in the modelling accounts for the feedback loops and influences between each process. For example a student who was disengaged and felt excluded is unlikely to be a source of positive



**Figure 3.** Hall et al.'s (2004) mapping from linguistic descriptions of 'performance' and 'confidence in judgement of performance' to interval values of performance.

Confidence judgment	Rating					
Very high	100%	(0.00,0.00)	(0.25,0.25)	(0.50,0.50)	(0.75,0.75)	(1.00,1.00)
High	75%	(0.00,0.05)	(0.20,0.30)	(0.42,0.58)	(0.70,0.80)	(0.95,1.00)
Medium	50%	(0.00,0.10)	(0.10,0.40)	(0.38,0.62)	(0.60,0.90)	(0.90,1.00)
Low	25%	(0.00,0.25)	(0.05,0.50)	(0.30,0.70)	(0.50,0.95)	(0.75,1.00)
Very low	0%	(0.00,0.60)	(0.01,0.80)	(0.05,0.95)	(0.20,1.00)	(0.40,1.00)
	Rating	0%	25%	50%	75%	100%
	Performance judgment	0	1 – strongly disagree	2 – disagree	3 – agree	4 – strongly agree

**Figure 4.** Conversion of scores from Likert scale to Sp and Sn values of evidence.

evidence for a teacher about how their practice is improving – or leaders who are not learning are unlikely to create a culture where teachers are learning. The significance of pairwise relationships or links between 'parent' and 'child' parameters/processes defining 'causes' and 'effects' in the hierarchy was modelled in an integrated approach using three detailed attributes of sufficiency, necessity and dependency defined by Davis and Fletcher (2000):

- The **Sufficiency** or relevance of the evidence to its parent process is judged as a single number in the [0,1] range;
- A sub-process is a **Necessity** if the parent process cannot succeed without it. Consequently, in the event of failure of the sub-process, the parent process fails; and

ID	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
										ID Number	Name	Description	Combined evidence sn	sp	FOS	UNC	NEG
1										10	54 LearnIt 0		0.845	0.961	0.85	0.14	0.02
2										11	ACADEMY 0		0.845	0.961	0.85	0.14	0.02
3	STRAND 4					TEMPLATE DAT15				12	ACADEMY 0		0	1	0.00	1.00	0.00
4	TITLE					Nmax 200				13	ACADEMY 0		0.943	0.996	0.94	0.05	0.00
5	STAKEHOLDER					Qmax 65				14	GENDER 1 0		0.887	0.996	0.89	0.11	0.00
6	YEAR									15	GENDER 2 0		0.358	0.915	0.36	0.56	0.09
7	PURPOSE					This questionnaire will help us to find out about your experiences of education at an Oasis Academy bet:				101	54 1 1. I was 1		0.889	0.996	0.89	0.11	0.00
8	ACADEMIES (A)	2								102	54 1 2. I felt p		0.888	0.996	0.89	0.11	0.00
9	QUESTIONS (Q)	30								103	54 3 3. I felt s		0.887	0.996	0.89	0.11	0.00
10	QUESTIONS (Q)	30								104	54 4 4. I enjoy		0.828	0.974	0.83	0.15	0.03
11	RESPONSES (N)	91								105	54 5 5. I did a		0.883	0.994	0.88	0.11	0.01
12										106	54 6 6. My bet		0.862	0.987	0.86	0.13	0.01
13	ACAD data ex: blanks	91	0	<total .. blank	zero .. A1 .. A3	0	24	0	67	107	54 7 7. I am ic		0.496	0.825	0.50	0.33	0.18
14	GEN data ex: blanks	91	0	<total .. blank	zero .. G1 .. G2	0	48	48	48	108	54 8 8. I have		0.871	0.99	0.87	0.12	0.01
15	RESPONSE data ex: blanks	2723	42	<total .. blank	zero .. 1 .. 2 .. 3 .. 4	0	31	244	1732	109	54 9 9. My su		0.887	0.996	0.89	0.11	0.00
16	NODES used ex: labels	2765			2765 N=Q=K=A+G+1					110	54 10 10. I was		0.981	0.999	0.98	0.02	0.00
17	LINKS used	10920			10920 N=Q=4					111	54 11 11. I deve		0.888	0.996	0.89	0.11	0.00
18	RESPONDENT RELIABILITY	50			(sample for Respondent N1)					112	54 12 12. The Ai		0.871	0.99	0.87	0.12	0.01
19	IMPORTANCE	1			(sample for N1 Q1)					113	54 13 13. I supp		0.611	0.986	0.61	0.38	0.01
20	SUFFICIENCY	0.4			(sample for N1 Q1)					114	54 14 14. I was		0.645	0.98	0.65	0.14	0.02
21	NECESSITY	0.3			(sample for N1 Q1)					115	54 15 15. I feel a		0.489	0.832	0.49	0.34	0.17
22	RUN FILE				C:\Users\ONS817\Documents\Clients\tranz\PS\54 Students (OAT15)023A.xlsx					116	54 16 16. My te		0.896	0.999	0.90	0.10	0.00
23	DASHBOARD									117	54 17 17. My te		0.873	0.99	0.87	0.12	0.01
24										118	54 18 18. I was		0.889	0.996	0.89	0.11	0.00
25										119	54 19 19. My pa		0.626	0.974	0.63	0.15	0.03
26										120	54 20 20. My pa		0.668	0.969	0.67	0.12	0.01
27										121	54 21 21. Senio		0.873	0.99	0.87	0.12	0.01
28										122	54 22 22. Everyt		0.826	0.973	0.83	0.15	0.03
29										123	54 23 23. The Ai		0.873	0.991	0.87	0.12	0.01
30										124	54 24 24. There		0.889	0.996	0.89	0.11	0.00
31										125	54 25 25. There		0.885	0.995	0.89	0.11	0.01
32										126	54 26 26. I enjo		0.888	0.996	0.89	0.11	0.00
33										127	54 27 27. The Ai		0.896	0.999	0.90	0.10	0.00
34										128	54 28 28. I have		0.648	0.961	0.65	0.13	0.02
35										129	54 29 29. I feel i		0.865	0.988	0.87	0.12	0.01
36										130	54 30 30. I coul		0.845	0.981	0.85	0.14	0.02
37																	
38																	
39																	



Figure 5. Sample of a Perimeta model datasheet.

**Table 1.** Sufficiency, necessity and dependency of relationships between processes.

'Effect' process	'Cause' process	Sufficiency	Necessity	Dependency
Overall performance	Responses to statements in questionnaires	0.3	0.4	1.0
Gender performance	Responses to statements in questionnaires	0.3	0.4	1.0
Academy performance	Responses to statements in questionnaires	0.3	0.4	1.0
Question performance	Responses to statements in questionnaires	0.3	0.4	1.0

- **Dependency** is the degree of overlap between sub-processes and describes the degree of commonality in the sources of evidence'.

On the basis of experience in modelling 'many to one' performance relationships and researcher understanding the values indicated in Table 1 were assigned to the data.

## Estimating System Performance

Full system models of the features of each of data strands 1 to 7 were developed in the Perimeta toolkit, combining the following features defined above:

- A hierarchy linking the responses to questionnaire statements (input processes) in turn to output processes representing the performance by question, by participating academies, by respondent gender (where given) as well as overall;
- Responses to questions using Likert ratings from 1 ('poor' or 'strongly disagree') to 4 ('very good' or 'strongly agree');
- Process performance functions using linguistic measures related to Likert rating scales; and sufficiency, necessity and dependency ratings for each cause and effect relationship.

The Perimeta models used the 'Juniper' algorithm to propagate the evidence and provide estimates of output performance by question, by gender, by academy and overall. A full explanation of the Juniper algorithm is beyond the scope of this article but it is described in detail by Davis and Fletcher (2000).

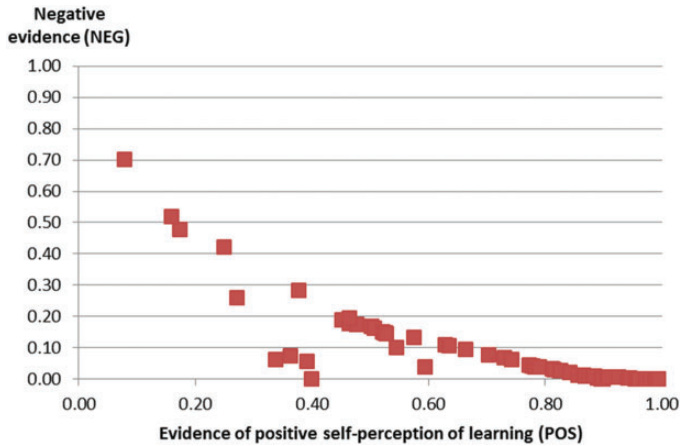
The Perimeta models were each configured to produce a dashboard summary view as well as tabulated results for each question, each academy, each gender and overall. A sample datasheet from a Perimeta model is reproduced in Figure 5 below.

## Example of Findings from Perimeta Model

In this section an example is provided from one of the strands of evidence: Year 8 and Year 10 students' perceptions of their learning experiences, teaching, and progress within the academy.

### *Perimeta Modelling of Students Perceptions of Learning from Strand 2*

The Perimeta modelling of Strand 2 revealed key insights into the self-perception of students in Years 8 and 10 of their development as learners. Figure 6 illustrates the distribution of all responses for Strand 2 according to the strength of evidence of self-perception of learning. The horizontal axis represents positive evidence (POS) and the vertical axis represents negative evidence (NEG),



**Figure 6.** Strand 2 – Distribution of responses.

so that the more uncertain (UNC) responses would appear towards the lower left of the graph. This indicates both the very wide range of strengths of perceptions (10% to 100%) and in the degree of uncertainty (0 to 60%). The results show a very strong bias towards positive perceptions, with average positive evidence (73%) outweighing average uncertainty (20%) and average negative evidence (7%).

There were clear differences between the participating academies and between the genders in the strength of self-perceptions of learning. The particular statements that students most readily agreed with, or were uncertain about, were substantially different between males and females.

Figures 7 and 8 demonstrate the Perimeta dashboard for students by gender and by academy for Strand 2. This example shows data for two of the three academies. This dashboard is one example of one particular query available to leaders wanting to explore evidence about progress towards the overall purpose of transforming learning. Leaders would ‘drill down’ into their Perimeta model for particular reasons in particular contexts, to raise questions, to understand the context and to inform strategy. Where there is greater uncertainty, that strategy might well be to gather more evidence and to revisit the Perimeta model. The scope of this article does not make it possible to provide greater detail – but each question, each student, each gender and each academy is represented in this model.

## Comparing Two Types of Analysis

Results were broadly consistent between the analysis using the Perimeta modelling and the descriptive statistics particularly when comparing standard deviation scores with the degree of uncertainty indicated by the Perimeta modelling. This comparison helped to build confidence in the Perimeta analysis. However, validation of results in the conventional sense proved to be challenging because the Perimeta modelling was not a substitute for more familiar statistical techniques but rather a complement to them. Detailed reviews with key stakeholders including members of the ECHO project team and senior Oasis Academy leaders were necessary to build confidence in the approach and its results, focusing on the following.



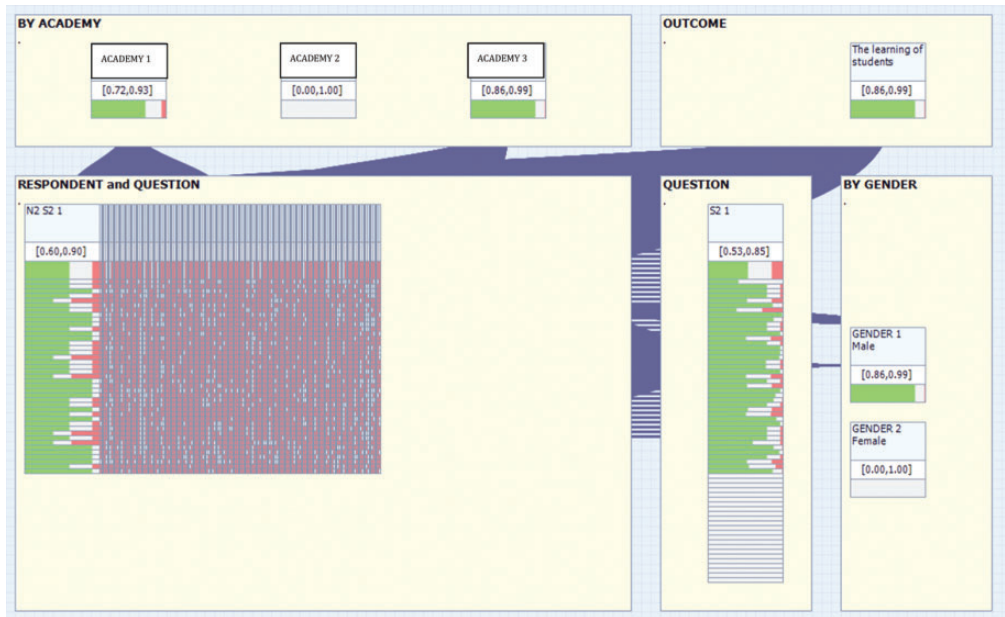


Figure 7. Male students dashboard: Strand 2.

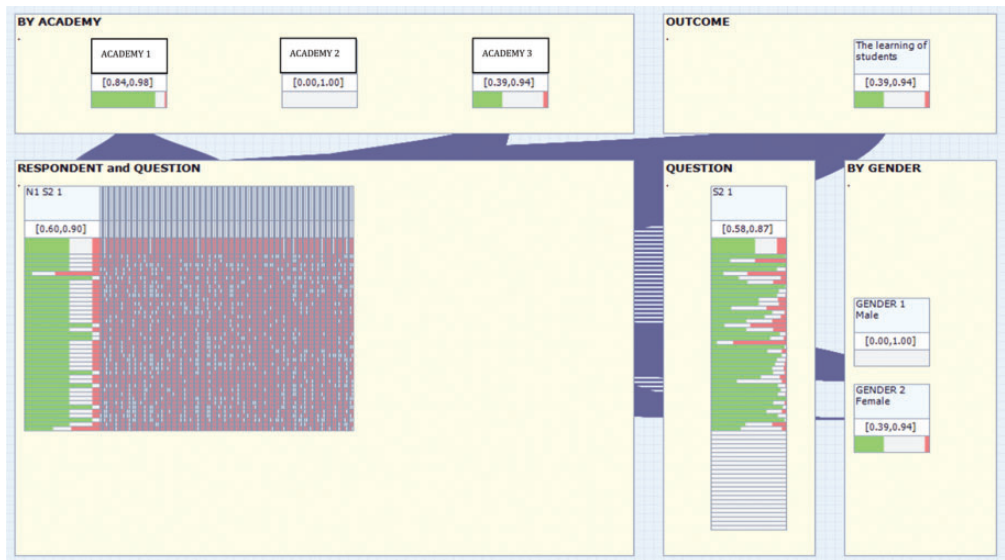


Figure 8. Female students dashboard: Strand 2.

- Understanding the theoretical difference between uncertainty and statistical variation.
- Recognizing the existence of uncertainty in the evidence provided by student responses.

- Appreciating that insights into uncertainty as well as positive and negative evidence of self-perception of learning could enhance the development of improvements in student experiences in Year 7.

Using Perimeta modelling uncertainty was recognized clearly in the responses from students about their experiences of learning in the Academy. Uncertainty was demonstrated to be a substantial factor accounting for around one fifth of the evidence obtained. Table 2 provides one example of a ‘drill down’ into the data looking at students reporting of their learning experiences in the academy.

## Deriving Actionable Insights from the Perimeta Dashboard

Having produced the Perimeta analysis and dashboard, academy leaders were able to interrogate the visualized data, exploring its meaning, identify possible explanations for the variations in the responses and make sense of it in terms of academy level interventions and change programmes. Technically it was possible for leaders to explore many different visualizations of the same data at differing levels of granularity. The following examples present a sub-set of insights which could be acted upon at different levels in terms of leadership and professional enquiry: further qualitative investigation into the uncertainty of students about feeling ‘safe’ in the academy; change strategies for supporting male students’ capability and willingness to get help in their learning when they needed it; insights into the relationships between student engagement in learning and particular approaches to curriculum. Table 2 below identifies a number of areas for further investigation by presenting the individual questions and their ‘Italian Flag’ status.

## Practical and Technical Challenges

The project was demanding in terms of resource. The software was not web based and there was no automated procedure for converting the data into the Italian flag nomenclature and entering it into the Perimeta model. In addition each academy collected the data within their own academy intranet and it was then manually integrated into the Perimeta model – there was a significant attrition rate because of this. Any development of this HPM for school self-evaluation would require significantly more automation than was possible in this pilot. A web-based version of Perimeta would enable the whole project team – practitioners and researchers – to utilize the dashboard as the project developed. It would also house more data. A rapid analytics platform, such as that being developed by the Learning Emergence research network with eXplorance,<sup>1</sup> capable of delivering surveys across organizations, while providing rapid feedback to individuals and teams, and exporting data into the Perimeta software would make the design and delivery of the measurement model achievable through a planned process over a school year. In principle a leadership team would design the model according to their particular shared purpose and be able to explore the Perimeta dashboard for leadership purposes as and when it was required. This could work across a group of schools formed as a Networked Improvement Community (Bryk et al., 2015). The automation of the software would significantly reduce the resource implications for schools.

Two further limitations of this approach are (1) the regulatory frameworks which constrain what is measured and valued leaving little scope for in school improvement and (2) the capability of both the research and professional community to understand and respond to data in new ways.

**Table 2.** Strand 2 – Statements for further investigation in Phase 2.

Q	Statement	Gender	Academy
Evidence of positive self-perception of learning > 75%			
S2 4	I like learning new things	G2	A1, A2, A3
S2 35	We are encouraged to care about the needs of other people in the local community and around the world	G2	A1, A2, A3
S2 9	I am confident in my learning	G2	A1, A2, A3
S2 2	I feel safe in the academy	G2	A1, A2, A3
S2 9	I am confident in my learning	G1	A1, A2, A3
Uncertainty in self-perception of learning > 50%			
S2 2	I feel safe in the academy	G1	A1, A2, A3
S2 31	I feel that my successes with activities outside the academy are recognized	G2	A1, A2, A3
S2 7	I like working with other students to help my learning	G2	A1, A2, A3
S2 42	I have had opportunities to lead other activities at the academy	G2	A1, A2, A3
Evidence of negative self-perception of learning > 25%			
S2 22	I get extra support to help my learning when I need it	G2	A1, A2, A3
S2 6	I am hard to distract	G2	A1, A2, A3
S2 18	I often have conversations with my teachers which help me to make better progress	G2	A1, A2, A3
S2 3	I feel valued and cared for as an individual in the academy	G2	A1, A2, A3
S2 8	I don't distract other students	G1	A1, A2, A3
S2 14	I come up with new ideas to help my learning	G2	A1, A2, A3
Strong/common themes in responses			
Student: confidence, learning from mistakes, learning new things, pride, responsibility			
Teacher: quality of teaching, expectations, sharing learning			
School/ leader: feeling safe			
Community/ parent: recognition of activities outside			

(Yeager et al., 2015) call for new research on methods to create and embed what they describe as 'practical measures' in networks of practitioners and researchers engaged in improvement research. They (2015: 39) conclude:

Educators need to be able to assess whether the instruction they deliver in a classroom is, in fact, leading to the changes they hope for, in real time, well before students become academic casualties. Although measurement for accountability is important for signaling a problem, relying on such measures for improvement is analogous to standing at the end of the production process and counting the number of broken widgets. The quality of the end product is an aggregate consequence of many discrete processes that operate within a complex production system. Quality improvement entails deeper information about system processes, where undesirable outcomes stem from, and targeting subsequent improvement based on this knowledge. Seeking to remediate the problem at the end of the line is neither an effective nor efficient solution. Educators need both more frequent data and also different kinds of information than they normally get—measures that can help them improve their actual practices.

The pilot study produced surveys that could be significantly and productively reduced in size, with new variables computed from validated scales which would both reduce the amount of data to

be entered into the Perimeta model while ensuring that it has more power as a measurement tool, rated with a higher confidence score in the translation from descriptive statistics into the Italian flag model

This study indicated that the question which we set out with was valuable and worth pursuing. The proof of concept was limited in its stakeholder engagement and in its feedback to a range of practitioners. However the findings were presented to the national leadership team of the multi-academy trust who were positive about the potential of the approach and its value for a Networked Improvement Community. It was clear that data of this type, aligned to Oasis Community Learning's unique purposes and feedback in a timely and usable manner, would be of significant value. The key ideas underlying Perimeta are those of responsible self-organization, local diversity, participation and seeking to manage emergence dependably. Perimeta does not 'do away' with the need for external regulation – rather it makes it possible for schools to 'comply' with the regulatory framework while focusing, at the same time, on the contextual issues which are of most concern locally.

By seeking to model 'uncertainty' Perimeta valorizes professional learning. What we do not know may be as important as what we do know and by definition the visualization of the Italian flag invites 'leadership decisioning'. Actionable insights derived from the model may be best addressed through professional enquiry and prototyping interventions designed to improve the issues which it flags up. Having an ongoing data collection process in place aligned to an agreed measurement model in turn may add rigour and discipline to professional enquiry.

These ideals are a long way beyond the completion of this proof of concept study. However, conceptually, practically, professionally and technically there is promise in this approach.

A unifying idea behind the project is the inter-relationships and interactions between the learning of leaders, parents/carers, teachers and students as they influence each other, and how these can be harnessed within a shared complex system, the purpose of which is to enhance the learning and achievement of students. The measurement model and analysis offered by Perimeta enables leaders to identify areas of uncertainty during the process of self-evaluation as well as aspects of performance that are more certain, or red and green to use the Italian flag metaphor. Clearly weaknesses (the red indicators) need to be addressed but schools will also benefit from exploring areas of uncertainty in more depth. Perimeta is designed to empower stakeholders to take responsibility for their own purposes and progress through designing and responding to relevant data, representing measurement models allied to overall purpose. It is essentially a decision-making tool, measuring in order to improve. It measures for the purpose of stimulating positive change according to locally defined need.

## **Summary and Conclusion**

The Perimeta software is a learning analytic designed to model diverse and complex processes driven by stakeholder purpose. It is capable of dealing with quantitative, qualitative and narrative data in evidence of success, failure and what is unknown. It provides a visual dashboard for making sense in the complex task of school leadership and it provides a framework for self-evaluation, dialogue and informed decisioning. It provides a visual environment for making sense in conditions of complexity and a framework for self-evaluation and dialogue

This project was a proof of concept: the beginning of a much longer journey. The main challenge theoretically has been in the paradigmatic worldview shift that is required in moving from a traditional social science approach to one informed by complexity thinking. The concomitant practical challenge how to measure and evaluate a wider set of student outcomes. Philosophically this shift requires a participatory paradigm which includes a critical realist epistemology, a theory of learning as becoming – rather than either acquisition or participation, and an approach to society which facilitates profound diversity, respecting community and tradition within a common framework. Methodologically the challenges of mixed methods are well known and the demands of inter-disciplinarity (for example, engineers working with educators and engaged research, theorists and practitioners) require time and careful listening. This represents the biggest challenge for researchers as much as school leaders in pursuing this agenda.

Complex data collection, representation and interpretation challenge traditional approaches in which a ‘post positive scientific method’ encourages a reductionist over-focus on a part, not the whole. Leadership requires the harnessing of collective intelligence – and the speed and complexity with which data can now be manipulated and represented presents unique learning challenges for leaders.

With advances in technology and learning analytics it is within our power now to create an online learning environment which both provides survey tools, and knowledge structuring tools, such as Perimeta and enables them to work together to link an organization’s purpose with its performance (Deakin Crick 2014). This then becomes a leadership decisioning tool available to leaders at all levels throughout the academic year, and particularly at key stages of the development planning cycle.

## Appendix A Data Strands and Sources of Evidence

Data strand	Stakeholder group	Source(s) of evidence	Notes
1	Students in Year 7	Research-validated questionnaire completed at the beginning and end of Year 7; $N = 300$	Focus on the development of students as learners
2	Students in Years 8 and 10	Questionnaire with some research-validated questions plus questions relating to the Charter $N = 600$	Questionnaire available online using the Oasis VLE; this allows for analysis according to gender, age, etc.
3	Students in Year 9	Recorded interview based on a few questions designed to get under the surface of learning $N = 30$	A resource-intensive attempt to explore deeper learning
4	Students in Year 11	Questionnaire designed to provide evidence about the impact of education at an Oasis Academy on the transition to FE, training and/or work and on into later life $N = 300$ (potentially)	Questionnaire available online and with incentives to complete it annually over several years; ambitious attempt to track the progress of students longitudinally and after leaving the Academy

(continued)

**Appendix A.** (continued)

Data strand	Stakeholder group	Source(s) of evidence	Notes
5	Parents/carers	Questionnaire about engagement with and support from the Academy for children's learning N = 100	This strand will include a group of parents/carers whose first language is not English
6	Teachers	Questionnaire and recorded interview about the impact of CPD on classroom practice and experience of the Oasis ethos (rhetoric or reality?) N = 30	An attempt to get under the surface of teacher learning and evaluate the extent of its impact on classroom practice and student outcomes
7	Leaders	Questionnaire for senior leaders initially, followed by discussion at a senior leadership team meeting about their role as 'leaders of learning', their impact on the culture and climate of the Academy and the learning of teachers and students N = 25	Student and teacher questionnaires also ask about the impact of the senior leaders on the culture and climate of the school with regard to learning; intention to extend this process to middle level leaders
8	–	Data from key performance indicators, for example, attendance, exclusions, student progress	Data for input to Perimeta will be relevant to each strand above
9	Students in Year 11	GCSE results N = 300	Data to complement strands 4 and 8 above

Yeager, D., A. Bryk, J. Muhich, H. Hausman, and L. Morales. 2015. 'Practical Measurement ' In. San Francisco: Carnegie Foundation for the Advancement of Teaching.

**Declaration of Conflicting Interests**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding**

The author(s) received no financial support for the research, authorship, and/or publication of this article.

**Endnote**

1. <http://www.learningemergence.com> [www.explorance.com](http://www.explorance.com).

**References**

- Argyris C and Schön D (1996) *Organizational Learning. 2. Theory, Method, and Practice*. New York: Addison-Wesley.
- Assessment Reform Group (1999) *Assessment for Learning: Beyond the Black Box*. Cambridge: University of Cambridge School of Education.
- Bateson G (1972) *Steps to an Ecology of Mind* San Francisco, CA: Chandler.
- Beare H (2001) *Creating the Futures School*. London: Routledge.

- Bereite C and Scardamalia M (1989) Intentional learning as a goal of instruction. In: Resnick L (ed.) *Knowing, Learning and Instruction. Essays in honour of Robert Glaser*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Blockley D (2010) The importance of being process. *Civil Engineering and Environmental systems* 27: 189–199.
- Blockley D and Godfrey P (2000) *Doing it Differently, Systems for Rethinking Construction*. London: Telford.
- Bore A and Wright N (2009) The wicked and complex in education: developing a transdisciplinary perspective for policy formulation, implementation and professional practice. *Journal of Education for Teaching* 35: 241–256.
- Bryk A (2009) Support a science of performance improvement. *Phi Delta Kappan* 90: 597–600.
- Bryk A and Gomez L (2008) Ruminations on reinventing an R&D capacity for educational improvement. *The Supply Side of School Reform and the Future of Educational Entrepreneurship*. Stanford, CA: American Enterprise Institute Conference.
- Bryk A, Lee V and Holland P (1993) *Catholic Schools and the Common Good*. Cambridge, MA: Harvard University Press.
- Bryk AS, Sebring PB, Allensworth E, Luppescu S and Easton JQ (2010) *Organizing Schools for Improvement: Lessons from Chicago*. Chicago, IL: University of Chicago Press.
- Bryk A, Gomez L, Grunrow A and LeMahieu P. (2015) *Learning to Improve: How America's Schools Can Get Better at Getting Better*. Harvard, MA: Harvard Educational Press.
- Buckingham Shum and Deakin Crick R (2012) Learning dispositions and transferable competencies: pedagogy, modelling and learning analytics. Proceedings of the 2nd International Conference on Learning Analytics and Knowledge. Vancouver, BC: ACM Press.
- Checkland P and Scholes J (1999) *Soft Systems Methodology in Action*. Chichester: Wiley.
- Clarke A and Collins S (2007) Complexity science and student teacher supervision. *Teaching and Teacher Education* 23: 160–172.
- Davies B (2011) *Leading the Strategically Focused School: Success and Sustainability*. London: SAGE.
- Davis B and Sumara D (2006) *Complexity and Education Inquiries into Learning, Teaching and Research*. London: Routledge.
- Davies B and West-Burnham J (2003) *The Handbook of Educational Leadership and Management*. London: Pearson.
- Davis J and Fletcher P (2000) *Managing Assets under Uncertainty*. Society of Petroleum Engineers Asia Pacific Conference on Integrated Modelling for Asset Management. Yokohama, Japan.
- Davis J, MacDonald A and White L (2010) Problem-structuring methods and project management: an example of stakeholder involvement using Hierarchical Process Modelling methodology. *Journal of the Operational Research Society* 61: 893–904.
- Day C, Sammons P, Hopkins D, Harris A, Leithwood K, Gu Q, Brown E, Ahtaridou E and Kington A (2010) *The Impact of Successful Leadership on Learning Outcomes: Ten Strong Claims*. Nottingham: National College of School Leadership, DfES.
- De Jaegher H and Di Paolo E (2007) Participatory sense-making: an enactive approach to social cognition. *Phenomenology and the Cognitive Sciences* 6: 485–507.
- Deakin Crick R (2012) Student engagement: identity, learning power and enquiry – a complex systems approach. In: Christenson S, Reschly A and Wylie C (eds) *The Handbook of Research on Student Engagement*. New York: Springer.
- Deakin Crick R (2014) Learning to learn: a complex systems perspective. In: Deakin Crick R, Stringer C and Ren K (eds) *Learning to Learn: Perspectives from Theory and Practice* (pp. 69–90). London: Routledge.
- Deakin Crick R, Jelfs H, Ren K and Symonds J (2010) *Learning Futures*. London: Paul Hamlyn Foundation.
- Deakin Crick R, Jelfs H, Huang S and Wang Q (2011) *Learning Futures Final Report*. London: Paul Hamlyn Foundation.

- Deakin Crick R, Huang S, Goldspink C and Ahmed-Shafi A (2015) Developing resilient agency in learning: the internal structure of learning power. *British Journal of Educational Studies* 63(2): 121–60.
- Dweck CS (2000) *Self-Theories: Their role in Motivation, Personality, and Development*. New York: Psychology Press.
- Goldspink C and Kay R (2007) Systems theory and the problem of structure and agency. *ANZSYS07 Conference*. Auckland, NZ.
- Goldstone R (2006) The complex systems see-change in education. *The Journal of the Learning Sciences* 15: 35–43.
- Hall J, Le Masurier J, Baker-Langman E, Davis J and Taylor C (2004) Decision-support methodology for performance-based asset management. *Journal of Civil Engineering and Environmental Systems* 21: 51–75.
- Harlen W and Deakin Crick R (2003a) *A Systematic Review of the Impact of Summative Assessment and Testing on pupils' Motivation for Learning*. London: Evidence for Policy and Practice Co-ordinating Centre Department for Education and Skills.
- Harlen W and Deakin Crick R (2003b) Testing and motivation for learning. *Assessment in Education* 10: 169–207.
- Jackson D and Temperley J (2007) From professional learning community to networked learning community. In: Stoll L and Seashore Louis K (eds) *Professional Learning Communities: Divergence, Depth and Dilemmas* (pp. 45–62). Maidenhead: Open University Press.
- James M and Gipps C (1998) Broadening the basis of assessment to prevent the narrowing of learning. *The Curriculum Journal* 9: 285–297.
- James M, McCormick R, Black P, Carmichael P, Drummond M, Fox A, Macbeath J, Marshall B, Pedder D, Procter R, Swaffield S, Swann J and William D (2007) *Improving Learning How to Learn: Classrooms, Schools and Networks*. London: Routledge.
- Little A and Veugelers R (2005) *Internationalization of R&D in the UK*. London: Office of Science and Technology.
- Macbeath J and Mcglynn A (2002) *Self-evaluation: What's in it for Schools?* London: RoutledgeFalmer.
- Mason M (ed.) (2008) *Complexity Theory and the Philosophy of Education*. Chichester: Wiley-Blackwell.
- National College for School Leadership (2004) *Learning-centred Leadership*. Nottingham: NCSL.
- Opfer V and Pedder D (2011) The lost promise of teacher professional development in England. *European Journal of Teacher Education* 34: 3–24.
- Pedder D (2006) Organizational conditions that foster successful classroom promotion of learning how to learn. *Research Papers in Education* 21: 171–200.
- Pedder D (2007) Profiling teachers' professional learning practices and values: differences between and within schools. *The Curriculum Journal* 18: 231–252.
- Pedder D, James M and Macbeath J (2005) How teachers value and practise professional learning. *Research Papers in Education* 20: 209–243.
- Pedder D and Opfer V (2013) Professional learning orientations: patterns of dissonance and alignment between teachers' values and practices. *Research Papers in Education* 28(5): 539–570.
- Reay DWD (1999) 'I'll be a nothing': structure, agency and the construction of identity through assessment. *British Educational Research Journal* 25: 343–354.
- Seely Brown J and Thomas D (2009) Learning for a world of constant change. 7th Gillon Colloquium. University of Southern California.
- Siegel D (2012) *The Developing Mind: How Relationships and the Brain Interact to Shape Who We Are*. New York: Guilford Press.
- Silins H and Mulford B (2001) *Reframing Schools: The Case for System, Teacher and Student Learning*. Freemantle: Australian Association for Research in Education.
- Sillitto H (2014) Architecting systems: concepts, principles and practice. College Publications Systems Series.



- Stoll L and Louis K (2007) *Professional Learning Communities: Divergence, Depth and Dilemmas*. New York: McGraw-Hill International.
- Tracy H (2014) Making it a great and effective place to work: co-creating the conditions for happiness and success. PhD thesis, University of Bristol.
- Wheatley M (1999) *Leadership and The New Science: Discovering Order in a Chaotic World*. (2nd ed.). San Francisco CA: Berrett-Koehler Publishers.
- Woolfolk H, Hoy W and Davis H (2009) Teachers' self-efficacy beliefs. In: Wentzel K, Wigfield A and Miele D (eds) *Handbook of Motivation at School*. Routledge, 727–654.
- Yeager D, Bryk A, Muhich J, Hausman H and Morales L (2015) *Practical Measurement*. San Francisco: Carnegie Foundation for the Advancement of Teaching.

### Author biographies

**Ruth Deakin Crick** is Professor of Learning Analytics and Educational Leadership at the University of Technology Sydney and Reader in Education at the University of Bristol. She is one of the originators of a learning analytic designed to measure and strengthen an individual's learning power and provide a framework for self-directed learning and teacher-led pedagogical change. Her research focuses on processes of learning, adaptation and feedback in complex social systems. She is Co-Investigator in the International Centre for Infrastructure Futures that brings together stakeholders involved in renewing the UK's infrastructure. She is a founding member of the Learning Emergence Partnership.

**Steven Barr** is a Chartered Engineer and holds a Masters degree in Operational Research and a PhD from research into dealing with complex problems of organisations. He is a Visiting Fellow of the Systems Centre at the University of Bristol and of Cranfield University. He is Director of ReasonUp, a collaborative consultancy that supports innovation, and was formerly National Director of the UK government's Manufacturing Advisory Service. He is a member of the Learning Emergence Partnership.

**Chris Goldspink** is Chief Scientific Officer of Incept Labs, Sydney which specializes in addressing complex problems in organizations. He holds a PhD from the University of Surrey in the UK and is a Founding member of the Learning Emergence Partnership. He led research into student engagement and quality pedagogy in the Department of Education and Children's Services for South Australia between 2001 and 2012.

**David Pedder** is Professor of Education and Director of Research at the University of Leicester School of Education. His research projects reflect his interest in classroom teaching, learning and innovation, teachers' professional learning, the organisational learning of schools, educational leadership and school improvement, and school-university partnerships. He was Principal Investigator for the Schools and CPD in England State of the Nation Study and is founding member of the Leicester Lesson Study Research Group.