To appear in ZDM – The International Journal on Mathematics Education
Special Issue: Classroom-based interventions in mathematics education, 2013

Chance by design: devising an introductory probability module for implementation at scale in English early-secondary education

Kenneth Ruthven & Riikka Hofmann

University of Cambridge
Faculty of Education
182 Hills Road
Cambridge CB2 8PQ, UK

Abstract

This paper reports the design of an introductory probability module intended for implementation at scale within the English educational system. It forms part of the *epiSTEMe* programme of redesign research aimed at improving the teaching and learning of mathematics and science at early-secondary level. The approach taken by the module is informed by the research literatures on effective teaching (with a particular emphasis on blending teaching components and exploiting dialogic discussion) and probabilistic thinking (with a particular emphasis on triangulating epistemic approaches and deconstructing fallacious reasoning). Recognising that scaleable innovation must take account of the current state and established norms of the educational system, module development was informed by such considerations. Advice and feedback from classroom teachers, as well as observation and recording of their lesson implementations, provided a basis for assessing the viability of proposed features of the module, and the adaptation required of teachers, so that guidance materials and professional development could be framed appropriately.

Keywords

design research, dialogic teaching, early-secondary school, England, improvement at scale, pedagogical design, teaching probability

1 Introduction

The goal of the *epiSTEMe* (Effecting Principled Improvement in STEM Education) project is to develop a research-informed pedagogical intervention in early-secondary physical science and mathematics, through "re-design" research that recognises that design for implementation at scale needs to take account of the existing state of the system: notably the people, structures, resources and practices already in place (Ruthven et al. 2010).

For over a decade, systemic improvement effort in England has promoted a pedagogical approach that combines tightly structured teaching to objectives with target setting for pupils, teachers and schools. The guiding model of 'direct interactive' teaching emphasises the role of the teacher in directing, instructing, demonstrating, explaining and illustrating, questioning pupils and leading discussion, evaluating responses and providing feedback, and summarising (Department for Education and Employment [DfEE] 2001). This improvement programme has not been an unqualified success. Longitudinal analysis of national and international findings on English system performance (Ruthven 2011a) suggests that, in mathematics, pupils' content knowledge and skills have been strengthened but not their broader subject literacy or functional capability; in science, these aspects of achievement have not changed. On attitude, while pupils' valuation of learning both subjects has risen modestly, their liking of mathematics and of science and their enjoyment of learning them has fallen substantially.

In this paper we will first outline the broad pedagogical approach embraced in the *epiSTEMe* project (§2) and the general apparatus developed to support its use (§3), and then illustrate the dialogic component which is its most distinctive pedagogical feature (§4). The specific focus of this paper on the design of a probability module will be anticipated in this section (§4), and then explicitly addressed through an outline of the iterative process of design (§5), which had to take account of established systemic frameworks for teaching probability (§6) as well as complementary epistemic approaches to the topic (§7). In this light, the evolving design rationale for the three-lesson teaching sequence forming the opening cycle of the module will be explained (§8-§10).

2 The *epiSTEMe* pedagogical model

Although important gaps remain in this literature, synthesis of the international research base on effective pedagogy in school mathematics and science (Seidel and Shavelson 2009; Schroeder et al. 2009; Slavin and Lake 2008; Slavin, Lake and Groff 2009) indicates that the following types of teaching component are particularly effective (Ruthven 2011b): *domain-specific enquiry* which poses authentic problems and takes student thinking seriously; *co-operative* or *collaborative group-work*, as long as students are properly prepared and activity well structured; *enhanced context* in which teaching makes strong links to student experiences and interests; and *active teaching* of the type informing the English model of direct interactive teaching.

The *epiSTEMe* project blends these teaching components in a model informed by more specific bodies of pedagogical research. In particular, the *epiSTEMe* model has been strongly influenced by UK research on classroom discourse and interaction (Mercer et al. 2004; Mercer and Sams 2006; Scott et al. 2006) which points to the value of dialogic small-group and whole-class discussion in encouraging pupils to talk in an exploratory way and to consider different points of view. Key aspirations for dialogic talk are that it be *collective* in involving all participants, *reciprocal* through considering different viewpoints, *supportive* of free expression and mutual assistance, *cumulative* in chaining and developing ideas, and *purposeful* towards particular curricular goals (Alexander 2008).

The pedagogical cycle at the core of the *epiSTEMe* model has phases of exploration, codification and consolidation (Ruthven 1989). In the opening exploratory phase, domain-specific enquiry tasks are employed to support informal development of target concepts. Dialogic small-group and whole-class discussion provides opportunity for pupils to express their thinking about a problem situation and to examine different perspectives on it. During such discussion, the teacher's principal role is to support the dialogic quality of contributions by pupils and exchanges between them. The cycle continues to a codification phase in which the teacher's role becomes a more authoritative one of explaining accepted mathematico-scientific approaches to the problem situation through active teaching which takes account of the thinking displayed during the earlier exploration phase. In a final consolidation phase, pupils tackle related problem situations more independently, with the teacher's role becoming one of checking pupil understanding and providing developmental feedback.

The *epiSTEMe* model incorporates further pedagogical principles. In view of the promising research findings for enhanced context, attention is given to conveying a sense of the wider human interest and social relevance of a topic, and to making relevant connections with widely shared pupil experiences and interests. Curricular prescriptions are also filled out to support the building of strong conceptual foundations: in particular, account is taken of what is known about informal knowledge and thinking related to a topic, and means provided of deconstructing common forms of fallacious reasoning.

3 The *epiSTEMe* intervention apparatus

The English school system is a highly regulated one in which departments and teachers are held publicly accountable through national testing and external inspection. This has produced a culture of compliance with official norms and of caution regarding pedagogical innovation (Prestage and Perks 2008). As official reports based on school inspections testify (Office for Standards in Education [Ofsted] 2008), in many schools a narrow focus on accountability requirements results in an objectives-driven pattern of teaching which emphasises immediate markers of superficial progress. Shortages of suitably qualified teachers also produce instability of staffing, particularly at early-secondary level. In many schools, mathematics and science departments operate under considerable pressure, with limited professional cohesion and developmental capacity.

Our goal, then, has been to support teachers and departments in working towards a renewed pedagogical approach without requiring significant reorganisation of work and substantial investment of time. We have devised a professional development and classroom teaching intervention, modest in scope, and packaged as a viable substitute for modules currently widely taught in schools. The focus is on Year 7, the first year of secondary education (during which pupils reach the age of 12): this is the period most distant from the inhibiting backwash of external testing, and the one during which teachers are actively shaping norms of classroom participation.

The *epiSTEMe* apparatus consists of the following components. An Introductory Module is designed to build teacher and pupil understanding of the value of talk and dialogue in supporting subject thinking and learning, and to develop rules and processes that support effective small-group and whole-class discussion. Two Topic Modules in each subject are designed to support and capitalise on such use of talk and dialogue, and to instantiate key pedagogical principles and processes. These modules are mediated by teaching materials designed to be educative in the sense of supporting teacher development as well as classroom activity (Davis and Krajcik 2005), and supported by a sequence of two one-day Professional Development events. The first event focuses on dialogic teaching and on how the Introductory Module supports its development; then, after teachers have undertaken the Introductory Module with one of their classes, the second event debriefs this experience and examines how the Topic Modules in their subject incorporate the pedagogical principles and processes of the *epiSTEMe* model.

The modules have been designed so that a wide range of teachers and departments will find them readily and robustly usable. In particular, the Topic Modules provide a full set of classroom materials which explicitly target the required curricular objectives through a sequence of lessons lasting for the period of time that schools typically give to that topic. They are designed to respect the parameters normally found within the system, but sufficiently flexible to be adaptable to other common variants. For example, the normal length of a timetabled session in English secondary schools is around 60 minutes, of which about 50 minutes is typically available for teaching. Lessons are designed to fit such a session length, but are structured as sequences of shorter activities to facilitate adjustment to another session length or to a differing lesson pace.

Equipment requirements have been limited to items known to be widely available and easily usable. For example, soundings indicated that classrooms typically had either an interactive whiteboard or computer projection, and that teachers were accustomed to using these, often with prepared displays, usually alongside a traditional whiteboard. This led to Projection Slides becoming the primary support for classroom implementation of lesson sequences. Nevertheless, during trialling it became clear that school computer systems could not be relied on: here the Study Booklet given to each pupil provides an important fall-back. Teaching Notes for each module support lesson planning, highlight key aspects of each activity and explain its rationale, and advise on handling pupil responses.

4 An illustrative dialogic activity

A central aim of the design process was to devise resources to support effective use of dialogic approaches to classroom teaching and learning. A ‘dialogic’ approach is one that takes different points of view seriously (Scott et al. 2006), encouraging students to talk in an exploratory way that supports development of understanding (Mercer and Sams 2006). Orchestrating such discussion to advance the learning of the whole class is acknowledged to be a significant challenge: pedagogical goals include facilitating public expression and respectful examination of students’ thinking; focusing – but not funnelling – discussion to prevent it becoming overly fragmented and incoherent; and guiding students towards accepted disciplinary norms of reasoning and communication (Franke, Kazemi and Battey 2007; Stein, Engle, Smith and Hughes 2008).

Amongst the techniques that we employed with a view to supporting development of dialogic classroom talk were: (1) often organising discussions around “concept cartoons” in which canonical examples are given of the differing types of viewpoint likely to be present amongst pupils; (2) preceding whole-class discussion by more informal small-group discussion between teams of pupils to allow time for preliminary exploration and verbalisation of ideas; and (3) establishing explicit ground rules aimed at supporting productive discussion.

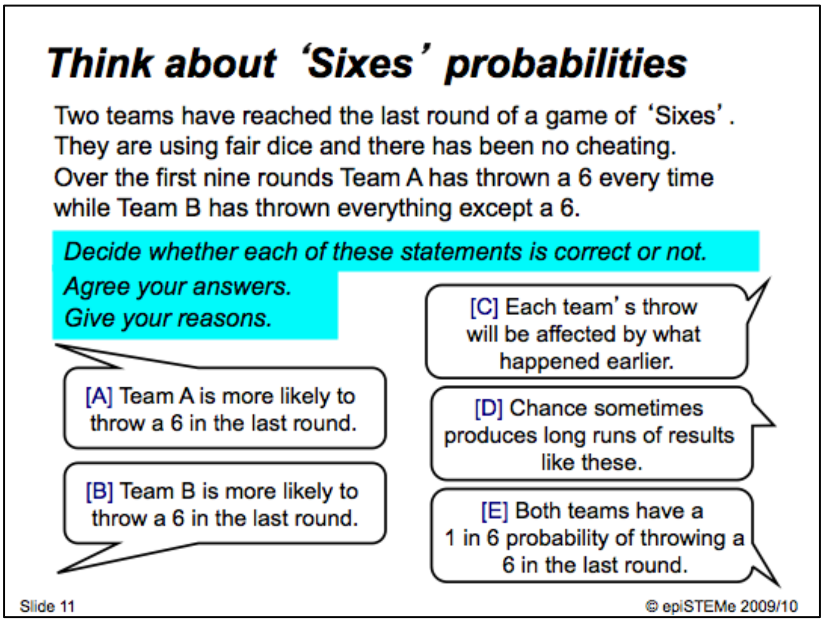

Fig. 1: Concept cartoon task under discussion in Box 1 dialogue

01	T:	Right. Chad?
02	P1:	We thought, well E, because there's no way that, if the dice isn't rigged, there's no way that it has; the dice can't choose which way it goes, can it?
03	T:	No.
04	P1:	So it's just chance really, it could go any way. So we thought E.
05	T:	I'm going to ask the girls. Girls, did you all agree with that?
06	Ps:	Yes.
07	T:	Okay. At this table, did you manage to come up with
08	P2:	Yes.
09	T:	And what did you agree?
10	P2:	That Team B is most likely.
11	T:	Okay, why did you think Team B was most likely to throw a six in the last round?
12	P2:	Because Team A would be tired of throwing sixes.
13	P3:	Well maybe not if it was
14	T:	Please don't jump in.
15	P4 :	How can you be tired of
16	T:	Don't jump in.
17	P2:	They get like unused to it, like they forget how to do it. And then they tell B, and then B know how to do it.
18	T:	So Riba's team thinks that Team B have thrown, you think that they're tired of not throwing a six, so they'll throw a six.
19	Ps:	Yeah.
20	T:	I need a group that disagrees with that to be able to come back and; yes Chad
21	P1:	I think that is totally out of order, because it's not their choice really whether they get a six or not. They're just throwing it and seeing where the dice lands.
22	T:	Okay. Riba, what do you think about what he's said?
23	P2:	I didn't hear it.
24	T:	Right, would you like to say it again, Chad, please?
25	P1:	I think what they're saying is totally out of order because you don't choose which way it goes, you can't all of a sudden get tired of not throwing a six and throw a six, it's up to the dice which way it goes.

Box 1: Excerpt from classroom discussion of concept cartoon task

For example, an important ground-rule is that pupils should try to come to an agreed position; even if they are unable to achieve this goal, honouring it calls for them to engage with points of view other than their own, and to develop an argument in support of their position. The classroom excerpt in Box 1 comes from a whole-class plenary discussion (of the task shown in Fig. 1) in which contrasting responses and arguments are being elicited from two groups, establishing a clear need for further whole-class discussion.

One feature of the exchange is the way in which the teacher contribution is primarily one of organising the expression and exchange of pupil ideas: by nominating teams and speakers (turns 01, 05, 07, 20, 22, 24) and maintaining an orderly and coherent flow of discussion between speakers (14, 16, 20, 22, 24); by seeking to probe and clarify pupil contributions (11, 18); and by accepting such contributions neutrally, rather than evaluating them (3, 5, 11, 18, 22). Supporting dialogic exchange is the aspect of the *epiSTEMe* pedagogical model that teachers have found particularly challenging. Because this approach emphasises developing mathematico-scientific thinking as its goal, not simply securing task performance, it requires many teachers to make significant shifts beyond the received ideas and habitual reflexes of established practice. For example, a dialogic approach calls for the teacher to be prepared to give time to multiple pupil contributions which can be persuasively fallacious or poorly formulated. Moreover, the teacher needs a nuanced understanding of the topic to support pupil conceptualisation and reasoning. To sustain productive discussion, the teacher must be able to identify and inter-animate the thinking behind different pupil responses, and steer progression in reasoning without closing down discussion.

The first *epiSTEMe* professional development event employs the video-recorded example of a plenary review to examine how teachers can support quality of classroom discussion. Research analysis (Ruthven, Hofmann and Mercer 2011) informed our choice of a sequence of short episodes to stimulate discussion with and between teachers, with the classroom dialogue transcribed to encourage attention to the fine grain of pupils' mathematical thinking and of the teacher's participation in exchanges. The emphasis is on encouraging teachers to "read" what is taking place as each episode unfolds so as to understand how pupils are thinking and responding, to analyse the quality of dialogic interaction, and to anticipate accordingly how the teacher might productively shape events and ideas. While the research analysis informs our contributions to the discussion with teachers during professional development, we do not explicitly present it.

During the design process, judgements about the effectiveness of lesson design had to be made on the basis of a small number of trials. Here, for example, how had the teacher's subtle recasting of the task as a matter of finding a single correct response affected matters? Nevertheless, this part of the lesson, and the task design in particular, can be viewed as successful inasmuch as they do generate different viewpoints which the teacher is able to identify and organise to develop exchange between them. At the same time, Riba's response (12, 17) to the teacher's probe does not articulate the type of common-sense (negative recency) reasoning that often underpins a type-B response. Did Riba just report what her team had agreed was "the right answer" (10)? Did the teacher's probe (11) push her to improvise some reason rather than report the perhaps forgotten discussion behind that answer (12)? What if the teacher had probed another member of the team (P3, for example, who interjected in turn 13)? On the other hand, the animated pupil reactions (13, 15, 21) to Riba's response do suggest that her contribution was starting to generate real mathematical debate at a level which may have been better adjusted to pupils in this set of below-average attainment.

In designing a topic module on probability, then, one of our central concerns was to craft lesson sequences, task materials and teaching guidance that would produce, in a dependable way, dialogic exchanges capable of contributing to intended learning. In the light of earlier research highlighting how delayed incubation effects underpin individual learning, a particular focus of our analysis of lessons was on the quality of these public exchanges, particularly the resources they afford to prime future problem solving, notably through representing alternative perspectives and approaches (Howe et al. 2005).

5 Iterative development of the module

Our aim was to take account of multiple considerations of national curriculum, class organisation, classroom equipment and teacher expectation in devising a module, following the *epiSTEMe* pedagogical model, that teachers would find a viable (indeed, preferable) substitute to their existing unit on the topic.

We worked with a group of mathematics teachers from five schools who critiqued and trialled successive versions of the module. This was not straightforward as trialling had to respect the strict annual timetable for teaching the topic in each school, and there was considerable attrition of participating teachers. By the start of the second year, 4 of the 8 mathematics teachers with whom we had started working had left their schools (2 to positions in organisations concerned with professional development, and 2 to schools outside the region) and a further 2 were no longer teaching Year 7 classes (1 taking maternity leave, and 1 redeployed to teach only older classes because of staffing shortages). To complete trialling, we recruited 3 further teachers to join the 2 remaining from the initial group.

A first version of the module was trialled by 2 teachers during summer term 2009, a second version by 3 teachers (one with 2 classes) during spring term 2010, and a third version by 2 teachers (one with 2 classes) during summer term 2010. Each wave of trialling led to revision of the module and it is on this experience that this paper draws. One consequence of teacher turnover was that only one teacher trialled the module in more than one version; this limited the scope for teachers to make comparative assessments of aspects of the module that had changed, and to become more experienced in handling those aspects that remained the same. This increased uncertainty about whether emergent difficulties should be attributed to module design alone or might also reflect teacher inexperience, making judgements about revision more complex.

Feedback from teachers was typically given orally, either on an individual basis following our observing them teaching a lesson, or in collective discussion during workshops convened to review the module after a wave of trialling. We were also able to observe or videotape some sequences of trial lessons. This more detailed information proved particularly useful in identifying specific changes to lesson designs that might be beneficial, and issues that deserved attention in teaching notes or professional development. In particular, such evidence ensured that the detail of design better acknowledged the ways of thinking and forms of language prevalent amongst pupils and teachers.

On the central issue of dialogic discussion, we used such records to carry out fine-grained analysis of how teaching strategies and tactics mediate exchanges during classroom discussion, and of the concrete form that these take in particular lessons (c.f. Ruthven et al. 2011). However, given the complexity and discontinuity that characterise ordinary school classrooms, design work also calls for considerable exercise of judgement in interpreting varied and sometimes inconsistent feedback, and in reconciling multiple and sometimes conflicting considerations. To address this multidimensionality and ambiguity within limited time and resources, we often had to rely on more holistic and informal assessment of the functioning of a lesson, empirically grounded in evidence from classroom implementations. In particular, we had to try to distinguish between situations in which disappointing classroom discussion arose from teachers not capitalising on promising pupil contributions (suggesting a need to sensitise teachers, typically through making the potential of such contributions more explicit in the teaching notes) as against situations in which the lesson task (as mediated by the associated classroom materials) did not adequately elicit such contributions from pupils (suggesting the need for change to the task or materials).

In the second and third waves we also conducted systematic evaluation of the learning gains made by pupils. A suite of tests was designed to provide balanced coverage of the objectives specified by the curriculum through items that did not replicate any task included in the module[1]. The 15-minute tests consisted of sets of items which were either identical or parallel in pre and post versions; two sets (worth 8 marks) required a structured response (responding ✓ or ✗ according to whether statements were right or wrong; matching event statements to positions on the probability scale), the remaining five sets (12 marks) required open responses in the form of short numerical answers. Because they too were undergoing calibration and refinement, the tests administered to each class were not identical, but were sufficiently similar for the results to be considered comparable. We focused particularly on learning gains between pre-test and deferred post-test, the latter taken unannounced a month after module completion[2]. The learning-gain effect sizes[3] for classes were consistently positive – 0.56, 0.62, 0.66, 0.76, 0.83, 0.98 – averaging 0.74. In the first two cases, the same test sequence was also taken by parallel classes following the school's existing probability module, with the results consistently favouring the intervention group – for an average learning-gain effect size of 0.59 against 0.22. In the future we expect to be able to report a wider range of findings from a randomised field trial involving 24 Year 7 classes and their teachers.

6 Established systemic frameworks

Subject teaching in English (state) schools is regulated by a national framework of curriculum, assessment and inspection. This regulation is particularly strong in core subjects such as Mathematics, taken by all pupils throughout the compulsory years of schooling. At the heart of this national framework is a model of level by level progression in each subject, with each level defined in terms of specified assessment objectives. The presumption is that pupils progress at the rate of about one Level every two years. Around 80 per cent of pupils enter secondary school having achieved at least Level 4 in Mathematics in the national tests at the end of primary schooling (with Level 5 the highest available at this stage during the period over which this work was conducted).

Each secondary-school department is expected to devise its own scheme of work which translates the official programme of study in the subject into more concrete plans for teaching units directed towards national assessment objectives. Generally, these schemes indicate the sequence in which topics will be taught and the amount of time to be devoted to each, organised into a timetable covering the whole school year which all teachers in the department

[1] The teacher of one class did not administer tests.
[2] In cases where the deferred post-test was not administered, we substituted results from the immediate post-test.
[3] Cohen's d, employing a pooled estimate of standard deviation across pupils in all classes.

are expected to follow closely. Most schemes of work recommend resources available for teaching each topic. Some provide a detailed schedule of lessons but the tendency is for teachers to be given considerable discretion. Most secondary schools form their mathematics classes by dividing pupils into "sets" according to assessments of their ability in the subject. Typically, then, a scheme of work will be differentiated into several tiers according to the Level which pupils in a class are "working towards".

Within the national curriculum for primary education, the topic of probability receives minimal attention. This reflects not just a prioritisation of the curriculum strands of number, space and measure, but a retreat from ambitious aspirations for the teaching of probability throughout primary education. Experience over the 1990s highlighted the difficulty of creating an effective probability curriculum capable of implementation at scale by largely non-specialist teachers. Not only was it found challenging to devise age-appropriate teaching sequences that were effective in socialising children's intuitions about chance, but tasks intended to foster and assess probabilistic thinking could too easily be stripped of any genuine stochastic dimension and reduced in practice to more superficial exercises (Threlfall 2004).

At early-secondary level, probability forms one – relatively discrete – component of a strand entitled "Statistics" in the official study programme and "Handling data" in the official attainment targets. The objectives for probability have remained stable for some time, as comparison of successive documents (Qualifications and Curriculum Authority [QCA] 1999; 2007) indicates (Table 1: *italicised phrase* appeared in 1999 only). Perhaps the most important point to note about learning progression is that the Level 5 objectives concern the use of basic ideas, terms and representations, and explicitly place in parallel the two dominant and contrasting epistemic approaches to conceiving and quantifying probabilities (as will be discussed more fully in §7). However, it is not until Level 7 (after, by implication, several years) that the sole objective calls for a co-ordinated and reflexive use of these approaches. The intermediate Level 6 objectives move beyond the direct formulation of simple probabilistic situations at Level 5 to the indirect derivation of sample spaces and probability values by using structural relations.

Table 1: Probability assessment objectives in the English national curriculum for Mathematics

Level	**Objective**
4	[NO ENTRY]
5	[Pupils] understand and use the probability scale from 0 to 1. Pupils find and justify probabilities and approximations to these by selecting and using methods based on equally likely outcomes and experimental evidence, as appropriate. [Pupils] understand that different outcomes may result from repeating an experiment.
6	When dealing with a combination of two experiments, pupils identify all the outcomes*, using diagrammatic, tabular or other forms of communication*. When solving problems, [pupils] use their knowledge that the total probability of all the mutually exclusive outcomes of an experiment is 1.
7	Pupils understand relative frequency as an estimate of probability and use this to compare outcomes of experiments.

To better understand patterns prevalent across the system in teaching probability, we examined school schemes of work for Year 7, and textbooks aimed at that year, as well as holding discussions with the mathematics teachers participating in the project. It became clear that probability was taught either through a single discrete block of lessons on the topic or through shorter – but still discrete – clusters of lessons within two or three larger blocks on "handling data". Over the whole school year, the time devoted to teaching the topic could be as low as 3 hours and as high as 9, but was typically between 4 and 6 hours. Because of the absence of probability objectives at Level 4 in the national framework, the topic sometimes received less attention in the lowest Year 7 sets. Broadly, however, schemes of work for middle and higher tier sets consistently covered the Level 5 objectives, and often at least broached the Level 6 objectives. We found no schemes of work or textbooks addressing the Level 7 objective in Year 7. These characteristics are, of course, consistent with the official expectations described above about the Levels towards which the great majority of pupils in Year 7 will be working. Finally, discussions with participating teachers also indicated that they considered it very important to satisfy pupils' strong expectation that work on probability would involve some playing of games.

Despite the widespread differentiation of school schemes of work into tiers, we chose to develop a common module, We judged that it was possible to devise a scheme that could be made accessible to all Year 7 pupils, helping to set suitably high expectations. Nevertheless, the module does provide scope for variation in depth and pace according to the response of pupils, and for extension either by exploring ideas generated by pupils or through undertaking supplementary tasks. The final sequence consists of 6 core lessons and one optional: the first three and a half lessons cover the Level 5 objectives, the remainder, the Level 6 objectives. Henceforth, we will focus on the first half of the module to explain and illustrate some fundamental aspects of the design and its iterative refinement.

7 Complementary epistemic approaches

The official Level 5 curricular objectives covered in the first half of the module convey a clear expectation that pupils will be exposed both to a classical (sometimes termed theoretical) approach to probability based on "equally likely outcomes" and to a frequential (sometimes termed empirical) approach based on "experimental evidence". In established practice, too, the injunction to "understand and use the probability scale" is often associated with a subjective approach in which probabilities are gauged on the basis of personal judgement and intuition.

One functional reason put forward for acknowledging the subjective approach is that it is applicable to a much wider range of situations than either of the others; the classical approach depends on the possibility of reasoning about a situation in terms of symmetry and necessity; the frequential approach requires the possibility of unrestrictedly replicating a trial. However, in the subjective approach, the process through which probability values are produced is less explicit and accountable than in either the classical or frequential approaches. Indeed, the subjective approach raises – or too often, in educational practice, ignores – tricky issues of the interpretation, calibration and justification of probability values generated in such a way.

The predominant pedagogical argument for acknowledging the subjective approach is as a way of building on pupils' intuitions. This depends, however, on finding effective ways to discipline these intuitions to create a more explicit and coherent system of probabilistic judgement; and such development must confront the counter-intuitive nature of accepted mathematical reasoning even about some simple probabilistic situations (Hawkins and Kapadia 1984). To discipline intuitions calls for systematic consideration of probability spaces, rather than of isolated probability values; and for triangulation of the subjective approach against one or both of the objective approaches within the narrower range of situations to which these latter approaches are applicable. In particular, taking Pratt's (2005) characterisation of intuitive bases for probabilistic thinking, the classical approach provides a framework particularly suited to disciplining intuitions of “fairness” and “unsteerability” (lack of control over outcome); just as the frequential approach does for intuitions of “unpredictability” and “irregularity”.

Thus, in the research literature on the development of probabilistic thinking, it is now widely accepted that teaching which focuses on the “reciprocal dynamic of theoretically computed probabilities and observed relative frequencies… may best contribute to the development of efficient probabilistic intuitions” (Fischbein and Gazit 1984, p. 3). However, one challenge of taking such an approach is that empirical feedback on theoretical predictions is not so readily interpretable in stochastic situations (particularly those traditionally realisable in the classroom) as in deterministic ones (Borovcnik and Peard 1996). Recently, exponents of the use of computer-based experimentation to underpin formation of probability concepts have sought to exploit the capacity of digital technologies to generate large samples in order to capitalise on the law of large numbers to improve the signal to noise ratio in quasi-empirical feedback from quasi-random simulations. Even so – and particularly where uniform distributions are concerned – reading the signal in a single large sample remains problematic (Konold and Kazak 2008). This means that, when teaching probability, discursive interaction between pupils and teacher plays a particularly important part in establishing inter-subjective agreement about relations between theoretical models and empirical patterns. More generally, such interaction is crucial for the appropriation of probability concepts through developing coordinated understanding of their associated semiotic tools and reference situations (Steinbring 2006). This makes probability a topic where a dialogic approach is particularly appropriate.

The literature suggests, then, that the formation of a versatile and systematic conception of probability depends on the appropriation and coordination of all three of the classical, frequential and subjective approaches. Against this background, we judged it important that early lessons in the module should examine types of stochastic situation that would allow relatively informal analysis in terms of all three of the approaches. Such situations, then, needed to be capable both of practical replication (for frequential analysis) and amenable to equiprobabilistic idealisation (for classical analysis). The principle of building on shared prior experience and ideas that pupils bring to lessons suggested analysis of familiar objects such as coins and dice: physical devices that materialise an idealised process of chance-making. Because dice offer more versatility on account of the greater degree of variation they make available – in the number of faces (according to shape) as well as their numbering; and in the position or combination of faces taken as constituting the target outcome (side faces as well as top) – they became the exemplary chance-making devices for the first cycle of lessons in the module. This, of course, is hardly innovative; but it reflects the compelling convergence of arguments towards such a didactical strategy.

8 Lesson 1: *What do you know about probability?*

This first lesson introduces the topic of probability in a way that seeks to draw on pupils' existing knowledge and wider experience. In the opening whole-class activity, *Spot the probability words*, a montage of illustrated news headlines (plus the first few lines of each news report) serves as a prompt for pupils to identify terms that they associate with probability and explain why:

Two more probable swine flu cases
Flu risk for indigenous peoples
Recession in Asia is 'unlikely'
Shares slide on uncertainty
'Tiny chance' of planet collision
Twins more likely for older mums

This also provides scope for discussion of the meaning of the headlines, relating probabilistic thinking to a range of human concerns. Trialling indicated that this opener engaged pupils and that many were willing to volunteer contributions, as illustrated by observation notes from a lower-attaining set where the teacher often found it hard to engage pupils mathematically (Box 2).

> One boy who has seemingly severe problems with accepting to even try to do the pre-test which he simply kept ignoring while it lay on his desk got very engaged once the discussion (after the test) started, made comments and also called out to comment on/correct other students' views (in a friendly and constructive way, with reasons) - and was correct.
>
> A girl (who [the teacher] later commented along the lines of 'she can't do the maths for her life') who before the "spot the words" called out she doesn't know what 'probability' is, and after [the teacher] had some students explain, said she still didn't understand, made several eager contributions once they started listing the words (and once [the teacher] wanted to move on, called out "There are still more probability words there Miss"). Another girl who after the beginning of the task after the pre-test called out "When do we get to the fun bit" [then] appeared quite knowledgeable about the topic and made well-reasoned contributions once the discussion got going.
>
> And really a lot of students contributed, although in a rather chaotic manner typical to this class… In my view the materials got the class really pretty engaged and wanting to make contributions and interested in the topic.

Box 2: Excerpt from field notes on lesson featuring introductory task

The lesson proceeds, for reasons already discussed, to explore pupils' knowledge of dice. The initial design involved discussion of the idea that it is "harder" to throw a 6 than other scores with (conventional cubical) dice (Watson 2005). This is usually attributed to experience of dice games in which players must throw a 6 to achieve some state; the idea of "waiting to get a 6" then disrupts symmetry of outcomes in the mental image of a die. In the original design, then, pupils working in small groups were tasked with deciding which of the following statements apply to a normal fair die:

It has a regular shape.
It is evenly balanced.
If 6 and 1 swapped faces their chances wouldn't change.
6 is ***less*** *likely to come up than 5.*
6 is ***as*** *likely to come up as 4.*
6 is ***more*** *likely to come up than 3.*

In practice, it transpired that the activity generated little good discussion, with pupils coming quickly to agreement in line with accepted mathematical thinking and reporting that they "knew all that already". Perhaps the framing of the activity prompted such a result, and it could be redesigned to generate more productive discussion. At the same time, teachers' reports on discussions, and our observations of them, suggested that the idea of 6 being less likely was neither commonly expressed nor strongly held. This may be because today's pupils have grown up playing computer games rather than dice games; or because the fallacious extrapolation from the experience of "waiting to get a 6" to the idea of 6 being less likely is not anchored in any deeper explanation of how such bias could arise. Indeed, it may be misleading to characterise this as a mis-*conception*.

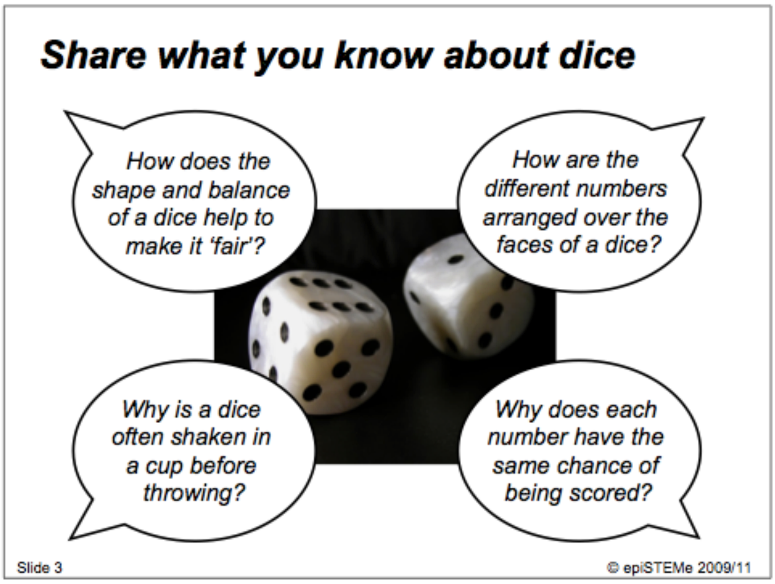

Fig. 2: *Share what you know about dice*

Perhaps, too, it was unwise on our part to place what was an unfamiliar type of discussion activity so early in the module. Equally, the feedback we received from teachers and pupils on this version of the module suggested that it contained too many discussion tasks of this type. Given that this activity took an appreciable period of the lesson, we decided that its learning benefit did not justify its time cost. Instead, we proceeded directly to a more structured whole-class activity, *Share what you know about dice* (Fig. 2) based on focused questions aimed at highlighting special properties of dice which make it possible to reason about their probabilities. The guidance in the Teaching Notes emphasises that rather than directly evaluating pupils' responses as right or wrong, the teacher should encourage explanation of them and exchange of views. It also suggests that bursts of pair exchange may be useful to generate ideas to feed into the whole-class discussion. Finally, the guidance recommends leaving the question on 'the same chance' until last, and prompting pupils to explain how their answers to this last question relate to earlier ones. Feedback from teachers and observation by researchers indicated that this activity proved more productive in bringing pupils to volunteer ideas and information, particularly about "crooked" or "loaded" dice, and about different types and shapes of dice that they were familiar with.

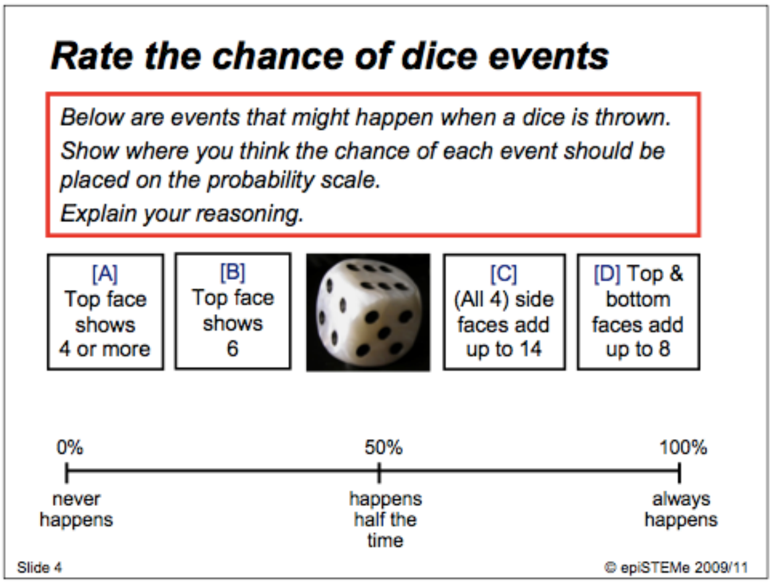

Fig. 3: *Rate the chance of dice events*

The lesson continues with pair, then class, discussion of a fairly conventional set of probability tasks, *Rate the chance of dice events* (Fig. 3), focusing on the idea of a probability scale. The first pair of items is designed to stimulate use of the classical notion of equiprobability, already introduced informally by the previous activity. Observation of trialling confirmed that, as expected, Event A typically elicits variant forms of argument which can then be compared in class discussion. A direct approach, often verbalised in recorded lessons (see Box 3) in terms of "as likely to happen as not" (turn 02), transpires here after teacher probing (03) to appeal to the symmetry of the situation, focusing on the number of outcomes but not explicitly referring to their equiprobability (04). The teacher elicits a further response, "three in six" (06), and then probes it in a way that anticipates the ensuing Event B (included to ensure that pupils tackle a problem which calls for some more generalisable form of unitary strategy) in the hope of eliciting such reasoning (07, 09). Though this is not forthcoming, the pupil response (08, 10) does at least provide an indication of where this task is positioned in their learning zone. On the basis, then, both of prior research and trial observation, the Teaching Notes accompanying tasks of this type guide teachers on clarifying with pupils the ideas surrounding such forms of thinking, and using them as a springboard to develop more explicitly probabilistic reasoning about the situation.

01	T:	Right, four or more on the top face. Would any group like to give me their answer? Yes [in response to indication from a pupil].
02	P1:	It's as likely to happen as not.
03	T:	As likely to happen as not. Why?
04	P1:	Because there are six numbers on a dice and if three of them are four, five and six which is more than four, then the other three must be one, two and three which is half.
05	T:	Okay, so you've got half and half. Yes [in response to indication from another pupil].
06	P2:	Three in six.
07	T:	Three in six. That's an interesting answer. How did you get three out of six?
08	P2:	Because there's three numbers that are above the four which is four, five, six, and there's three that is under.
09	T:	What's the chance of any one number coming up.
10	P2:	I don't know.

Box 3: Excerpt from classroom dialogue about dice probabilities task

The further pair of items was designed to break the conventional mental set surrounding dice probabilities, and to create interplay between classical, frequential and subjective approaches. Trialling confirmed that pupils' initial response to Event C is generally to make a (subjective) judgement by analogy with experience of apparently similar problems (such as the first pair of items); typically, pupils estimate the probability as lying towards the bottom end of the scale. Asked to test their prediction (empirically) by throwing the die issued earlier to each pair, pupils are usually very surprised to find (frequentially) that the numbers on the four side faces appear to always add up to 14, placing the probability at the top position on the scale. Virtually all pairs of pupils are able to reach this point. This finding provokes the question of why this is happening. Answering this requires a shift to some (classical) form of systematic analysis of possibilities, based on exhaustive consideration of cases prompted by inspection of the physical die and/or on noting that each pair of opposite faces bears numbers that sum to 7. While not all pairs of pupils resolve the problem for themselves, subsequent class discussion gives them access to the key ideas. Event D then provides an opportunity to reinforce this line of thinking, and exemplifies a probability occupying the bottom position on the scale.

In this activity, patterns of response often differ, particularly according to level of set. The Teaching Notes suggest gauging how pupils are coping with the first two items and bringing the class together to exchange ideas if many pairs are having difficulty. The contrast between Event A and Event B is also designed to provide a suitable opportunity to discuss and debug the common misuse of "fifty fifty" to refer to any uncertain event, or to any outcome for which there are other equally-likely ones (leading pupils to use such a phrase to describe Event B as well as Event A). However, if most pairs of pupils are producing confident and correct responses to these items, class discussion can be deferred and more emphasis placed on the later pair. Finally, the Teaching Notes provide a more open-ended extension task: trying to change the position of numbers on the die so that event C 'never happens' or 'happens half the time'. In this way, the activity is designed to offer progressively increasing levels of challenge to pupils, according to their response.

The underlying narrative of this lesson, then, is an informal development of the classical approach to analysing stochastic situations in terms of equiprobability and necessity. While the task-sets do provide points of entry for subjective and frequential approaches, these serve primarily to stimulate further attention to the classical approach. This can be seen, in particular, in relation to the handling of Event C, where the striking clash between subjective judgement and frequential evidence points to necessity rather than chance, motivating a shift back towards a classical approach. The second lesson, then, is designed around a complementary narrative in which frequential and subjective perspectives occupy the central place, but with openings created to triangulate these against a classical perspective.

9 Lesson 2: *The game of 'Fives'*

This second Lesson, *The game of 'Fives',* centres on playing and analysing a specially devised game (Fig. 4). Originally this was *The game of 'Sixes',* but, to avoid difficulties of communication that emerged (due to a proliferation of "six(es)" and "sixth(s)", often hard to distinguish in talk), and to bypass any idea of 6 being exceptional, the target value within the game was changed. The inclusion of games in the module was, of course, one of the key design criteria suggested by teachers in the light of pupil expectations that probability modules should be "fun". Trialling of the game proved that it certainly met this expectation. Indeed, pupils often got so excited about it that teachers could find them hard to manage. To address this we developed an even tighter structure for playing the game as a class, particularly to prevent cheating.

Learn to play 'Fives'

'Fives' is a dice game played by teams.

A round involves each team throwing a dice.

A team scores for every 5 that it throws.

The highest score over 10 rounds wins.

Throw	6	1	1	5	5	2	6	4	5	6

This team had to wait until the fourth round to score.

By the end of the game it had scored a total of 3.

Slide 6 © epiSTEMe 2009/11

Review 'Fives' results (spreadsheet)

TEAM	ROUND									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Team A	2	4	2	5	5	5	5	4	6	4
Team B	6	5	6	3	1	3	4	1	3	3
Team C	2	5	6	5	4	6	6	2	6	1
Team D	3	5	2	2	2	5	2	2	5	2
Team E	6	5	6	2	1	3	3	2	3	1
Team F	2	4	1	6	4	5	2	1	3	3

Slide 10 © epiSTEMe 2009/11

Review 'Fives' results

Look at the class results from your game of 'Fives'.
How many 5s were thrown in each round of the game?
Did all the possible numbers of 5s in a round occur?
What were typical numbers of 5s for a round?
What was the average number of 5s for a round?
Are the results surprising? Why (not)?

Suppose your class was going to play lots and lots more games of 'Fives'.
What do you predict would be the average number of 5s in a round?
What do you predict would be the least common number of 5s in a round?
Explain the reasoning behind your answers.

Slide 9 © epiSTEMe 2009/11

Fig. 4: *The game of 'Fives'.*

The game is designed to support empirical examination of chance patterns. Reflection on the game is scheduled at two points. After the first round has been played, stimulus questions focus on the "steerability" of the game and on its "fairness". In trialling, such discussions developed along the anticipated lines, serving to establish a shared appreciation of the chance character of the game and its even-handedness. Once the game has been completed, the second reflection centres on examining the results (presented as in the grid shown in Fig. 4) with the intent of identifying some kind of order behind their "unpredictability" and "irregularity". Here it took several iterations to establish an effective set of stimulus questions. The original questions examined both "horizontal" and "vertical" patterns: horizontally, the different number of throws that teams took to throw their first 5, as well as the differing numbers of 5s thrown; and, vertically, the varying number of 5s in a round. However, this set too broad and complex an agenda to generate productive discussion within a reasonable timespan; our response was to revise the questions to focus attention more narrowly and tightly.

The refined set of stimulus questions focuses on the round-by-round pattern of results, prioritised because it provides the most accessible representation of the key frequential concept of the "rate" at which 5s occur (Fig. 4). The first question serves to establish the 10 values constituting the empirical distribution, and the second to highlight the range of values observed. The third and fourth questions introduce complementary perspectives on the centre of this observed distribution. In the third question, "typical" was chosen as a more elastic term, affording discussion, but likely to elicit identification of particular values, or of a range of values, occurring more frequently than others. In the fourth question, "average" was chosen as a more precise term, likely to elicit estimation or computation of the mean of the 10 values.

Nevertheless, there is also a classical backdrop to this empirical examination. The second question explicitly seeks to triangulate the range of observed values against a theoretically derived range of feasible values. Implicitly too, the stipulation that the game should be played by exactly six teams introduces a triangulation of observed values against the equiprobable expectation of one 5 per round, one 5 per six throws, a simple natural frequency. In effect, this theoretical expectation provides a benchmark around which individual observed values can be seen to vary, and with which the "typical" and "average" values observed can be compared. Such triangulation is also important for a practical reason. While relatively unlikely, it is possible that results over only 10 rounds will diverge markedly from theoretical expectation. If so, it is rather important that the matter is not just left there. Equally, class discussion may give rise to questions or speculations about how results from further games might play out. To pursue such avenues, a spreadsheet-based simulation is provided, allowing further results to be generated in an efficient manner. For similar reasons, the final pair of questions in this set seeks to encourage triangulation of the observed distribution against theoretical expectations or subjective intuitions. A further set of questions then shifts attention to projections beyond the observed distribution. The Teaching Notes recommend offering a summarisation that focuses on the expectation that, over the long term, one sixth of the throws will produce a 5, but in an unpredictable pattern; so that in only some rounds will exactly one team throw a 5, whereas in others more than one team will do so, or no team at all.

The second activity in this lesson, *Think about 'Fives' probabilities*, is intended to stimulate examination of key misconceptions. Originally, the focus was on 'negative recency' (Fischbein and Schnarch 1997): the idea that the absence of an anticipated outcome over a run of trials renders that outcome more likely to occur in the near future. Such a notion can originate in the fallacious generalisation to the short term of the idea that relative frequency should be in line with theoretical expectations or extended observations. However, during trialling, pupils also displayed two further tenacious forms of thinking that represented fundamental obstacles. One was the elastic and ambiguous use of “fifty fifty” to characterise the chance, variously, of one amongst any number of equally likely outcomes, or of one of two complementary outcomes regardless of their relative likelihood, or, indeed, of any unpredictable event (Watson 2005). This latter usage could shade into the second form of antagonistic thinking, to the effect that “anything could happen” or “you can’t know what will happen”, which holds to the idea of making a determinate prediction rather than assessing the likelihood of possibilities (Konold 1989).

The form of the stimulus also evolved in response to trialling (c.f. the extract in Box 1). At an early stage, the situation under consideration was changed to resemble – a somewhat extreme and subsequently modified – one posited by a pupil which had provoked particularly rich discussion. At a later stage, the number of responses was reduced, “correct” ones removed, “incorrect” ones made more plausible, and a blank bubble added for “an improved statement”, because trialling had shown that too many pupils were using social cues rather than reasoning processes to align themselves with a “correct” response. The contents of the concept-cartoon speech-bubbles were also expanded to encourage pupils to give reasons in support of their conclusions, something which trialling indicated as needing particular emphasis (Fig. 5).

Think about ‘Fives’ probabilities

A game of ‘Fives’ has reached the start of the last round.
The teams are using fair dice and there has been no cheating.

TEAM	ROUND									
	1	2	3	4	5	6	7	8	9	10
Team A	5	4	6	5	2	3	5	5	5	
Team B	4	6	2	2	1	4	3	6	1	

Over the first nine rounds:

Team A has thrown a 5 five times and every other number except a 1.

Team B has thrown every number except a 5.

Slide 11 © epiSTEMe 2009/11

Think about ‘Fives’ probabilities

TEAM	ROUND									
	1	2	3	4	5	6	7	8	9	10
Team A	5	4	6	5	2	3	5	5	5	
Team B	4	6	2	2	1	4	3	6	1	

Agree whether or not each of these statements is correct. Explain your reasons. Write an improved statement if possible.

[P] Because the dice are fair, Team A and Team B both have a 50-50 chance of throwing a 5 in the last round.

[R] Because anything could happen, it’s impossible to know the chance of a team getting a 5 in the last round.

[Q] Because Team B hasn’t yet got a 5, they are more likely than Team A to throw a 5 in the last round.

Slide 12 © epiSTEMe 2009/11

Fig. 5: *Think about ‘Fives’ probabilities*

10 Lesson 3: *Tools for thinking about probability*

The opening two lessons in the module are intended to support exploratory activity aimed at informal development of basic ideas about probability through the example of throwing dice. The third lesson, *Tools for thinking about probability,* is designed to codify such ideas more authoritatively in the terms accepted in the field, and to consolidate their use by pupils. The guiding narrative is one of familiarising pupils with a system of representational tools that supports thinking about probability; these are specialised terms and diagrams which mediate accepted ways of thinking about the types of stochastic situation that pupils have already explored informally. One important reason for developing pupils' familiarity with these tools is to establish a shared representational repertoire to facilitate the clarity and cogency of the communication on which productive discussion and exchange depend.

Nevertheless, codification fits somewhat uneasily with an established English culture of mathematics teaching that emphasises learning informally through examples. This culture rarely gives priority to formulating mathematical principles explicitly, and tends to underplay the development of precise use of a shared technical language, as a rather confused extract from the official guidance to teachers on this part of the curriculum illustrates vividly (Fig. 6).

Know that if several equally likely outcomes are possible, the probability of a particular outcome chosen at random can be measured by:

$$\frac{\text{number of events favourable to the outcome}}{\text{total number of possible events}}$$

Fig. 6: Extract from official guidance on teaching probability (DfEE 2001, p. 278)

The first activity of this lesson has two aims. The first is to identify and organise key probabilistic ideas, using the reference situation explored informally in the preceding lessons. The second aim is to introduce systems of standard terminology and imagery that serve as tools for expressing these ideas effectively and using them to analyse situations efficiently. Originally it was envisaged that teachers would draw on experiences from the preceding lessons to cover an agenda sketched in the Teaching Notes, building up their own written summary over this part of the lesson. However, it became clear that such an approach was demanding on teachers, and that an explicitly prestructured approach might help them to pursue these aims in a more sustained and systematic way. The first structuring device is simply to provide a summary slide, *Summarise key ideas* (Fig. 7). Teachers use the interactive whiteboard facility to hide the text and gradually reveal sections; the intention is to organise whole-class interaction around reading a short section of the text, then putting it into one's own words and linking it back to experience from the preceding lessons. However, this was not a type of activity or style of interaction to which teachers and pupils were accustomed, and both teacher feedback and lesson observation showed that it risked degenerating into a leaden, cursory traversal of the key points.

Eventually, then, with a view to supporting more interaction with the text and the ideas involved, we moved the emphasis towards what was originally envisaged as an ensuing consolidation activity. *Adapt key ideas* involves revising the text of *Summarise key ideas* to relate it to the new example of a symmetrical spinner numbered from 1 to 8. Pupils work in pairs to fill the gaps in a parallel text (as exemplified in Fig. 7), followed by whole-class review. While teachers and pupils were again unaccustomed to this type of text-based activity, lesson observation indicated that it was more successful in generating engagement with the summary, primarily because it provided a structured task for pupils to undertake. However, in many respects, the experience that pupils gain through *Adapt key ideas* comes closer to routine consolidation than to the type of codification originally envisaged through *Summarise key ideas*.

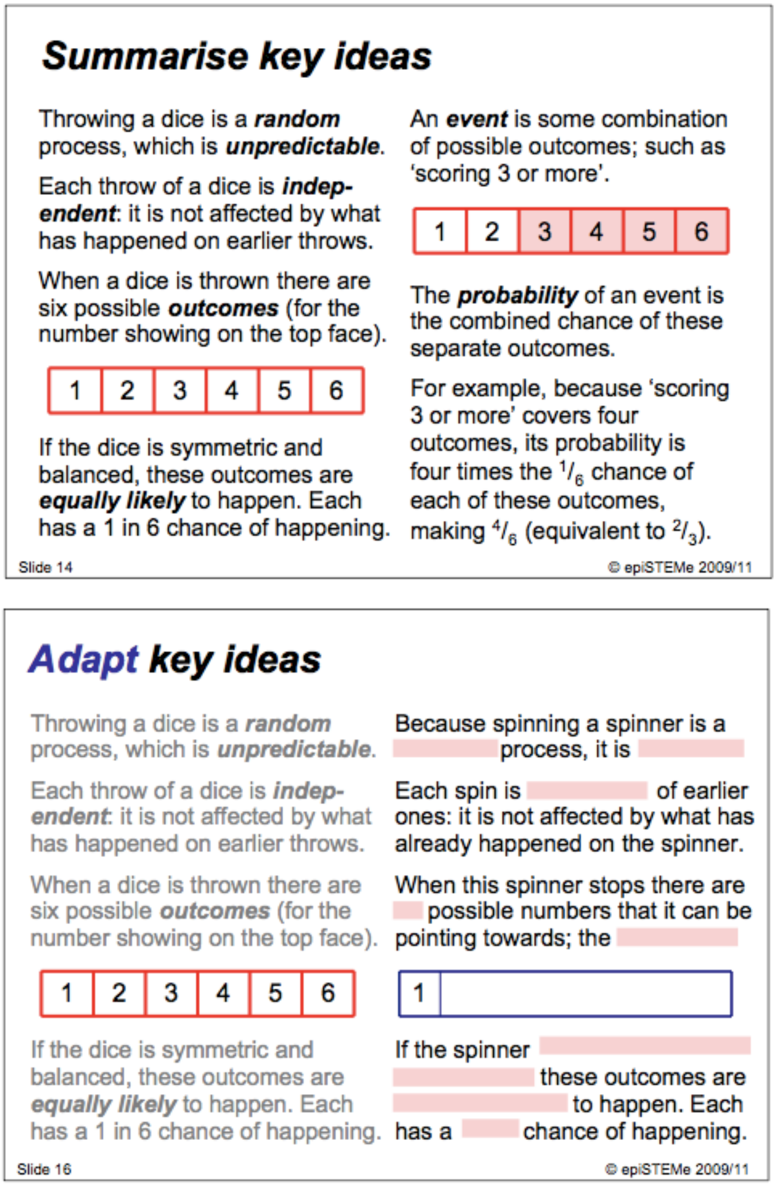

Fig. 7: *Summarise* and *Adapt key ideas*

Adapt key ideas provides a tightly structured lead into a central concern of the next activity in the lesson: developing the written organisation and expression of mathematical ideas. *Dodecahedral die* (Fig. 8) calls for pupils to work individually to solve a set of linked problems, and to present their solutions in writing. A whole-class review then considers teacher-chosen examples of pupil solutions, discussing both their mathematical substance and written expression. This activity is intended to consolidate use of the key ideas and terms made explicit in the earlier summary, and to develop pupils' capacity to provide an organised presentation of solutions in writing.

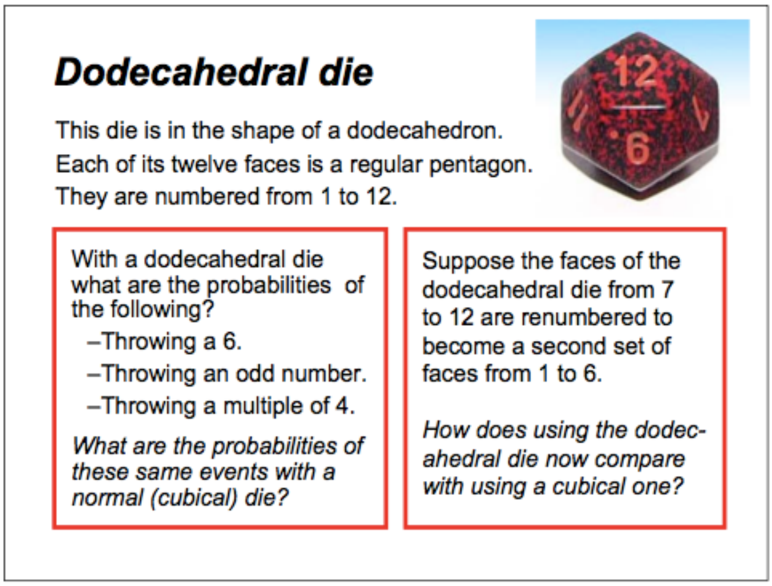

Fig. 8: *Dodecahedral die*

Further problems of similar type are provided as homework, with the suggestion that pupils discuss them with someone at home. For example, *Lucky numbers?* (Fig. 9) asks for reflection on the reasonableness of a statement made in promotional material for the National Lottery, as well as posing a more conventional set of probability problems.

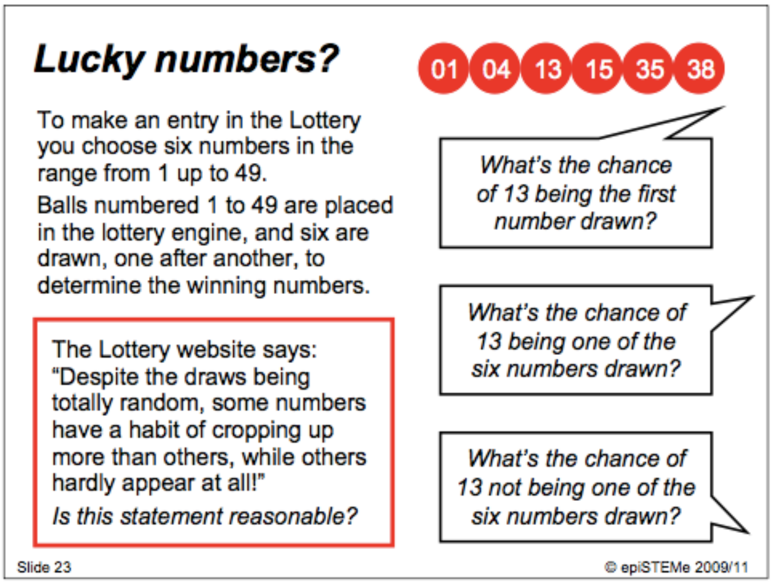

Fig. 9: *Lucky numbers?*

The last activity in this lesson, *Design your own probability scale* (Fig. 10), focuses on representing probability measures. It examines the coordination between figural, verbal and numerical components of representation, and counterposes differing verbal styles (such as the more frequential "happens half the time" against the more classical "even chance") and contrasting numerical forms (such as rational fractions against percentages). Feedback from teachers was that this small-group activity proved popular with pupils because it provides scope for personal choice in labelling the blank scale and choosing events to display on it. This type of mathematically reflexive task, which focuses attention on relationships within and between systems, provides a good basis for more explicit codification led by the teacher.

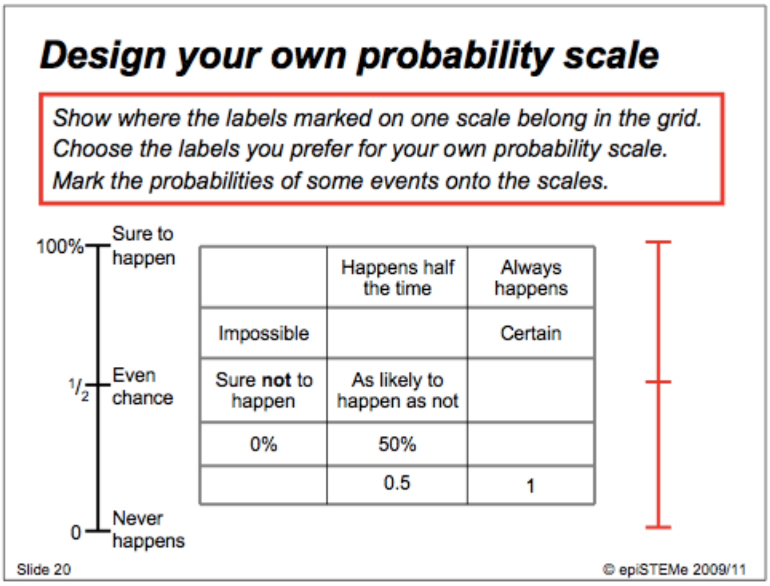

Fig. 10: *Design your own probability scale*

While this lesson marks the completion of the first cycle of exploration-codification-consolidation in the module, it does not quite complete the traversal of Level 5 topics. That takes place in the following lesson which develops the frequential approach to estimation of probability from empirical data, as well as addressing equiprobability bias (Lecoutre, 1992).

11 Conclusion

In this paper, we have sketched the wider framing and evolving design of the *epiSTEMe* probability module.

This wider framing positioned the module as one component of a research-informed pedagogical intervention intended to support improvement at scale. The core design brief was to produce a module that exemplified the *epiSTEMe* pedagogical model, took account of didactically relevant research, and respected crucial system constraints. This example of an evolving design contributes to developing wider understanding of what we call "re-design" research, conducted in ordinary school settings with relatively modest resources, and recognising that design for implementation at scale needs to take account of the existing state of the system – notably the people, structures, resources and practices already in place.

Perhaps the major design challenge was to incorporate the dialogic aspect of the *epiSTEMe* pedagogical model into the module. We took this as requiring a design that both introduced different perspectives on the topic and engineered productive dialogue between participants. A further contribution of our work, then, has been in developing research on dialogic talk and classroom teaching so as to take greater account of the managed development of pupils' understanding of the epistemic nuance of specific subject matter.

Previous research on probabilistic thinking provided valuable resources for developing this dialogic aspect of classroom exchanges. First, analyses of classical, frequential and subjective approaches to probability informed the interweaving of these perspectives over the opening lessons of the module. Second, analyses of fallacious reasoning about the specified curricular topics informed the design of tasks for small-group discussion and of advice on handling certain types of pupil response to other tasks. At the same time, in using current analytic frameworks for such thinking, our work identified what may be weaknesses in their value for design purposes: an overemphasis on foundational issues in the discussion of approaches (which is the reason why this paper uses ‘frequential’ rather than ‘frequentist’); too broad a use of ‘subjective’ to describe approaches ranging from the intuitive to those based on sophisticated judgement; and a disposition to attribute fallacious statements to deeper misconceptions.

Acknowledgements

Thanks are due to the Economic and Social Research Council which provided funding for the *epiSTEMe* project (RES-179-25-0003), to our colleagues in the *epiSTEMe* research team – Christine Howe, Stefanie Luthman, Neil Mercer, Fran Riga and Keith Taber – and to the teachers who generously volunteered to review and trial versions of this module.

References

Alexander, R. (2008). *Towards Dialogic Teaching: rethinking classroom talk* (4th edition). York: Dialogos.

Borovcnik, M., & Peard, R. (1996). Probability. In A. Bishop et al. (Eds.) *International Handbook of Mathematics Education* (pp. 239–288). Dordrecht: Kluwer.

Davis, E., & Krajcik, J. (2005). Designing educative curriculum materials to promote teacher learning. *Educational* Researcher, 34(3), 3-14.

Department for Education and Employment [DfEE] (2001). *Framework for teaching mathematics: Years 7, 8 and 9.* DfEE: London.

Fischbein, E., & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? An exploratory research study. *Educational Studies in Mathematics, 15*(1), 1-24.

Fischbein, E., & Schnarch, D. (1997). The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education, 28*(1), 96-105.

Franke, M., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. Lester (Ed.) *Second handbook of research on mathematics teaching and learning* (pp. 225-356). Charlotte, NC: IAP.

Hawkins, A., & Kapadia, R. (1984). Children's conceptions of probability: a psychological and pedagogical review. *Educational Studies in Mathematics, l5*(4), 349-377.

Howe, C., McWilliam, D., & Cross, G. (2005). Chance favours only the prepared mind: Incubation and the delayed effects of peer collaboration. *British Journal of Psychology, 96*, 67-93.

Konold, C. (1989). Informal Conceptions of Probability. *Cognition and Instruction, 6*(1), 59-98.

Konold, C., & Kazak, S. (2008). Reconnecting data and chance. *Technology Innovations in Statistics Education, 2*(1).

Lecoutre, M. (1992). Cognitive models and problem spaces in purely random situations. *Educational Studies in Mathematics, 23*(6), 557-568.

Mercer, N., Dawes, L., Wegerif, R., & Sams, C. (2004). Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal, 30*(3), 367-85.

Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve maths problems. *Language & Education, 20*(6), 507-27.

Office for Standards in Education [Ofsted] (2008). *Mathematics: understanding the score.* London: Ofsted.

Pratt, D. (2005). How do teachers foster students' understanding of probability? In G. A. Jones (Ed.) *Exploring probability in school: Challenges for teaching and learning* (pp. 171-190). New York: Springer.

Prestage, S., & Perks, P. (2008). HMI Ofsted report for Mathematics 2008, or why teenagers are maths dunces. *Proceedings of the British Society for Research into Learning Mathematics, 28*(3), 96-101.

Qualifications and Curriculum Authority [QCA] (1999). *Mathematics in the National Curriculum*. London: QCA.

Qualifications and Curriculum Authority [QCA] (2007). *Mathematics: Programme of study for key stage 3 and attainment targets.* London: QCA

Ruthven, K. (1989). An exploratory approach to advanced mathematics. *Educational Studies in Mathematics, 20*(4), 449-67.

Ruthven, K. (2011a). Research-informed pedagogical innovation at scale in school mathematics and science education. Paper presented at the British Educational Research Association. London, 6-8 September 2011. http://www.educ.cam.ac.uk/people/staff/ruthven/RuthvenBERA11paper.pdf. Accessed 12 Dec 2011.

Ruthven, K. (2011b). Using international study series and meta-analytic research syntheses to scope pedagogical development aimed at improving student attitude and achievement in school mathematics and science. *International Journal of Science and Mathematics Education, 9*(2), 419-458.

Ruthven, K., Hofmann, R., & Mercer, N. (2011). A dialogic approach to plenary problem synthesis. *Proceedings of the 35th Conference of the International Group for the Psychology of Mathematics Education,* 4, 81-8.

Ruthven, K., Howe, C., Mercer, N., Taber, K., Luthman, S., Hofmann, R., & Riga, F. (2010). Effecting Principled Improvement in STEM Education: Research-based pedagogical development for student engagement and learning in early secondary-school physical science and mathematics. *Proceedings of the British Society for Research into Learning Mathematics,* 30(1), 191-8.

Scott, P., Mortimer, E. & Aguiar, O. (2006). The tension between authoritative and dialogic discourse: a fundamental characteristic of meaning making interactions in high school science lessons. *Science Education, 90*(4), 605-631.

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T.-Y. & Lee, Y.-H. (2007). A meta-analysis of national research: Effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching, 44*(10), 1436-1460.

Seidel, T., & Shavelson, R. J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis research. *Review of Educational Research, 77*(4), 454-499.

Slavin, R., & Lake, C. (2008). Effective programs in elementary mathematics. *Review of Educational Research, 78*(3), 427-515.

Slavin, R., Lake, C. & Groff, C. (2009). Effective programs in middle and high school mathematics. *Review of Educational Research, 79*(2), 839-911.

Stein, M.K., Engle, R.A., Smith, M.S., & Hughes, E.K. (2008). Orchestrating productive mathematical discussions: Helping teachers learn to better incorporate student thinking. *Mathematical Thinking and Learning, 10*(4), 313-340.

Steinbring, H. (2006). What makes a sign a mathematical sign? An epistemological perspective on mathematical interaction. *Educational Studies in Mathematics, 61*(1/2), 133-162.

Threlfall, J. (2004). Uncertainty in mathematics teaching: the National Curriculum experiment in teaching probability to primary pupils. *Cambridge Journal of Education, 34*(3), 297–314.

Watson, J. (2005). The probabilistic reasoning of middle school students. In G. Jones (Ed.) *Exploring probability in school: Challenges for teaching and learning* (pp. 145-169). New York: Springer.