Chapter 13
Chance *Re*-encounters: 'Computers in Probability Education' Revisited

Dave Pratt

Institute of Education, University of London

Janet Ainley

School of Education, University of Leicester

Abstract In recognition of Rolf Biehler's contribution to probability and statistics education, we chose to re-visit his chapter in the edited book, Chance Encounters: Probability in Education. In particular, we examine three themes concerning the concept-tool gap, levels of access to concepts, which can be concealed by technology, and issues around demonstration and proof. We found many insights that resonate with current practice two decades later. Nevertheless, we argue that during that time, some progress has been made in how we can conceptualise the issues. For example, we discuss: (i) how it is now possible to unpick the metaphorical understanding that could emerge from the use of black boxes by reference to utility-based understanding; (ii) four principles that could inform how black boxes might be designed to support utility-based understanding; (iii) how the importance of explanation may overshadow a more traditional emphasis on proof.

Introduction

In 1991, Kapadia and Borovcnik invited leading researchers at the time to discuss issues on probability education. As a result, the edited book, *Chance Encounters: Probability in Education,* set out a critical review of the state of research in this field in 1991. In a sense the book demarcates an era dominated by the seminal research of psychologists such as Piaget, Fischbein and Kahneman and Tversky from a modern era in which educationalists in mathematics and statistics education have sought to resolve the many issues raised in the earlier work by conducting experimental studies to further theoretical development on how children learn to deal with uncertainty and how tools might be designed to support that learning.

In this chapter, we focus on Chapter 6, *Computers in Probability Education*, in the original book (Biehler 1991). Rolf Biehler reviewed the opportunities and constraints being offered by what was then a newly emerging technology. At the time of writing the book, Rolf was in his late 30s and now two decades later, it seems timely to revisit his text to consider whether the issues he raised have been resolved or at least come to be better understood, whether they have simply disappeared as the technology itself has improved, and indeed whether new opportunities and constraints have since emerged.

It is worth pausing to remind the reader about the context in 1991. We were early career researchers, studying the potential for using the latest digital tool, the laptop, in primary schools. Secondary schools in the UK possessed at best a room of computers, based on idiosyncratic operating systems, being increasingly used by departments other than mathematics for subjects like Computer Studies. Primary schools might have a single machine in each classroom, but with only a small monitor for use by individuals or pairs. Other than in unusual projects like our own, there was no handheld technology in schools apart from calculators; mobile phones were a distant dream (or nightmare for some of us). The internet was just beginning to emerge. It was not in use in homes or schools but university departments would have individual machines that could access 'JANET', the joint academic network.

This background context is very important as it would be easy to imagine that technological development since then has been so fundamental that little that was said about computers in probability education in 1991 could possibly be relevant today. Such a view would be far from the truth. As we will see, there remains a strong resonance between Biehler's analysis and the current situation, though our re-analysis perhaps also serves to emphasise where progress has been made.

Our approach to this re-analysis will not be exhaustive. We do not intend to work systematically through the many aspects of the original text, space would not allow such a method, but we do invite the reader to re-visit for themselves the original chapter. Instead, we plan to consider three themes, which emerged for us in our reading of the text. First, we will focus more explicitly on design by discussing the concept-tool gap, a construct that pervades the original Biehler chapter. Second, we will debate the extent to which technology hides or exposes underpinning statistical ideas by considering how design directs levels of access. It will be evident that the nature of the tools and indeed the tasks to which those tools are deployed provides a running theme throughout this chapter. Finally we will consider how demonstration, explanation and proof are played out when technology is used to support probabilistic learning. Inevitably such a discussion must be nuanced by considerations of design.

Design and the concept-tool gap

Biehler refers to the 'concept-tool' gap to articulate a concern that the meanings for uncertainty held by students are not well integrated with the performance demands made in curricula and examination syllabi. For young students, probabilistic algebra is still most often learned as a set of rules to be followed in order to get the right answer, leaving the concepts themselves distant and meaningless. The gap to which Biehler refers is, in this instance, between the concepts, such as independence, sample space and the Law of Large Numbers, and their use as tools to solve set problems. Later, students learn about hypothesis testing and confidence intervals but in most cases the learning is restricted to the application of little understood algorithms. Biehler raised this issue in 1991 but unfortunately the concept-tool gap continues, live and kicking, today. Although it is certainly not an issue only for probability education (consider, for example, the teaching and learning of algebra), Biehler was correct to identify this as a problem with particular relevance to probability education, when so often students and teachers conspire to answer examination questions in probability using methods that are understood only in instrumental ways.

Even in those early days, when computers were scarcely evident in schools, Biehler saw the special affordances that might mean the new technology could offer the potential to close the concept-tool gap. In particular, he listed three (p. 188):

1. the number of repetitions is easily increased so that uncertainty and variation in the results can be reduced; new kinds of patterns become detectable,
2. an extensive exploration is possible by changing the assumptions of the model, making further experiments, changing the way generated data are analysed etc.,
3. new and more flexible representations are available to express models and stochastic processes and display data with graphical facilities.

This list stands up to criticism even today. We might, for example, evaluate this list against the design of perhaps the most recent and innovative of computer applications in the field of probability education, *Tinkerplots 2*. In the earlier version of this toolkit, students were able to organize and graph data in largely intuitive ways. Tinkerplots 2 allows students additionally to model uncertain phenomena using probabilities represented as samplers, which can range from simple urn models to histograms and even to hand-drawn probability density functions. When a model has been created by the student, it can be used to generate data and analysed using the original Tinkerplots tools.

The example in Figure 13.1, taken from Konold, Harradine and Kazak (2007), uses urns and spinners to create a machine that could generate the varied attributes of cats. The likelihood of the two genders is set to be equal as are the probabilities of the three eye colours envisaged by the children. The various lengths of cats were represented as balls in an urn so that more common lengths could easily be modelled by the addition of extra balls with that length; older students might have used a sampler that would allow a continuous variable. Once the model has been created, the children are able to generate many cats with varying attributes.

The toolkit makes full use of affordance (1) above, in that it is very easy for students to generate as much data from their model as they like. Thus, small and large data sets can be compared. As imagined by Biehler, students can compensate for the lack of experience in the material world by generating extensive data in the virtual world of Tinkerplots 2. Furthermore, the use of different samplers whose parameters can easily be changed and methods of analysis whether based on collections of individual cases or grouped data exploits affordance (2) above. In response to affordance (3), Cliff Konold who leads the Tinkerplots development team not only allows students to use conventional representations of data, such as box, dot and scatter plots but also introduced novel representations in the light of research which suggests younger students might need transitional tools for organising and representing data prior to being able to intuitively manage the conventional approaches.

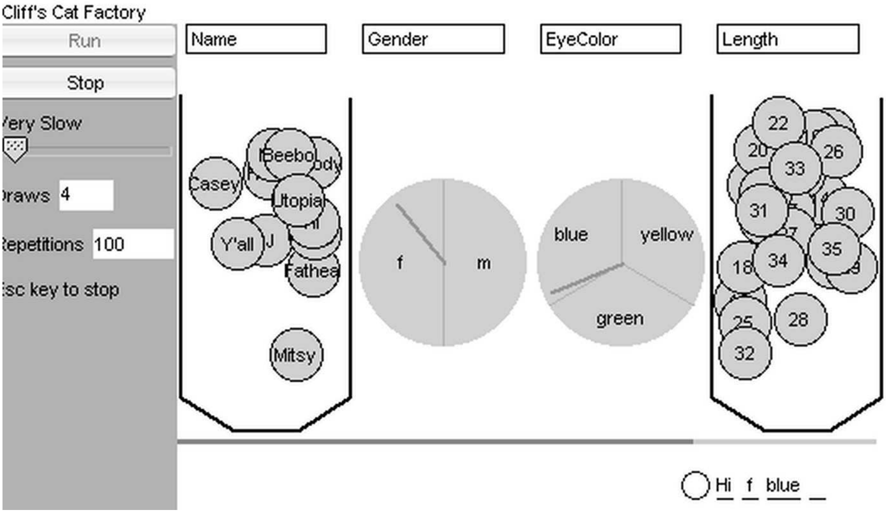

Fig. 13.1 A machine designed for making cats with various attributes

Biehler imagined that simulations of this type might enable students to close the concept-tool gap by making theoretical objects experiential. In a recent inaugural lecture, Pratt (2012) referred to 'making mathematics phenomenal' through the creation of on-screen objects and tools that can be explored and manipulated, showing the currency even now of Biehler's early insight. Indeed, both Biehler's and Pratt's ideas resonate with the constructionist movement (Harel and Papert 1991), which advocates the building of public entities by students to facilitate mathematizing. Papert's *power principle* (1996) argues that such an approach allows students to learn mathematics in a more natural way since they are able to draw on experience to construct mathematical meaning, which he sees as a parallel process to how humans most often learn outside of school mathematics. The constructionist vision is seen by Pratt (2012) as a challenge to design environments that facilitate making mathematics phenomenal and many of his examples come from probability education. For example, in building ChanceMaker, a domain of stochastic abstraction designed to research young children's meanings for chance, distribution and the Law of Large Numbers (Pratt and Noss 2002; 2010), he challenged 10-11 year olds to identify which of a range of virtual random generators were 'working properly'. The students were able to mend any so-called *gadget*, which in their view was not working properly, in particular by editing its workings box, an unconventional urn-type representation of the probability distribution. By editing the gadget, and reviewing the feedback in the form of the animation of the gadget, lists of previous results and charts aggregating those results, the students were able to learn *through use* about the nature of short-term and long-term randomness and come to situated understandings of the Law of Large Numbers and distribution.

Although, perhaps surprisingly, there is no mention of Papert's ideas in the 1991 text, there is direct reference to diSessa's (1986) work in mechanics. Biehler points out diSessa's defence against criticism that his Dynaturtle was not providing a 'real' experience of Newton's Laws. diSessa's counter-argument was that Newton's Laws are themselves a human construction; physics is not a direct perception but an intellectual abstraction, achieved over prolonged periods. The Dynaturtle offered an opportunity to construct Newton's Laws exactly because of its difference from the material world. It is often argued as criticism against virtual environments such as ChanceMaker that randomness on a computer is not the same as randomness as experienced through material objects like coins, spinners and dice. This is true but, because of the affordances identified by Biehler, virtual environments can offer experiences that facilitate the construction of powerful ideas like randomness more effectively than experience in the material world alone can afford.

In more recent years, Biehler and co-workers have designed integrated programmes of learning, built around simulation, in an attempt to bridge the gap (Batanero et al. 2005; Biehler and Hoffman 2011). Ainley, Pratt and Hansen (2006) have also looked at the concept-tool gap from the point of view of task design. In a sense, computer environments are only as good as the tasks to which they are deployed and so many of the design arguments that apply to the provision of tools also apply to how teachers and researchers choose and design tasks. They place some emphasis on the importance of designing the task so that it is seen as purposeful by the students. Many of the examples above would fit this design criterion but sadly many other tasks to which computer tools are deployed are no more engaging than their equivalents in traditional textbooks.

Ainley et al. would see the consideration of purpose as only part of the task design: the task must be directed towards the learning of mathematical or statistical concepts. Tasks that are purposeful may misdirect attention and can easily distract students from mathematics. They coined the term *utility* to describe the important element of task design that aims to focus the students' attention on the power of mathematical ideas to get stuff done. This type of learning is importantly different from tasks that focus on techniques, such as the instrumental learning of the Laws of Probability. It is also different from the type of relational understanding that enables students to appreciate the logic of why the mathematics works. To use a metaphor about car engines, utility would not relate to knowing precisely which controls to use, and in which order, to start the engine; neither would utility relate to understanding the engineering science of how an engine works. Utility would consist in an appreciation how useful an engine is in moving a car, which has obvious value in living one's everyday and professional life.

In some respects this example from everyday life appears ridiculous. The utility of an engine and the car it powers seems patently obvious: no one would embark on learning to drive without appreciating this utility. Ainley et al. argue that it is precisely the contrast between a real life example and the situation in mathematics and statistics classrooms, where students are generally required to learn to operate new tools before having any sense of their utility, which makes utility such an important idea. If mathematical and statistical ideas were experienced through purposeful tasks that made the utility transparent, their power might become rather more evident to students than is currently the case, and the concept-tool gap, if not closed, might be bridged in a different way.

However, this leads us in our re-analysis to a point of departure from the original text. Biehler imagined that simulation could lead to a closing of the concept-tool gap. We would now argue that this is over-simplistic. Simulations can provide purposeful tasks but students using tools such as Konold's Tinkerplots, diSessa's Dynaturtle or Pratt's ChanceMaker will learn about the scope and power of the statistical, mechanical or probabilistic concepts. Such knowledge is very important but is fundamentally different from the demands made by curricula and syllabi, which, we might argue, tend to heighten the concept-tool gap. It is a research question whether students with an appreciation of the power of probability will be better prepared to learn the formal knowledge and then grapple more effectively with the problems set in examination based on these curricula and syllabi. We would conjecture that this would indeed be the case but meanwhile the concept-tool gap remains largely inviolate.

Levels of access

Biehler (1991) in fact suggested that simulations could address the concept-tool gap by conceptualizing the simulation as a black box; "It may be possible that a metaphorical understanding of a 'black box' can guide a reasonable application" (p. 180). He made this argument in response to common criticism that concealing the computational algorithms relies on students accepting the outputs without knowing the underlying formulas, particularly common when people use statistical packages without understanding.

Indeed, the tendency for technology to hide the underlying mechanism is now common practice outside the classroom. Noss (1997) used his inaugural lecture to emphasise that technology often makes the mathematics that drives so much of everyday culture invisible, and that this trend demands the invention of new numeracies. A simple example takes place every minute of the day in the supermarket, when the customer goes through the checkout without either the shopper or the assistant needing to do any calculation. Not only does the computer total the cost of the goods purchased (and keep a stock check) but also the shopper hands over a credit card and no material exchange of notes or coins occurs. It is common for mathematics teachers to claim enthusiastically that mathematics is everywhere but, although more and more mathematics underpins the functions of everyday technology such as intelligent tills, never has this been less obvious to younger students, and indeed to the general public.

It is therefore an argument worthy of consideration that students cannot possibly come to appreciate the power of mathematics if the mathematics classroom mirrors the tendency to hide the algorithms, first through the use of calculators and second by wrapping the mathematics up in impenetrable black boxes. Nevertheless, in 1991, Biehler envisaged that through the use of, for example, a program that calculates Binomial probabilities, students might come to apply a metaphorical understanding without detailed knowledge of the formula.

What did Biehler mean by a ‘metaphorical understanding’? This is not clear but we would argue that students who were using the imagined Binomial program in this way might gain a sense of the scope and limitations of the Binomial Distribution. By playing with the program, they might never learn the formula, since that is concealed within the black box, but they might come to appreciate when the Binomial distribution is a useful tool and when it is not. We see a parallel here with the ideas of Meira (1998) who uses the idea of *transparency* (developed from the work of Lave and Wenger, 1991) as ‘an index of access to knowledge and activities’. Meira contrasts a view of transparency, or lack of transparency, being inherent in a tool or device, with one of transparency emerging through use:

> “… the transparency of devices follows from the very process of using them. That is, the transparency of the device emerges anew in every specific context and is created during the activity through specific forms of using the device.” (Meira 1998, p. 138)

Biehler’s ‘metaphorical understanding’ seems to be close to this notion of transparency. This is precisely the sort of knowledge and learning that Ainley et al. (2006) intended when they referred to utility-based understanding. Utility-based understanding will not emerge automatically and a good deal of careful design is needed to ensure that the tools and tasks lead to such an appreciation. Biehler of course referred to the Binomial program as an instance of a set of black boxes and indeed we also are not making a specific claim about the Binomial Distribution; on the contrary, we make a quite generic connection with utility-based understanding.

When considering the design of tools, one might ask what design decisions might be taken to avoid concealing the algorithms and instead foster utility-based understanding of probability? This question brings the reader explicitly to the issue of levels of access, headlining this subsection. We identify four elements of design, though implementation of any one in isolation is unlikely to be effective:

1. The use of evocative imagery that iconically carries a key aspect of the probabilistic concept;
2. The possibility of manipulating parameters of the simulation and receiving feedback that facilitates understanding of the behaviour of the probabilistic concept as instantiated on-screen;

3. Semantic layers that can be opened by the student who wishes to dig out a deeper meaning;
4. A blurring of control and representation.

An example of (1) above can be found in the innovative developments by Chris Wild's team to support the understanding of core concepts in statistical inference (http://www.stat.auckland.ac.nz/~wild/VIT/index.html). Their visual inference tools focus on how animation can be deployed to give a sense of variation. For example, a statistic such as sample mean can be plotted as a short vertical segment on a horizontal axis during continuous re-sampling. When the trace of the sample mean is captured, the effect is that gradually a box appears as the sample mean varies to the left and right, taking smaller and larger values. Some values of the sample mean occur more often and so there is increased intensity towards the centre. The vibrations of the box give an ongoing sense of the variation and yet the signal and noise apparent in the sampling distribution of the sample mean is visible.

In fact one can imagine a student altering the size of the samples being taken and re-taken and beginning to notice a relationship between the sample size and the width of the animated box that emerges. This is an example of (2) above. The facility to be able to manipulate the parameters in the black box is a minimal requirement if the level of access to the utility of the probabilistic concept is not to remain superficial. Above, Pratt's work with the ChanceMaker tool demonstrated how the students could change parameters such as the number of throws of the gadget and begin to appreciate from the graphical feedback that 'the more times you throw the die, the more even is its pie chart' (assuming of course that the probabilities of each outcome are in fact equal).

ChanceMaker allows the student to open up the gadget and change the workings (in other words the probability distribution). In one sense this is an extended version of (2) since it is possible to think of the workings box as a parameter. However, editing the workings box is a significantly more substantial act than simply changing a parameter. In that sense, deeper meanings may become apparent and so perhaps this is a simple example of (3) above. The design challenge here is to balance access with expressability. Wild's work with VIT is easy to use (though the underpinning ideas are far from simple) but students may be hindered as they have few opportunities to express their own ideas, a key design feature from the Constructionist perspective. ChanceMaker perhaps offers more opportunity for expressing personal conjectures about how the gadget should work through the challenge to mend the gadget, in particular by editing the working box.

A far more ambitious project to make available access to layer upon layer of meaning was diSessa's Boxer project (http://soe.berkeley.edu/boxer/). Boxer is a programming language, related to Logo but as part of a project aimed at developing an all-inclusive medium for computational literacy. The relevant feature here is how Boxer provides closets that by default are closed. Within these closets can be objects or pieces of code that may be best hidden from the naïve learner at the outset but may be of interest later. Since closets can themselves contain closets, there is potential for digging more and more deeply into the semantic layers of the concept. Although Boxer is not used widely, it is perhaps the tool which has taken most seriously the notion that the black box could be opened at the discretion of the user to uncover new layers of meaning. Biehler originally hinted that black boxes could, despite hiding underlying algorithms, prove useful to students who need to capture a metaphorical understanding. In trying to unpick that metaphor, we posit as one of the four design elements the notion that black boxes do not need to be so black but, aside from Boxer, few developments in this direction have yet materialised.

In fact, the Boxer closets could be opened to reveal the underpinning algorithms, offering a possibility of closing the concept-tool gap by, not only supporting utility-based understanding but beginning to appreciate the formulas themselves. Often principles in the design of tools can also inform task design. For example, a teacher might design a task involving measuring the time of flight of a sycamore seed dropped from different heights as part of an exploratory investigation. Several measurements of time might be taken for the same height and inserted into a spreadsheet. It would then be natural perhaps for the teacher to introduce the child to the ‘average’ function as a black box so that a better estimate could be calculated from a number of measurements. Inquisitive children might ask the teacher what the function does. A well-prepared teacher may have anticipated such a scenario and may have materials on hand that explain the mean average, and so open up the black box, either by exploring more systematically how average behaves for different number sets or by introducing methods of calculation.

When diSessa (1988) refers to integrating the formal and the informal, he has in mind, for example, that, in a computational medium such as Boxer, it is possible to execute mathematical formalisms within a generally expressive and creative environment. Two limited examples are: (i) the word processor, where it is possible to click a url embedded in the text and gain immediate on-line access to the internet; (ii) in Mathematica (http://www.wolfram.com/mathematica/), algebraic text can be executed to create simplifications and computations of that text. More generally, expressive environments that allow textual and graphical creation can also integrate mathematical and statistical formalisms so that the informal creative process, expressed through text and graphics, can be supported by the execution of formal algorithms. When such expressive environments contain tools and structures that behave in ways that reflect underlying mathematical or statistical principles, the tools are described by Noss and Hoyles (1996) as *autoexpressive*.

In (4) above, we take the notion of autoexpressive tools one stage further by operationalising this with the designer in mind. We suggest that a particularly felicitous way of integrating the formal and the informal is to blur control and representation. For example, in Logo, the young child controls the turtle by issuing commands such as `fd 50`. The command is both a means of control and a representation of how the turtle will move. In ChanceMaker, the workings box controls the behaviour of the gadget but, insofar as students learn to predict what will happen by inspecting the workings box, it becomes an unconventional representation of probability distribution. In both these examples, a key mathematical idea (linear distance; probability distribution) is placed exactly at the point of control so that the student can scarcely avoid engaging with the representation and gradually become aware of its meaning through use. These tools are indeed autoexpressive but to implement that the designer can privilege key representations by making them points of control.

By bridging the formal and informal in this way, the black box becomes rather more meaningful, especially in relation to utility-based understanding of the concept. These four design elements, especially if designed to work together, might begin to describe how the design of tools can create black boxes whose use could support metaphorical, or utility-based understanding, of probabilistic (and other) concepts.

Demonstration, explanation and proof

Biehler describes computers as supporting a new type of scientific method, which is characterised by an experimental style of working with models and data, and suggests that in such a method proofs may be valued to the extent that they offer explanations (p. 172). We understand this to mean that just as formulae may become black boxes when used instrumentally, so too may proofs, however rigorous, which do not support ways of understanding phenomena. In statistics, proof is evidence-based in contrast to mathematics where theorems are established logically from previously proven theorems. Nevertheless, a hypothesis test might be taken as a proof in its everyday sense but the test scarcely explains when the notion of a hypothesis test is obscure, the mathematical basis for the specific test used is not understood and when the procedure has been applied instrumentally. On the other hand, EDA techniques might enable re-presentation of the data that explains the basis of an inference.

We see Biehler's claim about an experimental style of working as closely linked to the ways in which computers support working with graphs and other visual images. Biehler himself provides a beautiful example of a visual simulation to solve the 'Abel and Kain' problem, which not only offers a way to calculate the relative probabilities of the two outcomes (1111, 0011) but also an explanation of why one is more likely than the other (p. 184).

The importance of explaining phenomena, rather than relying of the application of 'black box' formulae or proofs, is particularly pertinent in probability where intuitions play a significant role in thinking. Indeed 'Abel and Kain' only works as a problem because intuition might suggest that the two outcomes are, in fact, equally likely. An approach based on the application of formulae may produce an answer, but unless this solution explains the situation, intuitions are likely to be left untouched.

In the real world, the consequences of relying on intuitive understanding of probabilities may be problematic. Unfortunately Biehler's idea that computers may support an increased focus on subjectivist aspects of probability has not been widely realised, and school curricula still focus largely on coins and dice. An exception is the work of Pratt et al. (2011) on the design of computer environments to support learning and teaching about risk. They developed a tool to research mathematics and science teachers' knowledge about risk. They propose a scenario in which a fictitious young woman, Deborah, has a back condition, which might be cured through an operation. Deborah's dilemma lies in the fact that the operation could result in side-effects, some of which are severe. She needs to balance in some way the positive and negative outcomes with their likelihoods. The probabilities need to be estimated subjectively because the data that are provided give some indication of the likelihoods of the various outcomes but there are contradictions and discrepancies in the information. The teachers are encouraged to model what might happen if Deborah were to have the operation and what might happen if she did not. By running the model, the teachers are able to witness her possible futures and hence gain feedback on their subjective estimates and evidence about how the dilemma might be resolved.

In the years since Biehler's chapter was written, EDA has become a much more widespread pedagogical approach, supported by the substantial development in appropriate software, such as Tinkerplots and Fathom. The EDA approach foregrounds the use of visual methods to explore and identify patterns in data, though it is less clear the extent to which this encourages a search for explanation, rather than focussing on observing relationships. An issue that is becoming more widely recognised in statistics education (Pratt et al. 2008) is the potential for confusion between exploration of a set of data, which is *the whole population*, and exploration of a set of data, which is *a sample from which information about the population can be inferred*. Whilst it may be clear to the teacher that a task involves the second of these situations, careful design is required to produce tasks in which it is very clear to pupils that this is the case. The nature of what might count as proof, or convincing explanation will differ considerably in the two situations: if I am exploring the whole population I can make claims with a level of certainty which is inappropriate if I am working with a sample. Since the tools, both conceptual and technical, that are available can be applied in the same way to either situation, task design is crucial to enabling students to understand the importance of inference.

Conclusion

In re-visiting this chapter we have been struck many times by Rolf Biehler's vision in anticipating future possibilities offered by technology in the field of probability, and indeed in statistics and mathematics education more generally. For example, Biehler's list of affordances that position technology as being especially felicitous in closing the concept-tool gap is identifiable even in the most recent of statistics educational software developments. In particular, the notions that technology can make theoretical objects experiential and that virtual experience can enhance everyday experience still seem very current. Progress has been made on these notions so that the constructs of purpose and utility describe how such an emphasis generates a different sort of understanding from that typically demanded by curricula and syllabi. As a result the gap remains apparent and it may need further work on how to design levels of access, perhaps making use of how to blur control and representation, before those demands can be met in a pedagogically sound way. Some areas anticipated in Biehler's original text have not yet been developed so that, for example, there remains little evidence of the subjective use of probabilities being taught in schools.

There is also a realisation of how uneven and inconsistent the exploitation of that potential has been. In parallel to the development of flexible and creative applications such as Tinkerplots, much of the commercially produced software available to schools embodies approaches that prioritise technique and right answers above deep understanding. Technology which is hugely more sophisticated than that available in 1991 is often used for display and demonstration, rather than for developing the experimental style of working that Biehler envisaged. We would argue that in this context issues of design, of both environments and tasks, are as important as ever.

References

Ainley, J., Pratt, D., & Hansen, A. (2006). Connecting engagement and focus in pedagogic task design. *British Educational Research Journal*. *32*(1), 23-38.

Batanero C., Biehler R., Maxara C., Engel J., & Vogel M. (2005). Using simulation to bridge teachers´ content and pedagogical knowledge in probability. In *15th ICMI Study Conference: The Professional education and development of teachers of mathematics.* Retrieved from http://stwww.weizmann.ac.il/G-math/ICMI/strand1.html.

Biehler, R. (1991). Computers in probability education. In R. Kapadia, & M. Borovcnik (Eds.), *Chance encounters: Probability in education* (pp. 169-211). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Biehler, R., & Hofmann, T. (2011, August). *Designing and evaluating an e-learning environment for supporting students' problem-oriented use of statistical tool software*. Paper presented at the 58th ISI Session, Dublin, Ireland. Retrieved from http://isi2011.congressplanner.eu/pdfs/450437.pdf.

diSessa, A. (1986). Artificial worlds and real experience. *Instructional Science*, *14*(3-4), 207-227.

diSessa, A. (1988). Knowledge in pieces. In G. Forman & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Harel, I., & Papert, S. (1991). *Constructionism*. Norwood, New Jersey: Ablex.

Konold, C., Harradine, A., & Kazak, S. (2007). Understanding distributions by modeling them. *International Journal of Computers for Mathematical Learning, 12*(3), 217-230.

Kapadia, R. & Borovcnik, M. (Eds). (1991). *Chance encounters: Probability in education*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. *Journal for Research in Mathematics Education*, *29*(2), 121-142.

Noss, R. (1997). *New cultures, new numeracies,* London: Institute of Education, University of London.

Noss, R., & Hoyles, C. (1996). *Windows on mathematical meanings: Learning cultures and computers*. London: Kluwer Academic Publishers.

Papert, S. (1996). An exploration in the space of mathematics educations. *International Journal of Computers for Mathematical Learning*, *1*(1), 95-123.

Pratt, D. (2012). *Making mathematics phenomenal*, London: Institute of Education, University of London.

Pratt, D., Ainley, J., Kent, P., Levinson, R., Yogui, C., & Kapadia, R. (2011). Role of context in risk-based reasoning. *Mathematical Thinking and Learning,* 13(4), 322-345.

Pratt, D., Johnston-Wilder, P., Ainley, J., & Mason, J. (2008). Local and global thinking in statistical inference. *Statistics Education Research Journal*, *7*(2), 107-129.

Pratt, D., & Noss, R. (2002). The micro-evolution of mathematical knowledge: The case of randomness. *Journal of the Learning Sciences, 11*(4), 453-488.

Pratt, D., & Noss, R. (2010). Designing for mathematical abstraction. *International Journal of Computers for Mathematical Learning*, *15*(2), 81-97.