Supporting the Distributed Synchronous Learning of Probability: learning from an experiment

Eileen Scanlon*, Tim O'Shea*, Randall B.Smith† & Yibing Li*
*Institute of Educational Technology, Open University, Walton Hall, Milton Keynes, England, UK.
†Sun Microsystems Laboratories, Mountain View, California, U.S.A.

Abstract

This paper reports on the some experiments to explore the distributed synchronous learning of probability. We have been studying synchronous computer supported collaboration between adults using Kansas, a very general and powerful environment. We have implemented a shared simulation of a "game show" and used this with groups of adults ranging in size from 2 to 5 participants, collaborating through an audio or video link. We describe our research approach and give preliminary results from our ongoing experiments on the general question how collaborative work changes or is enhanced by different numbers of simultaneous users of such technology, and how collaborative work mediated by technology can support learning difficult concepts. We give examples of the strong prior conceptions of probability held by our students and how these influenced their work on the simulation. We indicate future directions to be taken in this work.

Introduction

In this paper we report on our current research with a system for supporting synchronous collaboration, Kansas, to help learners conduct joint experiments on a statistics task. The number of individuals simultaneously using such systems and whether or not they are co-located alters in fundamental ways the patterns of interaction. We are interested in a number of questions related to the use of such systems in education:

- How does such a system impact subjects' potential for solving problems,
- How does using such systems alter the pattern of interactions between subjects, and
- Are there pedagogical thresholds on scale of simultaneous use for particular learning environments that build on such technologies?

Our research approach involves selecting a technology to support collaboration, finding very hard problems with counterintuitive solutions and then conducting experiments with users systematically altering three key dimensions. The key dimensions of variation we are concerned with are

- the number of learners working together,
- whether or not they are physically co-located, and
- the bandwidth of the communication channels available to them.

Within this framework we are exploring the general question how collaborative work changes or is enhanced by different numbers of simultaneous users of such technology, and how collaborative work mediated by technology can support learning difficult concepts.

Kansas

We are working with a distributed classroom which allows several physically distributed users to move about in a 2D space (called "Kansas" because it is very extensive and very flat with several people it). It also has an acronym: Kansas: A Networked, Shared Application Space. Kansas supports real-time audio and video-links among participants as well as shared applications, so that moving together and apart in Kansas will make and break audio connections among users. In this way, a group can easily divide into subgroups for independent discussions. Users have their own window in which they can see their local portion of Kansas. This window may contain objects relevant to the subject matter. Each user sees a small local rectangular portion of Kansas. Users can scroll their viewpoints across the vast surface, causing their rectangles to overlap in order to collaborate, or can move away from others to work alone. Kansas supports interactive animation and its users can be anywhere in the physical world.

Smith (in press) writes *"It is designed to be simple and uniform, from the underlying object-oriented language (Self) through to the user interface. "* Kansas is written in itself - it is integrated with the Self programming environment, so anything can be modified arbitrarily while the system is running. Smith (in press) however points out that there is a tension between the uniformity of the system and *"the need to distinguish among the various roles (such as "instructor", "programmer" or "student.") "*

The experiments

The core of this study involves videoing adults at working together on a shared simulation implemented in Kansas in different physical locations. To complement these observations, subjects were interviewed and given individual pre and post test/questionnaires. The subjects were told.

"You are a game show contestant. You have won through to the final round and your final challenge is to choose one of three doors. Behind one but only one of the doors is a Mercedes. You announce your selection but before you open the door the game show host 'helpfully' opens one of the doors which was not the one you have chosen. It doesn't have a car behind it. What should you do, stick to your original choice or change?"

They were asked individually to make a prediction and to give a reason for that prediction. This instructional strategy was adapted from the Predict/Observe/Explain model for science learning (discussed in O'Shea et al., 1993). A similar prediction strategy also forms part of the distributed technology mediated learning model described by Vahey (1996). Then they were introduced to their partners and given a Kansas micro world to conduct experiments.

After the micro world experiment, students were asked to make a joint statement of their solution and to comment on whether it had changed from their individual statements. Then they completed an individual post experiment questionnaire to establish what their own opinion was. They were asked to state what they thought the best strategy for the game contestant to pursue. They also were asked to make a prediction about what the best strategy to pursue if the problem had four doors and the game show host opened two of them. In this paper we report our results on two dimensions of our study, scaleability and learning about probability.

We are also exploring the issue of the bandwidth of the communication channels of the communications channels that our subjects used. Previous experiments we conducted with SharedARK (Smith, 192, Smith et al., 1989) suggested the importance of eye contact in establishing successful joint working, but work by Whitelock et al. (1995) suggested that audio communication might be sufficient for students working on a shared simulation to establish adequate role and task division. Accordingly we have run groups where subjects communicated via audio and some with video communication. Our analysis of the differences between these groups will inform our discussion of bandwidth issues but these results are not presented in detail here. In this paper we focus on groups with audio communication only.

Scaleability

Current bandwidth limitations have restricted Kansas in other uses such as distributed tutored video instruction to user groups of size 10 or so, see Smith (in press). We have a version of Kansas which supports up to 5 simultaneously active users on our simulation. (This version could, of course, support up to 5 groups of users but we have not yet done this). These users can be given access to audio or both video and audio links between each of the 5 locations. We are therefore exploring what happens when groups of 2, 3 and 4 subjects work on a benchmark group statistical problem-solving task which is simple to state, understand and support in Kansas but which, like many of our standard tasks used in past studies (see O'Shea et al., 1993, Scanlon et al., 1993, Smith et al., 1992, Taylor et al., 1992) has a simple solution which is counterintuitive. The task we set involved deciding what advice to give a contestant who is in a game show situation. The situation is that, having chosen one of three doors which might lead to the grand prize, the contestant is shown behind another door which does not reveal a prize and is asked if they wish to change their choice to the third door. The groups of subjects are asked to explore the problem with the aid of a shared simulated game show setting, a shared note-taking tool and a remote human host who communicates via audio or video link who displays the consequences of their choices. The subjects hold many standard misconceptions associated with probabilistic reasoning. The experimental set-up works well and while the subjects find the technology easy to use, they mostly find the question difficult. We have collected data on pairs of subjects, threesomes and foursomes using the shared simulation augmented by audio communication. In addition we have collected data on pairs also using the video tunnel.

Preliminary Results on Scaleability

Learners behave differently when they use systems like Kansas from the way they behave when they are in the same room. The dynamics of groups of learners alter in very distinct and measurable ways, depending on the number of subjects who are simultaneously working together. One factor which seems to be important is the clarity and economy of the note taking methods devised or adopted by the group. Many groups made inaccurate summaries of data. In some groups, chronological and detailed records of every use of the simulation were made. The level of detail was such that they had to scroll in a shared window to view the data and they often "could not see the wood for the trees". These groups sometimes agreed on counterfactual summaries of the data that matched their expectations or prejudices, such as its "fifty fifty" either way. Other groups recorded unmanageable amounts of data and also agreed incorrect summaries. If groups devised simple highly focused ways of summarising the data that could be viewed in a single window, they were much more likely to identify and resolve mathematical misconceptions. For example, a two by two matrix of success/fail versus stick/change where the appropriate cell was updated by one of the subjects after each go at the game was particularly elegant. One group also used the shared note-taking tool to explore illuminating hypothetical cases such as the 100 door case. Our results so far indicate that not only do groups of size smaller that four solve the problem together more effectively using the technology, they also devise more effective shared tools to support their learning and exploration.

We have some indications that three students working together gain more incidental learning benefit than two or four. One three way division that was adopted was one person choosing the door, one person operating the simulation and the third person taking notes. We noticed that moving from groups of three to groups of four appears to add dramatically to the extra negotiation about how to organise the task. Our experience with groups larger than three is that the overhead of, for example, deciding which part of the task to do next, can grow out of all proportion to the actual problem-solving necessary to make progress. We have seen groups larger than two adopting a tactic of indicating preferences and choices by the location or movement of their mouse cursors on the shared screen. We are interested to discover whether other economical non-verbal methods for advancing the progress of joint problem solving are evolved by groups of learners or problem-solvers using systems like Kansas.

Learning about Probability

This part of the paper concentrates on an account of certain features of pairs working on the problem taken from our initial data collection episode involving 5 pairs of students working collaboratively on a shared simulation with audio contact only. These subjects all had a background of degree level or post-graduate work in science or psychology. This data forms part of our larger study looking at the scaleability of such shared simulations for collaborative working involving groups of between 2 and 5 users and with either video or audio communication (see also O'Shea et al, 1997).

The focus on the impact of investigative learning in statistics is a synthesis of skills, processes, procedures and concepts. Variable-based practical investigations help develop their knowledge and understanding of topics. Simulations on computers can allow many experiments to be conducted quickly to develop an understanding of statistical and probabilistic concepts. As this understanding can be affected by their ability to run and draw valid conclusions from statistics experiments, it is valuable to look at procedural competency in order to fully understand how this affects their conceptual development. So, in what follows we describe both the outcomes in terms of conceptual understanding and what the subject's behaviour in conducting the investigation tells us about their procedural competency.

Kahnemen and Tversky have worked for may years on the heuristics which people rely on while assessing the probability of an uncertain event (e.g. 1971, 1972). They say (1972, p431) "There is no simple story about how students reason." Falk (1983) presents evidence to show that intuition leads people astray in their judgement of randomness. In some ways, the use of heuristics can reduce the complexity of the tasks that people face when asked to make decisions based on their understanding of the likelihood of certain outcomes. However, research on these heuristics suggest that there are a number of tendencies which surface which lead to fallacious probabilistic reasoning. First, that people are willing to make judgements on small numbers of events, even on the basis of a single trial. Second, people appeal often to common-sense rather than mathematics. Third, people are influenced by how easily they can think of events happening- their own subjective judgement of frequency rather than the actual frequency determined by counting outcomes of trials. Fourth, negative evidence can be more powerful than positive. Fifth, people often make judgements based their views that some outcomes are more likely than others because of the processes involved. Sixth, people experience confusion between predicting the probability of an event and predicting what will happen at the next trial of an experiment (see e.g. Konold (1991) and Konold et al. (1993). They say judgements of 50% probability often indicate that a subject is unsure what will happen on the next trial and other very different percentages mean that the subject may think an event is either certain to happen or certain not to happen. Konold says (p 148) "50% is the midpoint of the yes/no decision continuum, and thus it means anything can happen.")

Some commentators have advocated building on the sound elements in students' intuitions rather than making them doubt their own judgement. For example, Falk (1982) asserts that students can guess very easily which of two events is more likely to happen. We would argue that it is difficult to see how this kind of approach would correct some of the fallacious reasoning tendencies described above. However, there is growing evidence that interactive simulations can improve students understanding of probability (e.g. Shaughnessy, 1992). Vahey (1996) has reported on the design of a system to deal with some of these reported problems. Our system was designed to allow subjects to engage collaboratively in experiments to solve a particular task. Working on the task revealed a number of initial viewpoints.

Initial Viewpoints

Subjects had a range of views about this task before working with the simulation. If we focus on the five pairs of subjects we studied, six out of ten believed that the best strategy was to stick to their original choice. Three believed that changing was the best strategy. One student argued that it made no difference.

Looking at the progress made by each student, the view expressed by their partner at the outset had a strong influence. Three out of five pairs were working with different opinions of the best strategy and this led in the main to productive exchanges. However in some cases the strongly expressed view of a partner was enough to outweigh the evidence gathered from working with the simulation. Another helpful feature of pairings with conflicting initial views was the beneficial effect that students realised they did not understand because they could not justify choices to partners.

Planning and executing the investigations

We found that the way subjects approached the planning and interpretation of the experimental results illuminating. There was a willingness to draw conclusions from small amounts of data (as little as 5 trials).

Students sometimes calculated their results to suit their prior beliefs. For example, one pair after conducting two experiments involving 20 trials had the following conversation.

"It's like 50% if you don't change and 2/3 if you do change. It's a bit strange."
"It adds up to more than 100."
"What I think is over the balance of time it will come to 50 each."
"I think its the other way round."
"As I said at the beginning you will have 50%. It doesn't matter how you do it, it won't change in the long term it will pile out to 50."
"I think if you do it very few times in a game show you change."

This extract illustrates some of the features of fallacious probabilistic reasoning which have been commented on before. The pair was experiencing difficulty in reconciling their past knowledge of probability e.g. that probabilities add up to 100%, that things work out over time with how they use this knowledge to judge what to do in the case of a single trial.

Reasoning about Probability

We found evidence of strong personal beliefs in luck. One pair who had conducted experiments to show that changing was always more likely to be beneficial, in the post test said they would still not change although they knew it would increase their chance of success because "it would be worse to lose if you changed from the right answer."

There was a strong tendency to over generalise the concept of randomness. One subject in particular had a very strong view of the concept of randomness. She interpreted it as meaning that as her choice of which door to open had been made randomly, it was not possible for any outside event to influence whether or not this choice was correct or not. *"I can't really see what the point is in trying to work out a strategy if it's all totally random."*

One pair ran 12 trials always changing, and one always "sticking". When encouraged to review their data they commented:

'We only won twice when we stuck....We won once up there and then we changed, we've won three, we've won four times when we've changed. I still think my first...I still think there's no strategy.

Although the data was in conflict with thesubject's view she persisted with it because of a belief that there could not be a strategy which would be of benefit.

Outcomes

Using the simulation with a partner was successful in moving the subjects' perception of what would be a successful strategy in four out of the five pairs. The fifth pair started with conflicting views about what would be the correct strategy to employ. One of them was so fixed in his view that either strategy was equivalent, resisted attempts to design an investigation to check this, rejected disconfirming data and succeeded in persuading his partner that his view was correct.

Concluding Comments

We are interested in how using this system has laid bare the prior conceptions held by our subjects on the topic of probability. We have presented evidence on how our subjects' interaction with each other and with the simulation influenced their understanding of the problem they were asked to solve. We are interested in exploring the possibility of scaffolding students at work on the simulation by augmenting the note taking and record keeping tools which they have available to them, and considering what further support we can offer students to help them visualise the problem more effectively. More generally we are interested in the usability of such systems for learning.

Environments such as Kansas that support synchronous distributed collaboration could play an important role in the support of distance learners. This potential role depends on the existence of working examples which allow us to study how much larger numbers of simultaneously active learners, distributed as individuals and small groups can productively learn together. On the other hand, there are simple arguments that say there must be threshold limits to the number of individuals and groups who can usefully be co-active. We are exploring whether scaling up of collaborative group size works and exploring the limits of this scaleability. We are interested in exploring further the ideas of scaleability. Some teaching and learning approaches scale much more easily than others. It is necessary to analyse each new example of computer supported collaborative learning with respect to how many students should be associated with each running instance, whether human support is necessary and what the effective human learning guide support arrangements should be.

Our initial findings indicate that the groups can become less effective in both process and product terms as the number of group participants increases above two or three. Larger groups find it harder to engage in constructive task division and in situations where there is a lot of initial consensus the larger groups display inordinate cognitive fixity and become totally committed to erroneous solutions that are contradicted by the operation of the micro-world. On the other hand, larger heterogeneous groups that are not dominated by individuals can display more flexibility and creativity than smaller groups. These findings are similar to those found in co-present groups working together but our preliminary finding is that these differences in group problem-solving dynamics are amplified by the use of the collaborative support technology. We are currently continuing with our experiments involving larger groups and our experiments involving video as well as audio communication so that we can further explore the dimensions of variation we have laid out as our current research agenda. We are considering as our refined research agenda for the future to add 2 new dimensions of variation and explore the way that groups of different sizes with participants chosen with matched or unmatched problem-solving styles and with relevant or no subject expertise work together using this new type of distributed computer supported collaborative environment.

References

Falk, R. (1983) Probabilistic reasoning as an extension of common-sense thinking. In *Proceedings of the Fourth International Congress on Mathematical Education*, Birkhauser, Boston, p 190-198.

Kahneman, D. and Tversky, A. (1972) Subjective probability: a judgement of representativeness. *Cognitive Psychology*, 3, 430-454.

Konold, C. (1991) Understanding students' beliefs about probability. In E. von Glaserfield (ed.) *Radical Constructivism in Mathematics Education*, 139-156. Kluwer Academic Publishers.

Konold, C., Pollatsek, A., Well, A., Lohmeier, J., Lipson, A. (1993) Inconsistencies in students' reasoning about probability. *Journal for Research in Mathematics Education*, 24 (5), 392-414.

O'Shea, T., Smith, R., Scanlon, E., and Li, Y. (1997) Exploring the scaleability of object-oriented enviroments that support synchronous distributed collaboration, Paper presented at EDMEDIA-97, Calgary.

O'Shea, T., Scanlon, E., Hennessy, S., Twigger, D., Byard, M., Driver, R., Draper, S., Hartley, R., Mohamed, R., O'Malley, C. (1993) Twenty nine children, five computers and a teacher. In Edwards, D., Scanlon, E. and West, R. (eds.) *Science education: teaching, learning and assessment,* London: Paul Chapman Publishing.

Scanlon. E., O'Shea, T., O'Malley, C., Smith, R. and Taylor, J. (1993) Running in the rain- can a shared simulation help to decide? *Physics Education*, *28*, 107-113.

Shaughnessy, M. (1992) Research in probabilistic and statistics reasoning: reflections and direction. In Douglas Grouws (ed.) *Handbook of Research on Mathematics Teaching and Learning.* Macmillan: New York.

Smith, R. B, O'Shea, T. and Scanlon, E. (1987) Experiments in teaching physics with alternate realities. *Proceedings of the Third Artificial Intelligence and Education Conference*, Pittsburgh, June.

Smith, R.B. (1990) The Alternate Reality Kit: An Animated Environment for Creating Interactive Simulations. In O. Boyd-Barrett and E. Scanlon (eds), *Computers and Learning* Addison Wesley: London.

Smith, R.B. (1992) A prototype futuristic technology for distance education. In Scanlon, E. and O'Shea, T. (eds.) *New directions in educational technology* Springer Verlag.

Smith, R.B. (in press) Kansas: a dynamically programmable multi-user virtual reality. In O'Shea, T. and Scanlon, E. (eds.) *Virtual Learning Environments and the Role of the Teacher*, UNESCO Press.

Taylor, J., O'Shea, T., O'Malley, C., Scanlon, E. and Smith, R. (1991) Discourse and Harmony: preliminary findings in a case study of multi-media problem solving. In R. Glanville and G. de Zeeuw (eds.) *Interactive interfaces and human networks , University of Amsterdam*, 1993.

Tversky, A. and Kahneman, D. (1971) Belief in the law of small numbers. *Psychological Review*, 76, 105-110.

Whitelock, D., Scanlon, E. and Taylor, J. (1995) The influence of audio communication on students solving problems on collisions. *Journal of Association for Learning Technology*, March.

Vahey, P. (1996) How students' conceptions of fairness are informing the design of an interactive instructional probability environment. *Paper presented at the American Educational Research Association conference*, April, New York.

Acknowledgements

We gratefully acknowledge the technical support of David Perry during the current round of experiments with Kansas at the Institute of Educational Technology and thank Sun Microsystems Laboratories for access to the Kansas software.

Author's Addresses

*Eileen Scanlon**, *Tim O'Shea**, *Randall Smith*†
*& Yibing Li**

*Institute of Educational Technology,
Open University, Walton Hall, Milton Keynes,
Bucks, MK 7 6AA, U.K.
e.scanlon@open.ac.uk

†Sun Microsystems Laboratories,
2550 Garcia Ave.,
Mountain View,
CA 94043,
U.S.A.
randall.smith@sun.com