PERGAMON

Computers & Education 33 (1999) 153–169

COMPUTERS & EDUCATION

www.elsevier.com/locate/compedu

Help seeking, learning and contingent tutoring

H. Wood*, D. Wood

Centre for Research in Development, Instruction and Training, Department of Psychology, University of Nottingham, Nottingham NG7 2RD, UK

Abstract

The focus of this paper is the application of the theory of contingent tutoring to the design of a computer-based system designed to support learning in aspects of algebra. Analyses of interactions between a computer-based tutoring system and 42, 14- and 15-year-old pupils are used to explore and explain the relations between individual differences in learner–tutor interaction, learners' prior knowledge and learning outcomes. Parallels between the results of these analyses and empirical investigations of help seeking in adult–child tutoring are drawn out. The theoretical significance of help seeking as a basis for studying the impact of individual learner differences in the collaborative construction of 'zones of proximal development' is assessed. In addition to demonstrating the significance of detailed analyses of learner–system interaction as a basis for inferences about learning processes, the investigation also attempts to show the value of exploiting measures of on-line help seeking as a means of assessing learning transfer. Finally, the implications of the findings for contingency theory are discussed, and the theoretical and practical benefits of integrating psychometric assessment, interaction process analyses, and knowledge-based learner modelling in the design and evaluation of computer-based tutoring are explored.

1. Introduction

QUADRATIC is a computer-based learning environment designed to help 14- and 15-year-old pupils with little prior experience of algebra to develop some understanding of the quadratic function. This aspect of algebra was chosen for two main reasons. The first is that we know that it is hard to learn and that, once understood, it plays an important role as a basis for

* Corresponding author.

PII: S0360-1315(99)00030-5

further mathematical development. It has applications in many other fields of knowledge such as geography, physics and economics. Consequently, it provides an educationally significant and challenging domain for testing theories of instruction. The second reason is that the research literature offers a well developed theoretical account of the role of multiple representations or 'modes of knowing' in teaching and learning the quadratic function, coupled with associated materials and activities. Since our aim here is to evaluate principles of tutoring rather than to develop new curricular approaches and materials, the quadratic function offered a well prepared testing ground.

The approach to system design underpinning QUADRATIC is unusual in that it implements and evaluates principles derived from investigations into human tutoring. This contrasts with the more common strategy for developing tutoring systems, which typically derive their pedagogical principles from a model of learning. The relative merits of these two approaches, together with the advantages we could expect to achieve by combining them, will be the subject of much of the discussion section.

1.1. Contingent tutoring and help seeking

The tutoring principles implemented and evaluated within QUADRATIC are based on the concept of contingent support for learning, and derive from numerous empirical investigations of face to face tutoring (Wood, Bruner & Ross, 1976; Wood, Wood & Middleton, 1978; Wood and Wood, 1996a) and our earlier work on computer-assisted tutoring (Wood, Shadbolt, Reichgelt, Wood & Paskiewicz, 1992; Wood, 1996).

The basic ideas are deceptively simple to state, though hard to achieve in practice. When a learner has been set or is trying to achieve a goal and seems to be 'in trouble', then the contingent (human) tutor immediately offers help. If the learner does not seem to understand that help and remains 'in trouble', then more explicit instruction immediately follows. After about three cues of increasing levels of explicitness (or 'depth' or intrusiveness) the tutor will provide the answer or physically demonstrate the next step in the task. Moving in to give more and more help in time of need is one aspect of contingency. The other is in drawing back, 'fading', in order to leave more and more responsibility to the learner. Whenever the learner succeeds, the level of help is decreased. Most of the time this may involve the tutor doing nothing more than confirming correct responses, but if the learner has just needed a great deal of help with one problem, this is treated as a signal that s/he may need a bit of (unsolicited) help with the next.

Although this description of contingent instruction is couched in terms of principles consistent with what the *tutor* does, the achievement of contingent tutoring is in reality an *emergent property* of the interactions between tutor and learner. Both are responsible for the attempt to maintain focus on a common task, the pursuit of shared goals, and the maintenance of mutual understanding. Where tutors are perfectly contingent, it follows that it is the learners who 'drive' them to act as they do. Thus, a relatively knowledgeable pupil might proceed successfully through the tutorial tasks with little or no help. A struggler would probably, if they were seeking help strategically, drive the system to support them with explicit hints and answers until they reached the point where they were able to proceed alone.

In face-to-face interaction the tutor is able to exploit not only a learner's task actions but also their verbal and non-verbal acts of communication, both deliberate and spontaneous, in order to assess whether they seem likely to be on task, or are confused, concentrating, thinking or resting. In computer-based tutoring, the program does not have access to much of this activity. More specifically, information about when the learner might need help is severely limited, and it is particularly difficult to know how to program the system to make decisions about what to do when the learner remains inactive.

In the absence of information about the learner's likely state of mind, one option is to impose some arbitrary limit on the time allowed for inactivity, after which the tutor intervenes. However, this would be unlikely to result in suggestions which are contingent upon the learner's mental state. The other option is to leave decisions about *when* to seek help to the learner. The tutor then decides *what* help to provide and at what depth.

This option, which is the one taken in QUADRATIC, places more responsibility for regulating tutorial interactions on the learner than is likely to be the case in face-to-face tutoring where the tutor is on hand to 'read' and interpret any significant signs of the learner's likely mental state. The risk here is that, by leaving help seeking to the learner, the additional cognitive demands, together with the potential threat to self-esteem, could act as impediments to learning.

Investigations of help seeking in face-to-face tutoring suggest that these risks are real ones and, further, that lower achieving learners are most likely to be disadvantaged (Nelson-Le Gall, 1985; Nelson-Le Gall, Kratzer, Jones, & DeCooke, 1990; Puustinen, 1998; Winnykamen, 1992, 1993). For example, differences in prior achievement may be associated with less learner awareness of their need for help, and/or a greater reluctance to seek help when in need. This being the case, it follows that higher achievers are more likely to help the tutor to locate the 'upper bounds' of their 'zone of proximal development' (or, in our terms, to help define their own region of sensitivity to instruction) by signalling when they are in need of assistance. If so, such learners are likely to perform better under tutoring both because they start with a more robust knowledge of the domain, and because they are better able to help the tutor to create a learning environment which is contingent on their needs.

Looked at from this perspective, the potential theoretical significance of learner help seeking is considerable. It provides a means of specifying in more detail how learners contribute to the emergence of more or less contingent tutorial interactions. It also identifies processes of interaction which could serve to integrate a social and communicative dimension into our theories of learning; a dimension which may play a part in explaining individual differences in achievement and the processes whereby these impact on collaborative learning.

One aim of this investigation is to assess the extent to which learners' use of contingent, on-line help is instrumental in helping them to learn. We ask, for example, whether there is evidence that any learners were 'abusing' help by driving the tutor to provide answers without first trying to solve the problems themselves, or, conversely, evidence for learners 'refusing' help by persisting in unsuccessful actions when requests for help might facilitate their learning. Further, by comparing the performance of learners with relatively high and low levels of prior knowledge of the domain we will explore the possibility that errors and help seeking may exert a differential impact on learning outcomes for Low and High groups, i.e. that we will find evidence of 'aptitude $\times$ process' interactions.

A second aim is to show how on-line help and help seeking can be exploited to evaluate the generalisation of learning. QUADRATIC was designed to explore the impact of tutoring on tasks which are conceptually and mathematically related to the expansion of the quadratic function. Specifically, the tests of generalisation included problems which deal with squares which are getting smaller rather than bigger (i.e. which involve quadratic expressions with negative terms, such as $(x-n)^2$), and with cubic algebraic expressions such as $(x+n)^3$. Is there evidence that contingent tutoring with the initial quadratic problems facilitates learning on these tasks? Our main aim here is a methodological one: to demonstrate how a learner's use of on-line help can augment indexes of learning based on success and failure. Put another way, we seek to show how it is possible for a tutoring system to exploit evidence from its learner's patterns of help seeking to draw inferences about the extent to which it is supporting transfer of learning from one set of tasks to another. With QUADRATIC, the aim is to help enhance the learner's conceptual understanding of algebraic notation and to develop some sense of how changes in more than one dimension of geometric figures can be modelled using algebra. In the next section, we summarise the theoretical background for this aim, together with more details about how QUADRATIC was designed to achieve it.

1.2. Designing the learning domain

The logic underpinning the use of representations in QUADRATIC derives from principles advanced by Dienes (1960; Dienes & Jeeves, 1970) and implemented as an intervention programme by Bruner and Kenney (1965). This general conceptual framework has subsequently been extended to form the basis of more ambitious research and intervention programmes by Lesh and others (Lesh, Landau & Hamilton, 1983; Behr, Herel, Post & Lesh, 1992). In relation to our computer-based implementation of this framework, the analyses offered by Kaput (1992) and Ainsworth (this volume) provide a context within which to locate both the claims for, and the acknowledged limitations of, the multiple representations used in QUADRATIC.

QUADRATIC exploits area as a 'reference field' (Kaput, op. cit.) for 'grounding out' algebraic expressions of the quadratic function. The quadratic function, in turn, is intended to provide a mathematical 'model' for the perceptual effects of operations which serve to increase the area of squares. By varying the absolute size of squares used as a referent for the x square ('x^2' term) the idea is to exploit Dienes' principle of perceptual variability (since size is variable) in relation to a series of cases where the *configuration* formed (e.g. for an $(x + 1)^2$) is perceptually similar, in an attempt to get across the idea that the configuration has invariant properties beyond the surface differences in size (see Fig. 1).

In this way, one pedagogical ambition of QUADRATIC is to exemplify equivalencies implicit in mappings between the algebraic and geometric representations. For example, both the 'unexpanded' $(x+n)^2$ and the elements of the expanded expression (x^2, $2nx$ and n^2) refer to the completed configurations in Fig. 2. The terms on each side of the equation also, of course, express the equivalence between the two algebraic expressions. Thus the equivalence expressed by the '=' sign in the algebraic equation can be exemplified by two 'readings' of the geometric figure (one demanding attention to the $(x+n)^2$ treated as a gestalt, and the other to the elements of the square as equivalents to the elements in the expanded side of the equation). Since this duality of mappings holds true across all the examples used, despite variations in the

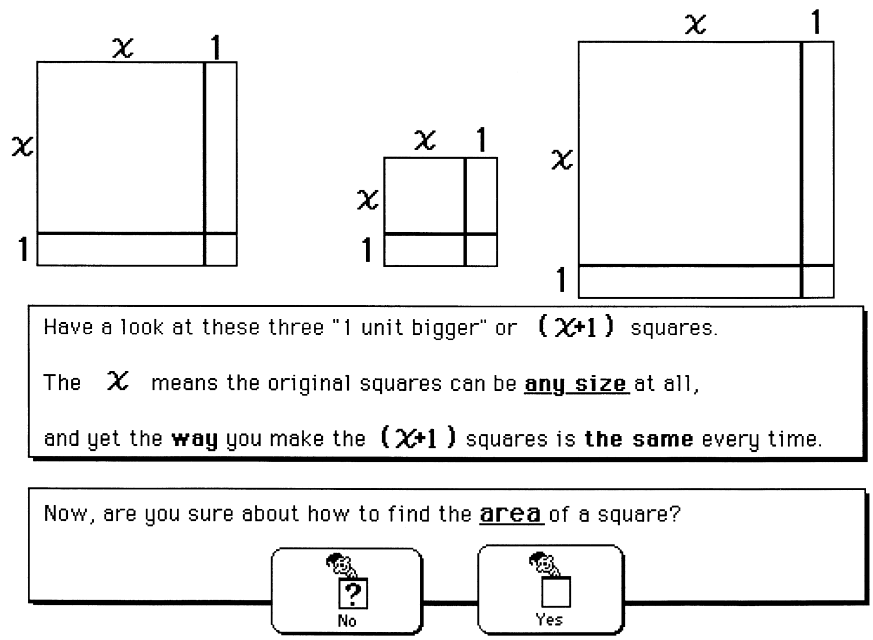

Fig. 1. Review of $(x + 1)$ squares.

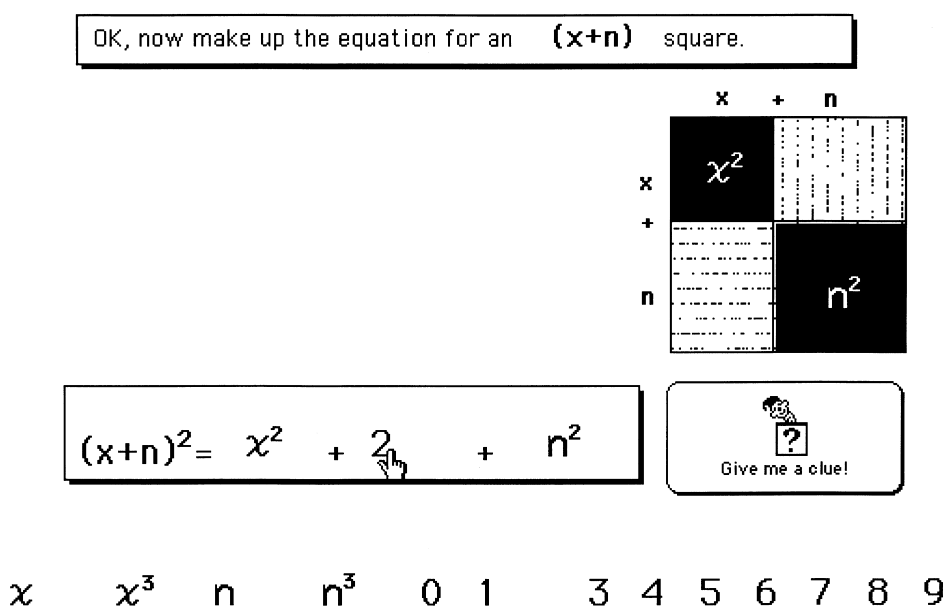

Fig. 2. Constructing the expansion of $(x+n)^2$.

absolute size of 'x' and the numbers of elements added to it, the pedagogical ambition is also to help the learner to begin to build up an intuitive sense of a 'variable' as a 'theorem in action' (Vergnaud, 1982). Having assembled the equation for $(x+n)^2$ (Fig. 2) learners were then invited to choose their own values for x and n (in the range 0 to 9, producing changes in the size of the squares and in the equation) and further, to choose their own letters for variables: "Just in case we have led you to believe 'x' and 'n' are special letters...". The aim is to exploit the use of these representations as a basis for initial learning of quadratic expressions, as well as providing an opportunity for the more knowledgeable learner to 'apply' this algebra to novel situations.

What QUADRATIC does *not* attempt to do is to teach any rules or algorithms for expanding the quadratic expressions themselves. Thus, there is no attempt to teach 'syntactic rules' which serve to map the pairs of equivalent algebraic expressions on to each other (i.e. where either side of the equation is the only 'reference field' for the other). This neglect of algorithms is not based on any assumption that teaching such 'within sign system' syntax is pedagogically undesirable or unsound. It is simply the case that they are not a focus of theoretical interest within the current QUADRATIC investigation. Similarly, it would be quite possible, and perhaps pedagogically desirable, to exploit Cartesian, graphical representations to serve as another reference field (or modelling representation) to elaborate upon the patterns and regularities which can be found, for example, in different 'families' of quadratic expansions. Here, too, the effects of such elaborated uses of representations are not a focus for interest; not because they are considered unnecessary for understanding, but because we needed to keep the design and evaluation of QUADRATIC within practical limits.

1.3. Design of the contingent tutoring regime

The tutoring component of QUADRATIC rests on an implementation of the rules for contingent provision of help on request, as outlined above. Help involves one of five levels of hint, varying in specificity from immediate and automatic general feedback (either to acknowledge a successful operation or to mark an action which the system does not recognise as task appropriate) through increasingly specific instructions, culminating in a demonstration of the next step. Any help sequence is solicited by the learner. The level of help provided by the system is determined by the learner's recent history. Thus, if the system had provided a level 3 hint which the learner followed with a successful move, it would next offer a level 2 hint. If, instead, the learner had requested further help, the system would offer level 4, and so on.

Learners are free to work at their own pace. Although the sequence of problems is predetermined as outlined above, the learner is allowed to bypass any problem, or to abandon a problem which they have started, whenever they wish.

2. Design

2.1. Participants

The analyses reported in this paper are based on a sample of 42, 14- and 15-year-olds drawn from two medium-sized secondary schools; one in a mainly middle-class residential area and the other from an urban area on the edge of a large industrial city.

2.2. Procedure

Each participant had two 30–50 min sessions with QUADRATIC, one day apart. Before and after each tutoring session, and again after about four weeks, pencil and paper 'probe tests' were administered, in which learners were asked to expand the quadratic function $(x+n)^2$.

Each participant was also pretested using an algebra test based on the *Chelsea diagnostic mathematics tests: Algebra* (Brown, Hart & Kuchmann, 1984). This was administered again after about four weeks (alongside peers from the same classes who served as a matched, non-treatment control sample).

On the basis of the pre-test the sample was divided into two subgroups: A 'High' knowledge group who scored at or above the median score (n = 22) and a 'Low' knowledge group (n = 20).

2.3. Learner–system interaction measures

Most analyses involve only the data from interactions between learners and QUADRATIC up to and including the expansion of $(x+n)^2$ (all learners got this far on the first day). Separate analyses involve similar data from the same learners when they moved on to tasks involving $(x+n)^3$ and $(x-n)^2$.

The system logs every operation (success, error and request for help) by the learner together with the times between them. This yields the following process measures:

- mean time per operation (used to assess speed of progress through the tutor)
- frequency of successes
- frequency of errors
- frequency of help seeking
- average level or 'depth' of help provided by the tutor
- latency of help seeking (the average time between requests for help and their previous action with the tutor)
- the tendency to seek help rather than risk an error is estimated: help/(errors + help)

2.4. Learning outcome measures: expansion of the quadratic function

Scores were derived from each of the five probe tests. For each attempt to expand the quadratic, one point was awarded for each of the five elements in the $(x+n)^2$ expansion: x^2, 2,

n, x and n^2. In addition, an average probe test measure, the mean score over the four attempts following any tutoring, was calculated. It is possible that learners might do nothing more than simply memorise this pattern, in which case we would expect no relationship at all with prior knowledge or any of the process measures.

3. Results

3.1. Prior knowledge, help seeking, errors, speed and learning

Table 1 provides the general profile of how pupils of varying knowledge levels performed with the tutor and how far these process measures then relate to recall of the expansion of $(x+n)^2$.

Pupils with more prior knowledge of the domain proceeded faster through the tutorial tasks, asked for help less often and made relatively few errors. These pupils also tended to master and remember the expansion of the quadratic function more easily. Conversely, as we would expect given a strategic use of help, learners with less prior knowledge sought help more often, and were provided with more specific help, than their peers with more prior knowledge. The frequency of successful problem solving did not correlate with pre-test, outcomes or the other process measures since almost all pupils (eventually) succeeded on all tasks and subtasks.

Multiple regression analyses allow us to investigate whether a combination of these basic measures can give us a model of learner ability which might be used in future systems to complement other means of testing progress. Putting together time per operation with frequency of errors and help seeking gives us a model which accounts for 63% of the variance in pre-test scores (and all three process measures are independent contributors). Thus if test scores were not available, future systems might use such simple measures to estimate pre-existing knowledge levels.

How does this model fare in predicting performance on the probe tests? It can account for 30% of the variance, and if we add latency of help seeking this rises to 43%. The pre-test *alone*

Table 1
Correlations between process measures and ability and outcome measures[a]

	Pre-test	Average of probe tests
Pre-test		0.71***
Time per operation	−0.51***	−0.26
Number of successes	−0.07	−0.22
Number of errors	−0.49**	−0.48**
Number of help requests	−0.34*	−0.16
Help/(errors + help)	−0.29	−0.07
Average depth of help	−0.49**	−0.27*
Latency	−0.13	0.23

[a] *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

accounts for 50%, and if we add this to our four process measures, we can now account for 62%.

We also checked that our four-factor process model predicted outcomes within both High and Low knowledge groups (57%, $P < 0.01$, and 35%, $P < 0.05$).

3.2. Partialling out prior knowledge

Pre-test scores predict process measures and outcomes so well that in order to look for independent effects of help seeking, etc., on outcomes we must statistically partial it out. This leads to a quite different story than when considering raw correlations alone.

With variance due to prior knowledge partialled out, error frequency (but not help seeking) remained a significant, negative correlate of outcomes. More surprisingly, the mean time per operation, our measure of speed of progress through the tutor, changed from a significant negative to a significant positive correlate of outcomes. Thus, because more knowledgeable learners worked their way more quickly through the tasks and also learned more, the relation between the 'raw' time per operation measure and outcomes was a negative one; learners who worked faster had better outcomes. However, with the effects of prior knowledge partialled out, positive learning outcomes were associated with a slower rate of operation. Another process measure, which was independent of prior knowledge, was time spent before seeking help. Here too, slower times were associated with better learning outcomes.

3.3. Aptitude × process interactions

As the results in Table 2 illustrate, the relations between the learner–system interaction process measures and outcomes varied as a function of prior knowledge. The Fischer z scores tell us whether the difference between the groups achieves canonical statistical significance.

For the Low but not the High knowledge group, errors had a significant and negative effect, whereas help seeking was associated with significantly more positive learning outcomes. Thus those who really do not know much about the domain fare better if they have a tendency to seek help and/or avoid error (help/(errors + help) in Table 2).

Table 2
Partial correlation coefficients between process measures and performance on the average probe test measure for both High and Low knowledge groups[a]

	Whole sample of 42	High group (n = 22)	Low group (n = 20)	Fischer z	Significance
Time per operation	0.36*	0.32	0.45		ns
n errors	−0.31*	0.04	−0.51*	1.85	$P < 0.06$
Depth of help	−0.00	−0.30	0.35	2.01	$P < 0.04$
Latency	0.41**	0.59**	0.24		ns
Help/(errors + help)	0.14	−0.22	0.52*	2.42	$P < 0.02$

[a] *$P < 0.05$; **$P < 0.01$;.

For both High and Low groups, a slower pace of working is positively (though not significantly) correlated with performance on the probe test measure, whereas help seeking latency is only significant for the High group, and again, the difference between the coefficients for the two groups did not achieve significance.

3.4. Consistency of help seeking activity

As the Pearson Product Moment coefficients shown in Table 3 demonstrate, learners who sought help most frequently on the initial, positive quadratic items were also significantly more likely to seek help frequently on both cubic and negative quadratic problems. To a lesser extent those making most errors on the positive quadratics also made most errors on the other items.

On the second day, pupils could do the quadratic tasks again, or skip to where they had left off. Those who had sought most help on the first day did so again on the second while the same was not true for the making of errors. These results indicate systematic individual differences in the tendency to seek help.

3.5. Help seeking and learning transfer

Given that a contingent tutoring system is designed to ensure that every learner receives sufficient support to complete all learning tasks, the incidence of successful solutions to the transfer problems in QUADRATIC does not provide a means of assessing learning. More informative is an index of the efficiency with which solutions are achieved. This takes account of the frequency of help requests and errors which precede successful task completion. In Table 4, the average percentage of successful operations to all operations (successes plus help requests plus errors) are given for positive quadratics, cubics and negative quadratics. For each category of problem, performance measures for the initial problem encountered and for the general case are shown.

An analysis of learners' solution efficiency (using the sign test) indicates that the sample found the initial cubic case harder to complete than the quadratic. For the general case, however, the cubic proved easier. For both cases of the QUADRATIC, the increased difficulty in going from specific to general approached statistical significance. For the cubic case, however, the general case was significantly easier.

Table 3
Correlations between frequency of help seeking and errors for quadratic, cubic and negative quadratic cases[a]

	n	Help seeking	Errors
Quadratics and cubics	38	0.58***	0.33*
Quadratics and negatives	19	0.48*	0.44
First and second days	42	0.51***	0.14

[a] *$P < 0.05$; ***$P < 0.001$.

Table 4
Efficiency of successful solutions on quadratic, cubic and negative quadratic cases, plus z scores for differences between them[a, b]

	$(x+1)^2$	$(x+n)^2$	$(x+1)^3$	$(x+n)^3$	$(x-1)^2$	$(x-n)^2$
% success	72	62	57	70	65	69
z: $(x+1)^2$		1.85*	2.18**			
z: $(x+n)^2$				2.09**		
z: $(x+1)^3$				2.83***		

[a] $^{*}P < 0.06$; $^{**}P < 0.05$; $^{***}P < 0.01$.
[b] Since less than 50% of the sample worked their way to and through the negative quadratic cases (mainly the more able learners), the performance indicated by these measures and those for the other cases are not comparable.

Since changes in these indexes of efficiency could come about through variation in the relative frequency of errors and/or help seeking, raw error and help frequencies for each learner on each problem were also compared.

For the quadratic, it was errors which increased as pupils struggled to cope with a general equation, involving two variables, for the first time (23 showing an increase and only 6 a decrease, $z = 2.96$, $P < 0.003$). When it came to the cubic cases, it was help seeking which declined significantly (18 seeking less help and 4 more, $z = 2.77$, $P < 0.006$).

These results provide evidence that transfer is occurring on the task of working out how the initial case relates to the general case.

3.6. Prior knowledge and contingent help seeking

The correlational analyses demonstrate that learners with less prior knowledge sought help more often. However, since these pupils also made more errors, it is necessary to explore the relations between error and help seeking in more detail to assess the extent to which the two groups might differ in help seeking activity. On average, the Low group made twice as many errors as the High one (16 against 8). The High group was significantly more likely to follow an error with a correct response, i.e. to self-correct (0.59 against 0.33; $P < 0.01$). Objectively, the Low group, making more errors, and being less likely to self-correct, should therefore need more help. However, the data indicate that the opposite was the case, with children in the high group being more likely to seek help after error (0.27 against 0.11, $P < 0.02$). Although this comparison excludes those learners in the High group who made no errors or who self-corrected all but the odd one (11 in total), it suggests that greater prior knowledge of the domain is associated with more effective help seeking after errors.

To rule out the possibility that such differential help seeking might result from less effective help provision by the tutor for the low scorers, the probability of success after help for each learner was also determined. This revealed no significant difference between the likelihood of success after help for the two groups (0.62 against 0.60).

3.7. Illustrative individual cases

The analysis of grouped data, such as that just presented, is necessary to assess whether or not claims about the impact of the design of the tutoring system hold up in the general case. However, the basic aim of our work on contingent tutoring systems is also to demonstrate how systems can be designed to adapt their support to fit the demands of individual learners. In this section, we illustrate how a profile of learner–system interaction process measures can be used to identify both 'typical' and potentially problematic cases (i.e. examples of individual performance which seem at variance with the results found for the general case). The aim of such case studies is to suggest ways in which contingency theory could be extended to deal with such problematic cases.

Miles' case is illustrative of a pupil whose learning outcome and profile of interaction with the tutor reflect the overall pattern of results for the High knowledge group. Although he made frequent errors, his performance was fast and he sought help relatively infrequently. When he did seek help, he did so some considerable time after his previous action on the system. On the positive and negative quadratic and cubic cases his success rate remained high (over 80%) and he showed learning gains at post-test commensurate with his initial performance at pre-test.

Hazel illustrates both the benefits and potential challenges of on-line tutoring and assessment. Hazel would have taxed a human tutor to the limit. She sought help continually, even on the simplest of tasks. Though in the Low knowledge group, she made only one slip. Thus, almost every help request came before and/or after a successful move. However, she did eventually abandon this seeming 'abuse' of help and went on to achieve one of the most marked learning gains. Although there may be circumstances in which such 'help abuse' signals an over reliance on the tutor and a superficial approach to learning, Hazel's learning outcomes suggest that she was learning extremely well, providing no evidence that the tutor should have been prepared to refuse her help. However, this assumption could be tested by evaluating the impact of withholding help by the system on her subsequent measures of performance and transfer.

Max, on the other hand, was one of the learners in the Low group who showed least progress on the outcome measures, despite seeking help quite often. He made many errors, including one error sequence 13 moves long. He was also one of only three pupils who repeatedly abandoned a problem. Either the help did not help or he really needed more than he was prepared to ask for. For Max, changes to the tutor's on-line help regime (triggered by his high error–error rate) could be tried. He could be given help automatically after error or given hints that help might prove useful. Either way, the impact of instruction on his seemingly poor self-regulatory skills could be evaluated on-line.

Peter, though initially amongst the highest scorers at pre-test, fell below the average score on learning outcomes for the whole group. In general, he worked relatively slowly with the tutor, and both his errors and requests for help were infrequent. On these measures his profile fits that associated with good outcomes for his (High) group. His latency of help seeking was relatively short, however, and may have been a factor attributing to his poor outcome. One pedagogical strategy with Peter (in addition to feeding the profile back to him and his teacher) would be to use the tutor to suggest that he takes more time before seeking help, giving himself space to consider why he might be in difficulty. Such tutorial tactics could then be evaluated by monitoring his ensuing performance.

4. Discussion

4.1. Help seeking, co-constructing the ZPD, and tutoring

One of the main motivations for this investigation was to explore the relations between prior knowledge, learner–system interactions and learning outcomes in computer-based tutoring. Overall, learners with less prior knowledge sought help from the tutor more frequently than their more knowledgeable peers. However, an analysis of individual differences in learner responses after they had made an error did not reflect this general association. In this specific context, learners with more domain knowledge not only made fewer errors, and were significantly more likely to self-correct their errors, they were also more likely to seek help. These findings suggest that the accuracy of learners' judgements about their need for help after an error (and/or their readiness to seek help) reflects prior knowledge.

Previous analyses, both with QUADRATIC and in other learning contexts (e.g. McKendree, 1990), have shown a negative effect of errors on learning. The results of the current analyses have shown that this negative effect is statistically mediated by the impact of errors on learning outcomes for a group of pupils with relatively little prior knowledge of the domain. At the same time, frequency of help seeking was associated with positive effects on outcomes for this group but a negative one for those with more prior knowledge. Thus, the *balance* struck between committing errors and seeking help proved to be a highly significant process indicator for the Low knowledge group but not for the High one. Conversely, whilst learning outcomes for the High knowledge group were correlated with individual differences in the amount of time each learner spent before seeking help (such that more time was associated with better outcomes), help seeking latency was not a significant correlate of learning outcomes for the Low group.

Further, our analyses have shown that the effects of (contingent) help on performance are similar for both high and low knowledge learners. Help sought and received increases the chances of successful actions for all learners up to a similar, high level. For the Low knowledge group, seeking help leads to a much greater increase in the chances of success over that gained from trying to recover alone from error. Thus, the positive impact of help on the task performance of the learner with relatively little domain knowledge may be to prevent them from getting deeper into confusion. If this interpretation is sound, it should follow that encouraging less knowledgeable learners who are 'help refusers' to seek help more strategically will reduce their confusion and improve their learning.

The positive relation between relatively long help seeking latencies and learning outcomes might, we suggest, reflect processes of self-explanation which are an index of effective learning strategies (Wood, 1999). If this speculation proves to be correct, then the current findings show that the phenomenon is significantly more marked in learners with relatively high levels of domain knowledge. This finding may indicate that, given more relevant domain knowledge to evaluate, those learners who take time to reflect upon what they know before deciding to seek help are exhibiting a more thoughtful approach to learning. Although speculative, these hypotheses about the interpretation of aptitude × process interactions are testable.

The general findings from QUADRATIC are consistent with the theoretical position taken by Winnykamen and others who hold that learner help-seeking should be viewed as an adaptive and instrumental process whereby the learner exploits knowledge and skill in collaborative knowledge acquisition to determine the nature of their own social learning environment. An analysis of help seeking identifies one role that the learner can play in constraining the activities of the tutor as they seek, collaboratively, to locate and work within their own region of sensitivity to instruction or ZPD. On the basis of the current evidence, it seems to be the case that some lower achievers, both in the present computer-based investigation and in face-to-face tutoring, are likely to impose weaker constraints on this process.

Although research in face-to-face tutoring motivated our analyses of aptitude × process interactions, the pattern of results just discussed have not, to the best of our knowledge, been found before. Given the very detailed level of performance analysis, and the accuracy of timing needed to access such interaction processes, this is not surprising. Undertaking such analyses in face-to-face rather than computer-based situations would be a daunting task. Further, our investigation of aptitude × process interactions relied on the fact that the rules driving the tutorial process were held invariant over different dyads. Ensuring this degree of control in face-to-face interactions would be nigh impossible. Given the absence of evidence against which to compare our current findings and their interpretation, we obviously cannot offer firm generalisations about the detailed relations between knowledge, help seeking and learning found here. However, the more general argument that we wish to push further on the basis of such findings does not rest or fall on the generality of such detailed results. This more general argument concerns the role of feedback in learning and the design of more contingent tutoring systems.

4.2. Feedback, learning and contingency

Simple correlations between measures of prior knowledge, indexes of the collaborative learning process and learning outcomes show that, in general, a faster rate of progress through the tutoring session, with fewer errors and requests for help, is predictive of good performance. However, although these process measures predict variance in learning outcomes it obviously does not follow that they identify the processes which mediate effective learning itself. Providing feedback which, implicitly or explicitly, encourages a learner to be fast, accurate and autonomous will not guarantee sound advice to the learner about how they should try to regulate their own learning. In some contexts, taking more time in order to plan, reflect or evaluate may lead to more effective problem solving. Asking for help at an opportune time can serve to avoid confusion and enhance learning.

In their discussions of integrated learning systems, Wood, Underwood and Avis (this volume) explore the hypothesis that experience with computer-based systems which give feedback about the speed and accuracy in performance may have unintended, negative consequences on learners' implicit theories about the nature of learning itself. To the best of our knowledge, the effects of richer feedback which draws the attention of learners and teachers to aspects of self-regulation other than speed, success and error, have not been evaluated. Indeed few studies have evaluated the real impact of feedback on computer-based learning (McKendree, 1990). Feedback to learners and their teachers about the nature and timing of help seeking and its effects provides one means of exploring the impact of a more enhanced representation of learning processes on assumptions about the nature of learning.

In the preceding sections where we discussed general results and case studies from the current investigation, we have suggested a number of heuristics which could be used to design and evaluate more contingent tutors. One way of testing hypotheses about individual differences in tutor-regulation by the learner would be to build and evaluate computer-based tutoring environments designed to help the learner to improve their tutor-regulation practices. We have identified (on-line) measures which can be used to determine when it might be appropriate for a tutor to try to help learners who use on-line help too much, or too infrequently, or too quickly, to try to change their approach. The analyses suggest specific hypotheses which could be evaluated by examining the impact of such advice from the tutor on the individual learner's subsequent help seeking activities, and on their learning outcomes. For example, where a learner exhibits both frequent error and little self-correction, and seeks help infrequently, it would be possible to explore the impact of (contingent) hints about how they might make more effective use of help. Where a learner shows little evidence of taking any more time before seeking help than they do before a successful action or an error, it would be feasible to explore the impact of suggestions about why they might want to think more about if and why they need help.

Analyses based on help seeking activity have also been used to illustrate how measures based on the use of help and its effects on performance can be exploited to investigate the transfer of experience across different tasks. An assessment of the amount of help required to meet new demands provides a finer 'grain size' of more discriminating measures than one based on a dichotomy between success and error. A learner who cannot succeed alone but who needs only the merest of hints in order to succeed is likely to be at a quite different knowledge level than one who drives the tutor to give very explicit hints and clues.

Although, in the current investigation, we cannot make strong claims about the nature of the impact of QUADRATIC tutoring on transfer (due to a lack of experimental control over the order in which tasks were set), the analyses demonstrate how help seeking can be used to assess transfer, and show that, in a contingent tutoring context, learner help seeking is sufficiently systematic to provide a reliable basis for inferences. As Wood, Underwood and Avis argue, it is necessary to evaluate the impact of tutoring on a number of different measures of learning outcomes to support inferences about the nature of what it is that learners have and have not learned. Current analyses with QUADRATIC are investigating other, off-line assessments of learning outcomes. These should support more confident inferences about the impact of contingent tutoring on the nature of learning. Finally, the current analyses have illustrated how the use of help seeking measures, taken on line, might also be exploited to provide a means of evaluating the impact of tutoring on individual learners, including effects on aspects of their self-regulation activities.

4.3. Help seeking and models of learning

Our investigation of QUADRATIC has demonstrated how analyses of help seeking and the effects of that help on performance can be used to complement measures of time, success and error in relating learning activity to psychometric assessments of prior knowledge (and of learning outcomes). The same information could be used to enhance on-line models of learner knowledge. Tutoring systems developed and evaluated by Luckin and her colleagues illustrate one way of doing this (Luckin, 1998; Luckin and du Boulay, 1999, forthcoming). Their tutors exploit Bayesian networks which take such information to inform the tutoring system's 'beliefs' about learner knowledge. Information from this network is then exploited to support both the system's domain planning (i.e. what learning tasks to provide next for an individual) and instructional planning (i.e. what help to provide for that individual).

Developed further, knowledge-based models of the learner could be used to explore and evaluate hypotheses put forward on the basis of the learner–system process measures used in QUADRATIC. On-line learner modelling which proves powerful enough to specify precisely what individual tutees know and are learning could be informed by, and serve to inform, the analysis and exploitation of information on help seeking. The approach to learner modelling advanced by Anderson and his colleagues (Anderson, 1987; Anderson, Boyle, Farrel & Reiser, 1987; Anderson, Corbett, Koedinger & Pelletier, 1995), for example, could be augmented to capitalise on information from help seeking and its effects in much the same way as Luckin and her colleagues have done. Such a model could then be compared with learner help seeking activity in 'real time' to assess whether a given learner was likely to be 'abusing' help (i.e. requesting information on problems where their chances of autonomous success were estimated to be high) or, conversely, 'refusing' it (i.e. failing to request information on problems where their chances of success were estimated as low). Such an agenda could bring together and integrate the investigation of learning processes and psychometric assessment, as done here, with knowledge-based approaches to learner modelling.

If such a rapprochement of different approaches were to succeed, it would also help to integrate at a theoretical level models of 'cognitive architectures' with a social-constructivist perspective on learning. We have already argued that the analysis of principles for supporting effective tutoring derived from these two traditions show marked convergence, and have suggested that Anderson's analysis of procedural learning as skill acquisition provides a plausible theoretical account of the nature and origins of what Vygotsky called 'inner speech' (Wood and Wood, 1996b). Further theoretical integration along the lines envisaged here could also shed new light on to the nature of self-regulation. The learner's use of help seeking to co-regulate the process of tutoring, and the impact of tutoring on this process itself, provides a potential way in to the study of Vygotsky's claims about the nature and origins of self-regulation and about the formation of an inner dialogue.

References

Anderson, J. R. (1987). Production systems, learning and tutoring. In D. Klahr, P. Langley, & R. Neches, *Production system models of learning and development*. Cambridge, Mass: The MIT Press.

Anderson, J. R., Boyle, C. F., Farrel, R., & Reiser, B. J. (1987). Cognitive principles in the design of computer tutors. In P. Morris, *Modelling cognition* (pp. 93–133). Oxford: John Wiley.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: lessons learned. *The Journal of the Learning Sciences*, *4*(2), 167–202.

Behr, M. J., Herel, G., Post, T., & Lesh, R. (1992). Rational number, ratio, and proportion. In D. A. Grouws, *Handbook of research on mathematics teaching and learning* (pp. 296–333). New York: Macmillan.

Brown, M., Hart, K., & Kuchmann, D. (1984). *Chelsea diagnostic mathematics tests: Algebra* Originally published by NFER–Nelson, now available from Prof. K. Hart, Shell Centre for Mathematical Education, University of Nottingham.

Bruner, J. S., & Kenney, H. J. (1965). Representation and mathematics learning. *Monographs of the Society for Research in Child Development*, *30*(1), 50–59.

Dienes, Z. P. (1960). *Building up Mathematics*. London: Hutchinson.

Dienes, Z. P., & Jeeves, M. A. (1970). *The effects of structural relations of transfer*. London: Routledge and Kegan Paul.

Kaput, J. J. (1992). Linking representations in the symbol systems of algebra. In D. A. Grouws, *Handbook of research on mathematics teaching and learning* (pp. 167–294). New York: Macmillan.

Lesh, R., Landau, M., & Hamilton, E. (1983). Conceptual models in applied mathematical problem solving. In R. Lesh, & M. Landau, *Acquisition of mathematics concepts and processes* (pp. 263–343). New York: Academic Press.

Luckin, R. (1998). Knowledge construction in the zone of collaboration: scaffolding the learner to productive interactivity. In A. Bruckman, M. Guzdial, J. Kolodner, & A. Ram, *Proceedings of the International Conference of the Learning Sciences* (pp. 188–194). Charlottesville, VA, USA: AACE.

Luckin R., du Boulay B. (1999). Capability, potential and collaborative assistance. To appear in *Proceedings of the Seventh International Conference on User Modelling*. UM 99, Banff, Canada.

Luckin R., du Boulay B. (forthcoming). Designing a zone of proximal development. To appear in *International Journal of Artificial Intelligence and Education*, Vol. 10.

McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human Computer Interaction*, *5*, 381–413.

Nelson-Le, Gall S. (1985). Help-seeking behavior in learning. *Review of Research in Education*, *12*, 55–90.

Nelson-Le Gall, S., Kratzer, L., Jones, E., & DeCooke, P. (1990). Children's self-assessment of performance and task-related help seeking. *Journal of Experimental Child Psychology*, *49*, 245–263.

Puustinen, M. (1998). Help-seeking behavior in a problem-solving situation: Development of self-regulation. *European Journal of Psychology of Education*, *XIII*(2), 271–282.

Vergnaud, G. (1982). A classification of cognitive tasks and operations of thought involved in addition and subtraction problems. In T. P. Carpenter, J. M. Moser, & T. A. Romberg, *Addition and subtraction: A cognitive perspective*. Hillsdale, NJ: Lawrence Erlbaum.

Winnykamen F. (1992). *Les interactions de guidage: Le mediation par la tutorat*, Bulletin de L'Association Francaise des Psychologues Scolaires. Actes du congres de Lille (18 Octobre, 1991) .

Winnykamen, F. (1993). Gestion socio-cognitive du recours a laide d'autrui chez l'enfant. *Journal International de Psychologie*, *28*, 645–659.

Wood, D. (1999). *Theories of learning and computer-based teaching*. Paper to the Seventh International Conference of the International Association for Cognitive Education. Calgary, Canada, July 1999.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100.

Wood, D., Shadbolt, N., Reichgelt, H., Wood, H., & Paskiewicz, T. (1992). EXPLAIN: Experiments in Planning and Instruction. *AISB Quarterly*, *81*, 13–16.

Wood, D., & Wood, H. (1996a). Commentary: Contingency in tutoring and learning. *Learning and Instruction*, *6*(4), 391–397.

Wood, D., & Wood, H. (1996b). Vygotsky, tutoring and learning. *Oxford Review of Education*, *22*(1), 5–16.

Wood, D. J., Wood, H. A., & Middleton, D. J. (1978). An experimental evaluation of four face to face teaching strategies. *International Journal of Behavioral Development*, *1*, 131–147.

Wood H. (1996). Contingent tutoring tutor: A computer based tutorial on contingent tutoring. Hypercard interactive demo. Nottingham, UK, Centre for Research in Development, Instruction and Training.