Effects of representational modality and thinking style on learning to solve reasoning problems

Padraic Monaghan (pmon@cogsci.ed.ac.uk)
Centre for Cognitive Science, University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, UK

Keith Stenning (keith@cogsci.ed.ac.uk)
Human Communication Research Centre, University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, UK

Abstract

Individual differences in the abilities and preferences of students have an influence on their responses to information presented in alternative ways. Explanations may appeal to differences in representation or in strategy. This paper reports an experiment that compares the response of students to two computationally similar methods of teaching syllogisms that rely on different external representations of the premiss information. The use of both representations can be broken down into the same stages: translating-in; manipulating; and translating-out. We show that the ease of acquisition and the understanding of the methods relate to a measure of spatial ability and also to preferences for serialist/holist styles of learning. We find that spatial ability and learning style relate to different stages in the two teaching methods, and are therefore complementary contributors to effective learning. In addition, a further test that predicts diverse responses of students to learning the same information from different modalities was used. This is found to relate specifically to stages of translating-in and manipulation of representations. The results of this study support the view that providing a computational account of reasoning and learning requires an acknowledgement of individual differences in the 'starting state' of the individual. These differences can be explored through measures of ability and learning style. This study also supports accounts of problem-solving that distinguish modality and strategy of information processing.

Introduction: Processing modality and style

From the educational perspective, knowing what styles of information presentation are effective for learning is a practical imperative. But to characterise the cognitive processes that are invoked by different presentations of information is also an important theoretical aim. The individual differences which are of such practical importance to education may provide a tool for analysis of mental processes. Assessing the influence of modality of information on learning, and providing computational descriptions of the processes, invariably leads to the well-publicised problems associated with studying internal representations (Pylyshyn, 1973). Paivio (1986) has argued that the style of processing associated with a strategy rather than the representation modality is the more fundamental issue in providing computational accounts of information processing. The modality and style perspectives have remained independent accounts of performance in complex tasks.

One approach to studying individual differences in complex problem solving has concentrated on issues of modality. According to Kirby, Moore and Schofield (1988), some students prefer to process information visually through graphics, diagrams or illustrations, whereas others prefer to process information verbally. Preferences for visuo-spatial representations of problems have been linked to high spatial ability in the sentence-picture verification task (MacLeod, Hunt & Mathews, 1978). Individuals also vary in their ability to process information presented in the verbal or the spatial modalities, indicated by factor-analytic studies of the independence of verbal and spatial ability (Guilford, 1967; Poltrock & Brown, 1984). That verbal and spatial processing resources are, at least partially, distinct is supported by theories of working memory that posit different components for phonological and visuo-spatial storage and processing (Baddeley & Gathercole, 1993; Shah & Miyake, 1996).

An alternative approach has focussed on individual differences in strategy use during reasoning tasks. Research into students' responses to Hyperproof (HP), a multimodal logic tool which enables proofs to be constructed as hybrids of graphical situations and first order logic expressions of those situations, allows detailed analysis of strategies of using external representations. Cox, Stenning and Oberlander (1994) compared students' performance on a GRE Analytic Reasoning test before and after following either the HP course, or a version of the course that used HP with the graphical interface disabled. Although the relevant GRE questions are verbally set and answered, they are usefully solved by constructing a diagram to support reasoning, as the premisses constrain a unique, or nearly unique, logical model. They found that on one measure of near-transfer, those that scored high on the GRE improved their reasoning performance by following the HP course, but performed worse if they followed the non-graphical HP course.

These differences in response to diagrammatic presentations can be better accounted for by students' strategic skill in knowing *when* to translate from sentences into diagrams, and when in the reverse direction, than by preferences for one modality or another (Oberlander et al., 1996). In fact their 'verbal' students who failed to gain from diagrammatic presentation did so because they translated into the diagrammatic mode *more* often at inappropriate junctures.

Monaghan (1998) shows that different proof styles exhibited in the HP course can be related to the individual differences literature through the holist-serialist dimension (Pask, 1988). Holists tend to use strategies that formulate connections between all aspects of the situation at once. Serialists tend to use strategies where the situation is broken down into units of information, and connections are formed between these units. In HP, holists use the graphical interface to gradually build up concrete situations, using nested assumptions.

Serialistic proofs use a series of fully concrete situations, where little or no nesting of assumptions is used. High GRE scorers use holistic strategies, low GRE scorers construct serialistic proofs.

The present study set out to explore these aptitude treatment interactions in the teaching of reasoning in a radically different domain in which our understanding of the equivalences and differences between graphical and sentential reasoning methods have recently been clarified. Stenning and Yule (1998) show that for categorial syllogisms, Euler's graphical method can be emulated operation-for-operation by a simple natural deduction method in a fragment of propositional calculus. This enables components of reasoning strategies to be equated across modalities of external representation.

This simpler domain allows the separation of translation-in; manipulation; and conclusion-forming stages, each of which can be identified in both diagrammatic and sentential presentations. Since teaching this domain can be accomplished in about half an hour, (as opposed to the 10 weeks of the HP course), it is possible to study the details of the tutoring process and its impact on immediately successive performance. As the teaching was videoed, the material provides detailed new data on the effects of teaching with different external representations. We also sought to explore the relations between established measures of 'strategy' use: the GRE Analytical Reasoning test, and the serialist-holist test; and a test of processing ability within the spatial modality (the paper folding test (PFT)), and to see whether these different tests might be sensitive to performance on different components of the reasoning process, or distinguish between representational and strategic differences. To the extent which similar individual differences are shown to distinguish learning in this domain and the HP domain, the study provides evidence of the generality of these aptitude treatment interactions in learning to reason.

Two methods of solving syllogisms

Syllogisms are reasoning problems about three properties, A, B, and C. The A property is related to the B property by the first premiss, the B and the C properties are linked by the second premiss. The subject's task is to say what, if anything, follows from the given information.

Stenning and Yule (1998) describe two methods for solving syllogisms that depend on different representations of the premisses. These two methods are 'computationally similar', in that for each stage in one method there is a reciprocal stage in one-to-one correspondence in the other method. One method relies on graphical representations based on Euler's Circles (EC), the other method is based on sentential forms of the premisses, derived from a fragment of natural deduction (ND). Figure 1 shows the EC representations of the four kinds of syllogistic premisses, Table 1 shows the ND translations of the premisses.

The EC method relies on a few simple rules for putting the separate premiss diagrams together, adding crosses to indicate existence in areas of the resulting diagram, and reading off a conclusion. The ND method utilises a fragment of propositional calculus in order to reach a conclusion. To form a conclusion either *modus ponens* or *modus tollens* is used.

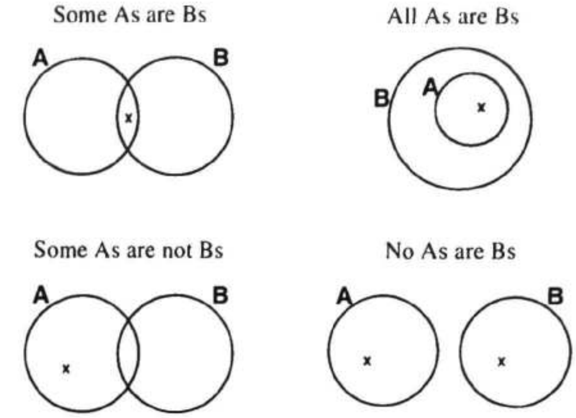

Figure 1: EC representations of syllogism premisses.

Table 1: ND representations of syllogism premisses

premiss	translation
All A are B	$A \rightarrow B$
Some A are B	A & B
No A are B	$A \rightarrow \neg B$
Some A are not B	A & $\neg$B

Combining the EC premisses into an integrated representation corresponds with use of one of these two rules. Correspondences between these two methods emerge through consideration of the property of *case-identifiability*: each method isolates critical types of individuals and then makes conclusions about them with respect to the A and C properties. In the ND method, this is done serially, with each premiss being treated separately. In the EC method, all the premiss information has to be integrated, and all individuals treated simultaneously. Three stages are easily isolated and comparable between the methods: (1) translation from the premisses into the represented form; (2) manipulation of the represented forms in order to isolate the critical individuals; and (3) translation from the final representation to form a conclusion. The correspondence of these stages in the two methods mean that different key stages in the use of external representations can be assessed.

Hypotheses of the study

If the GRE test is a general measure of aptitude in using graphical representations, then it is predicted that high GRE scorers would better learn the EC method over all stages of its use. In accordance with the HP data, it was predicted that high GRE scorers would perform worse on the ND method. The PFT test was predicted to relate to the stage of manipulating the graphical representations, as the test involves only manipulations of spatial representations, therefore translating between modalities is unimportant. The serialist/holist preference was predicted to relate to all stages of the methods, holists learning better from the EC method as graphical representations promote holistic processing, and serialists being better fitted to the ND method which supports serial strategies.

Method

17 first year undergraduate students from a wide variety of academic backgrounds took part in the experiment. Each student was paid for participating in the experiment.

Pre-tests

Ability to process in the spatial modality was assessed by the paper folding test (PFT) (French, Ekstrom & Price, 1963). This requires the subject to imagine the array of holes resulting from a piece of paper being folded, then having holes punched in it before being unfolded again. High and low spatial ability students were discriminated along a median split.

Processing style was assessed by using Ford's (1985) measure of serialist/holist learning style. This distinction has been shown to discriminate different strategy uses and responses in a number of domains, including human-computer interaction (Helander 1990).

The GRE test was constructed from the 'analytical' subscale of the analytical reasoning test. High and low GRE scoring students were again discriminated by a median split.

The syllogistic reasoning task

All students were given 8 syllogisms to solve, including five with valid conclusions and three with no valid conclusion. These syllogisms were selected to cover a range of difficulties according to the findings of Johnson-Laird (1983). In addition, the syllogisms were selected to include some items that 'spatial' reasoners found easier to solve than did 'verbal' reasoners, and vice versa, according to Ford's (1995) criteria and data. While solving the problems students were requested to voice their thought processes. The teaching sessions were videoed and transcribed. Students were paired according to scoring similarly on the pre-tests, and then each assigned to a different teaching group, where they had to solve the same eight syllogisms using either the EC or the ND method. Nine students were instructed in the EC method, and eight followed the ND method. All students were taught for the same length of time (the session ended after one hour), and not all students finished all the problems.

The video protocols were marked for errors that the student made in applying the method to solve the syllogisms, and for the number of corrective or directive interventions that the instructor had to make during the course of solving the problems. A pilot experiment indicated that the number of syllogisms solved by the student was related to general pace of working rather than to the ease or accuracy of acquisition of the method. Number of errors measures the ability of the student to utilise the method correctly whereas interventions are more an indication of the student's understanding of the application of the method to the problem.

Results

The PFT and GRE scores were not significantly correlated ($r(15) = 0.32$, $p>0.2$). Holists scored higher than serialists on the PFT ($t(14)$[1] $= 3.79$, $p<0.005$), and holists also scored higher on the GRE test ($t(14) = 2.47$, $p<0.05$). These results support general hypotheses that spatial information is conducive to holistic processing, and that the GRE relates to ability to use spatial representations to support reasoning processes.

The data were analysed with two-way unrelated ANOVAs with the number of errors made in applying the taught method on each stage, and the number of interventions made by the instructor in the course of the stage as the respective dependent variables. The two independent variables were the taught strategy and one of the three pre-test measures (serialist/holist learning preference; high-low PFT score; or high-low GRE score). Only significant interactions are reported in the results.

Translating-in stage

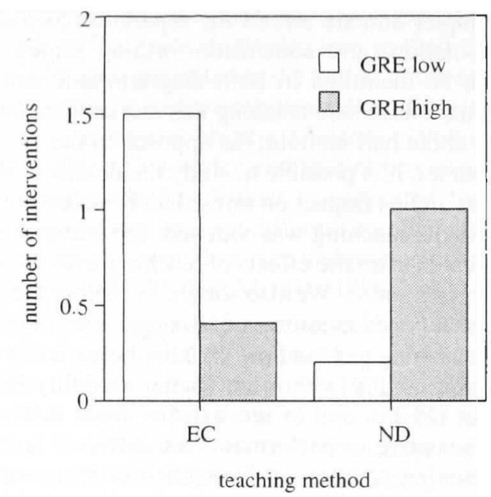

Figure 2: number of interventions for translation-in stage by GRE score for the teaching methods:$F(1, 13) = 5.30$, $p<0.05$. *N.b.*, different x-axis scale.

On the translation-in stage, there were few errors made or interventions required, as this stage in the procedures can be accomplished by reading off the translations. One interaction was, however, significant for this stage: GRE group by taught method on the number of interventions required (Figure 2): $F(1,13) = 5.30$, $p<0.05$, with no main effects. Those who score high on the GRE test require more interventions for this stage in the ND method, but fewer interventions for translating into the EC representations.

Manipulation of representations stage

For the manipulation of representations stage, both GRE score and serialist/holist style interacted with taught method for both dependent variables.

GRE and taught method interacted both for number of errors (Figure 3) ($F(1, 13) = 7.45$, $p<0.02$, with no main effects) and interventions (Figure 4) ($F(1, 13) = 5.50$, $p<0.05$, with a main effect only of taught method ($F(1, 13) = 11.84$, $p<0.01$)). High GRE scorers made fewer errors and required fewer interventions on the EC method, but made more errors and required more interventions on the ND method.

For the serialist/holist variable, number of errors by taught method on this stage were significant (Figure 5): $F(1, 12) = 5.89$, $p<0.05$, with a main effect only of learning style (F(1,

[1] One student could not be classified as serialist or holist as she demonstrated no preference for either learning style, so was omitted from these analyses.

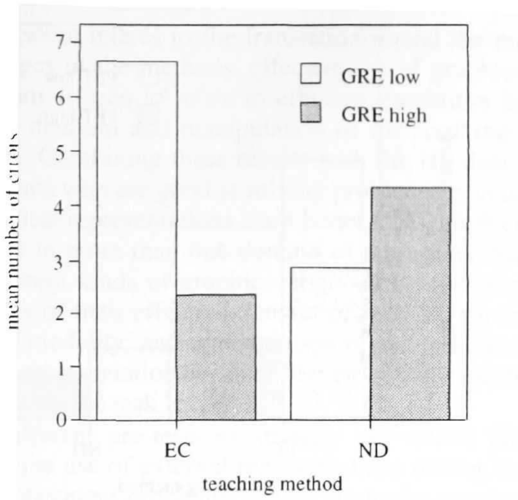

Figure 3: number of errors for manipulation stage by GRE score for the teaching methods: $F(1, 13) = 7.45, p<0.02$.

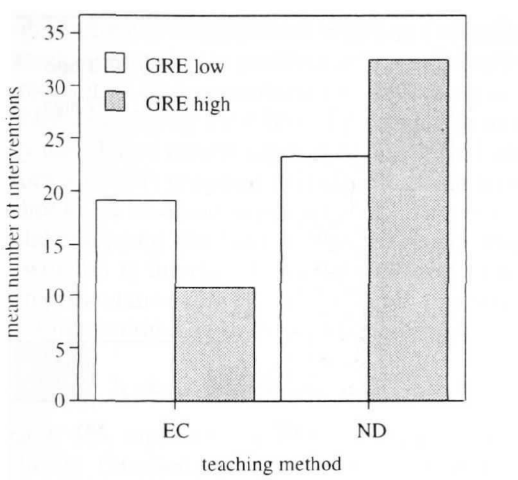

Figure 4: number of interventions for manipulation stage by GRE score for the teaching methods: $F(1, 13) = 5.50, p<0.05$.

12) = 7.12, $p<0.05$). For number of interventions (Figure 6), $F(1, 12) = 5.46$, $p<0.05$, with a main effect only of taught method ($F(1, 12) = 15.05$, $p<0.01$). Serialists made more errors and required more interventions than holists on the EC method. The serialist/holist distinction does not seem to differentiate performance on the ND method.

Translating-out stage

For the translating-out stages, both serialist/holist style and PFT score interacted with taught method.

For the serialist/holist learning style, number of errors were significant (Figure 7): $F(1, 12) = 5.23$, $p<0.05$, with a main effect only of learning style ($F(1, 12) = 5.23$, $p<0.05$). Also for number of interventions (Figure 8), $F(1, 12) = 9.82$, $p<0.01$, with a main effect again only of learning style ($F(1, 12) = 5.89$, $p<0.05$). As with the manipulation stage, serialists made more errors and required more interventions on the EC method than did the holists. The ND method was not distinguished by this pre-test grouping.

For PFT score there was an interaction with taught method

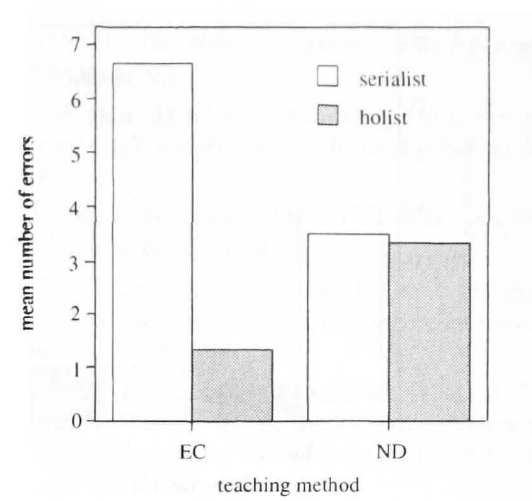

Figure 5: number of errors for manipulation stage by serialist/holist preference for the teaching methods: $F(1, 12) = 5.89, p<0.05$.

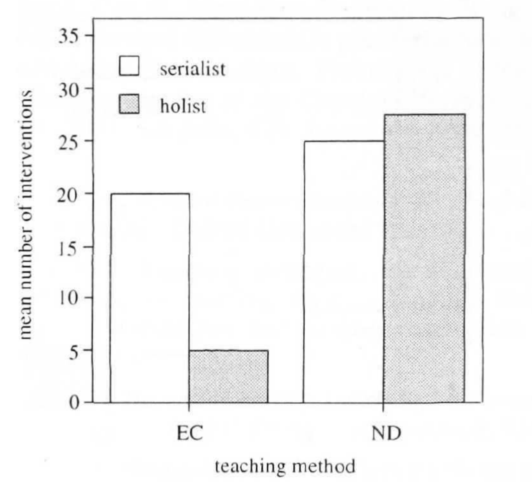

Figure 6: number of interventions for manipulation stage by serialist/holist preference for the teaching methods: $F(1, 12) = 5.46, p<0.05$.

both for number of errors (Figure 9): $F(1, 13) = 4.90$, $p<0.05$, with a main effect of taught method ($F(1, 13) = 9.96$, $p<0.01$) and PFT group ($F(1, 13) = 9.31$, $p<0.01$)) and number of interventions (Figure 10): ($F(1, 13) = 7.16$, $p<0.02$, with a main effect only of PFT group ($F(1, 13) = 6.66$, $p<0.05$)). High spatial ability students require fewer interventions on the EC method, and low spatial ability students make more errors on the EC method for this stage.

Discussion

The absence of main effects of teaching method for most of the ANOVA results provide important evidence that our implementations of these two teaching methods are comparably good.

As anticipated, spatial ability (PFT score) influenced performance on the EC method. However, it is perhaps surprising that PFT score only influenced the conclusion forming stage. This suggests that the PFT measures a student's

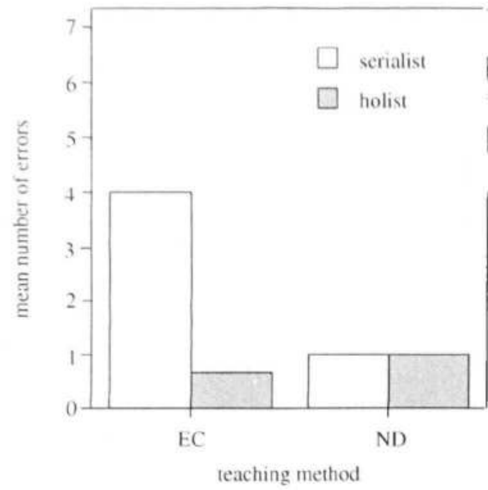

Figure 7: number of errors for translating-out stage by serialist/holist preference for the teaching methods: $F(1,12) = 5.23$, $p<0.05$

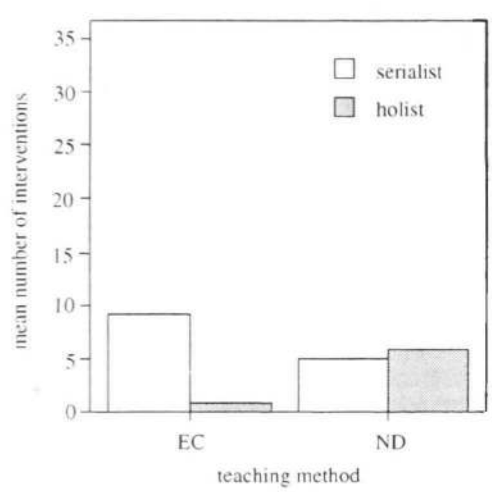

Figure 8: number of interventions for translating-out stage by serialist/holist preference for the teaching methods: $F(1, 12) = 9.82$, $p<0.05$.

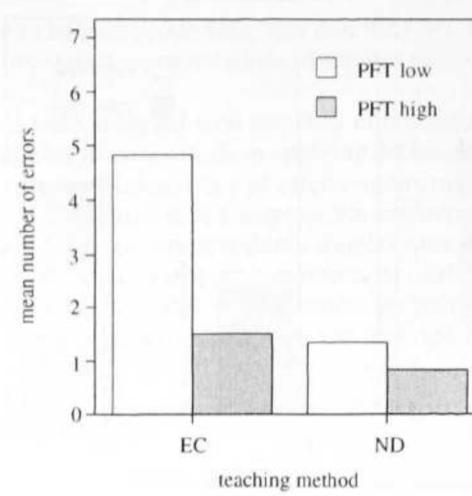

Figure 9: number of errors for conclusion stage by PFT score for the teaching methods: $F(1, 13) = 4.90$, $p<0.05$.

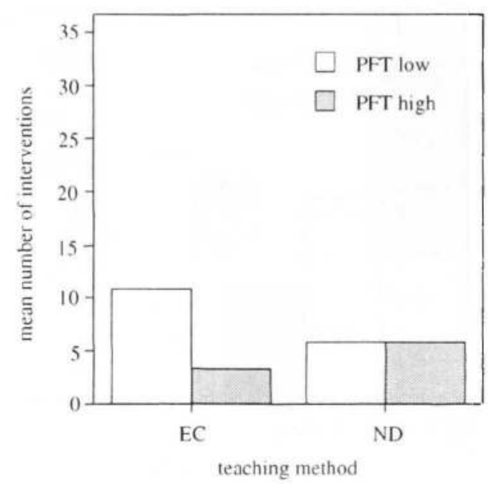

Figure 10: number of interventions for conclusion stage by PFT score for the different teaching methods: $F(1, 13) = 7.16$, $p<0.02$.

ability to translate accurately and with little assistance from the graphical modality into sentential form. PFT ability is therefore most influential at the point where graphical representations have to be interpreted into language. This rather suggests that far from measuring 'purely spatial operations' the PFT measures comparisons between verbal 'folding narratives' and spatial patterns (see Snow, 1978), and is open to strategic interpretations as opposed to representational ones.

The serialist/holist style distinction predicts different responses to both the manipulation stage and the translation-out stage. Students with a preference for holistic strategies make fewest errors and require fewest teaching interventions on the manipulation stage of the EC method, and most errors/interventions for the ND method. This result is consonant with expectations about the representational style of the EC method being conducive to holistic strategies in that the premiss information has to be integrated into one diagram. For the translation-out stage, serialists make more errors when forming conclusions from the EC method, perhaps due to the simultaneous consideration that has to be given to all the properties in order to describe the critical individual, a task that these students find difficult. The holists require fewest interventions for the EC method on this stage for similar reasons: the processing of all the information concurrently is advantageous to these students. When this style of information is not available (as in the ND method), neither group is advantaged or disadvantaged.

The effect of GRE score on teaching different methods of solving syllogisms shows that the results of the HP research apply to other domains of reasoning, and to much briefer teaching interventions. High GRE scoring students acquire the EC method more easily and with fewer errors, but perform worse with the ND method. Low GRE scoring students indicate the opposite tendency: acquiring the ND method with greater ease, but finding the EC method more difficult. The GRE can be taken as a predictor of the effectiveness of graphical representations for learning to solve reasoning problems.

As GRE score relates to the translation-in and the manipulation stages of the methods, effective use of graphical representations is seen to relate to efficient translation into the graphical medium and manipulation of the resulting representation. Combining these results with the HP data shows that students who are good at solving problems by constructing graphical representations learn better from graphical presentations in more than one domain of reasoning, and with very different kinds of graphic. High GRE scorers' ability is in terms of their effective transfer of information into the graphical modality, and manipulation of the representation: when the representation does not fit their aptitudes then these students learn the task less well.

The different pre-tests of strategy use reflect different stages in the use of external representations during learning to solve reasoning problems. These strategies are shown to interact with representational modality, and can be seen as testing different subprocesses of the reasoning task. The research on HP meshes with the current study to clarify the ways in which modality of presentation has an influence on learning: this influence is manifested in terms of effecting different strategic approaches to problem solving. The effects of different modalities of representation on reasoning are supplemented by considering the effect of presentation modality on strategy use. These results support Alesandrini, Langstaff and Wittrock's (1984) proposal that effects of modality and strategy should be assessed separately. If there is a single fundamental reasoning mechanism then the representational forms by which it is interfaced and the strategies of reasoning these representations invoke are of equal importance in providing computational descriptions of reasoning.

Acknowledgements

This research was supported by ESRC research studentship R00429634206. Grateful thanks to Jon Oberlander for helpful comments during analysis.

References

Alesandrini, K. L., Langstaff, J. J., & Wittrock, M. C. (1984). Visual-verbal and analytic-holistic strategies, abilities and styles. *Journal of Educational Research, 77*, 151–157.

Baddeley, A. D., & Gathercole, S. E. (1993). *Working Memory and Language.* Hove: Lawrence Erlbaum Associates.

Cox, R., Stenning, K., & Oberlander, J. (1994). Graphical effects in learning logic: reasoning, representation and individual differences. *Proceedings of the Sixteenth Annual Conference of the Cognitive Science Society* (pp. 237–242). Georgia: Lawrence Erlbaum Associates.

Ford, M. (1995). Two modes of mental representation and problem solution in syllogistic reasoning. *Cognition, 54*, 1–71.

Ford, N. (1985). Learning styles and strategies of postgraduate students. *British Journal of Educational Technology, 16*, 65–79.

French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Kit of reference tests for cognitive factors.* Princeton, NJ: Educational Testing Service.

Guilford (1967). *The Nature of Human Intelligence.* New York: McGraw Hill.

Helander, M. (Ed.) (1990). *Handbook of Human Computer Interaction* (2nd edition). North Holland: Elsevier Science Publishers.

Johnson-Laird, P. N. (1983) *Mental Models.* Cambridge: Cambridge University Press.

Kirby, J. R., Moore, P. J., & Schofield, N. J. (1988). Verbal and visual learning styles. *Contemporary educational psychology, 13*, 169–184.

Macleod, C. M., Hunt, E. B., & Mathews, N. N. (1978). Individual differences in the verification of sentence-picture relationships. *Journal of Verbal Learning and Verbal Behavior, 17*, 493–507.

Monaghan, P. (1998). *Holist and serialist strategies in complex reasoning tasks: cognitive style and strategy change.* (Research Paper EUCCS/RP-73). Edinburgh: University of Edinburgh, Centre for Cognitive Science.

Oberlander, J., Cox, R., Monaghan, P., Stenning, K., & Tobin, R. [1996] Individual differences in proof structures following multimodal logic teaching. *Proceedings of the Eighteenth Annual Meeting of the Cognitive Science Society*, (pp. 201–206). La Jolla, CA: Lawrence Erlbaum Associates.

Paivio, A. (1986). *Mental Representations: A Dual Coding Approach.* Oxford: Oxford University Press.

Pask, G. (1988). Learning strategies, teaching strategies, and conceptual or learning style. In R.R. Schmeck (Ed.), *Learning strategies and learning styles.* New York: Plenum Press.

Poltrock, S. E., & Brown. P. (1984). Individual differences in visual imagery and spatial ability. *Intelligence, 8*, 93–138.

Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin, 80*, 1–24.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: an individual differences approach. *Journal of Experimental Psychology: General, 125*, 4–27.

Snow, R. E. (1978). Eye fixation and strategy analyses of individual differences in cognitive aptitudes. In A. M. Lesgold, J. W. Pellegrino, S. D. Fokkeman, & R. Glaser (Eds.), *Cognitive psychology and instruction.* New York: Plenum Press.

Stenning, K., & Yule, P. (1998). Image and language in human reasoning: a syllogistic illustration. *Cognitive Psychology, in press.*

Sternberg, R. J., & Weil, E. M. (1980). An aptitude × strategy interaction in linear syllogistic reasoning. *Journal of Educational Psychology, 72*, 226–239.