The Interaction between Informational and Computational Properties of External Representations on Problem-Solving and Learning

Shaaron Ainsworth (Shaaron.Ainsworth@nottingham.ac.uk)
School of Psychology, University of Nottingham,
Nottingham, NG7 2RD, UK

Gareth Peevers (gjp@psychology.nottinghm.ac.uk)
School of Psychology, University of Nottingham,
Nottingham, NG7 2RD, UK

Abstract

This paper reports an experiment that explores if the way instructions for operating a complex device are represented influences problem-solving and learning about the task. Instructions were presented in one complex representation or in multiple simpler ones. The form this information took was tabular, diagrammatic or textual. Participants found the optimal solution more often when given instructions in text representations or in a single representation. However, the single text representation was associated with significantly slower performance. Participants recalled more about the task with text representations, irrespective of how the information was distributed. This experiment confirmed that representations that display instructions in such a way as to increase the cost of operating with them can paradoxically lead to better performance.

Introduction

There are many claims for the benefits that multiple external representations (MERs) can bring to learning and problem-solving (*e.g.* Cheng, Lowe & Scaife, 2001; Kozma, Chin, Russell & Marx, 2000). However empirical support for the benefits to learners is mixed with some studies reporting benefits of either constructing or using MERs in educational environments (*e.g.* Cox & Brna, 1995; Mayer & Sims, 1994; Tabachneck, Koedinger & Nathan, 1994) and some not (*e.g.* Ainsworth, Bibby & Wood, 2002; Tabachneck-Schijf & Simon, 1998).

One key factor that may underlie any benefits of MERs is that of computational non-equivalence. Representations that theoretically contain the same information (informational equivalence) differ in the ease with which people can extract this information. Larkin & Simon (1987) contrasted interpretation of graphical and textual ERs in terms of search, recognition and inference. They showed how search processes are considerably more efficient in diagrams. They propose that text has a high cost of perceptual enhancement when compared to the diagrams. Tables tend to make information explicit, emphasize empty cells that directs attention to unexplored alternatives, and allow quicker and more accurate read-off (*e.g.* Cox & Brna, 1995; Meyer, Shinar & Leiser, 1997). These findings help explain the match-mismatch hypothesis (Gilmore & Green, 1984) that argues that no notation is perfect; it is only good with respect to a particular task. Bibby & Payne (1993) showed that different ERs that explained the operation of a complex device facilitated alternative aspects of the task. Identifying a faulty component was helped if instructions were presented as procedures whereas a topological diagram improved recognition of a misaligned switch. Consequently, MERs may offer benefits for solving complex problems by allowing for these alternative perspectives.

However, the majority of the research on MERs has addressed only the computational properties of ERs without also considering how information is distributed. The combinations of ERs studied are almost invariably informationally equivalent. Yet, MERs are commonly used to distribute information over a number of separate ERs, so that each one displays a subset of the total information. This is often done when one ER would be very complicated if it expressed all the necessary information (Ainsworth, 1999). It is likely that there may be a match-mismatch effect for informational as well as computational properties of ERs.

One possibility is that separate ERs that isolate dimensions of information will allow learners to focus on separate aspects of the tasks. Each aspect could be learnt separately and then integrated with others when it was fully understood. Ainsworth, Bibby & Wood (1997) found that two ERs that each displayed a single dimension of information (no informational redundancy) allowed children to learn to perform an estimation task faster than two ERs that simultaneously displayed two dimensions of information (full redundancy). This was true when the information was represented in both mathematical expressions and pictures. In this case, it would appear that multiple simple ERs allowed learners to incrementally build their knowledge of the task, so facilitating performance.

However, a single ER that contains all the necessary information may facilitate problem-solving in other circumstances, as it: a) minimizes the number of ERs to be learnt, which can be demanding if these ERs have complex format and operators; b) often makes explicit the relations between dimensions of information and; c) presents this information in close proximity, which has been shown to reduce a split-attention effect (*e.g.* Chandler & Sweller, 1992). Research within the cognitive load paradigm would suggest that a single integrated ER (normally text and graphics) improves learning by reducing the working memory demands associated with integrating information (*e.g.* Kalyuga, Chandler & Sweller, 1999).

One further prediction that follows from the match-mismatch hypothesis is that ERs, which enhance the efficiency of problem-solving, may not be the same ones that facilitate learning. Svendsen (1991) showed that a command-line interface to the Tower of Hanoi task led problem-solvers to spend more time per trial, but they engaged in more planning and subsequently were able to remember more about the task. O'Hara & Payne, (1998) similarly found that increasing cost of operations in the 8-puzzle led to increased planning. Representations that increase the users' cognitive load may require learners to remember, evaluate, and design better solutions to problems.

This experiment explores whether the informational and computational properties of external representations interact to influence problem-solving and learning. The ERs provided instructions for operating a complex device – an Alchemist's Factory. The task required chemicals to bought (or made), combined in specific ratios, placed in the appropriate apparatus (which must be paid for) and then set to correct values. Furthermore, participants were requested to create gold in the cheapest possible way, as there are multiple correct solutions for this task that differ in their cost and complexity. Accordingly, the task requires information to be integrated (for planning solutions) but then efficient operation of the factory requires focusing on single factors. Instructions were presented either in one complete ER or in MERs, which each presents only a subset of the necessary information. They were displayed in text, diagrams or tabular forms.

Hypotheses

1) *Does the form of representation influence efficiency and effectiveness of problem-solving*? Given the prior research on the computational properties of ERs, it is proposed that the diagrams should enhance search and recognition compared to text. Hence, participants given diagrams should solve the problem faster. The structuring information provided by tables should also make them more efficient than text.

2) *Does the complexity of representation influence efficiency and effectiveness of problem-solving?* We hypothesized that finding the best solution to the task would be facilitated if all the information was available at once as participants would be able to more directly compare the different solutions paths. However, there may be no difference in efficiency, as although planning may be more effective with one ER, performing the task should be facilitated by separating information as this will reduce cost of navigation and may enhance memorization.

3) *Does the form and complexity of representation interact to influence efficiency and effectiveness of problem-solving?* We hypothesized that text, given its high cost of search and recognition, will take longer to study when presented as a single complex ER. However, diagrams and tables should be more efficient when presenting complete information as the additional costs of navigating more complex ERs will be minimized given the way diagrams and tables facilitate search and recognition. However, if the research on adding "interface costs" to encourage planning also applies to these ERs, then it possible that there will be reverse effects of the effectiveness of problem-solving: *i.e.* better solutions will be found when ERs are costly.

4) *Does the form and complexity of representations interact to influence learning?* Two alternative hypotheses are proposed. Learning may be facilitated when working memory demands associated with search and recognition are lower as this will promote transfer to long term memory. This suggests that diagrams and tables will enhance learning (Cognitive Load account). Alternatively, learning may be facilitated by ERs that are associated with higher costs of search and recognition as participants will be encouraged to remember this information. In this case, the complexity of the single text ER may lead to enhanced learning (Interface Costs account).

Method

Design

A two factor between groups design was used. The first factor (*complexity*) had two levels: *single*, which presented instructions in one complex ER and *multiple*, which presented the same instructions but in four simpler ERs. The second factor (*form*) had three levels: *table*, *diagram* and *text.*

Participants

Eighty-four students took part in the experiment for payment (£4). There were 56 females and 28 males with an average age of 21.5 years.

Materials

Participants were requested to create gold using an Alchemist's Factory in the cheapest possible way. A typical step (in single text form) is "*Blue Vitriol is a liquid and costs 22sl per litre. Mercury is a liquid and costs 3sl per litre. These can be mixed together in the* *mixer to form Vermillion. The Mixer costs 5sl per use. 1 unit of Vermillion is made from 2/3rd of a unit of Blue Vitriol and 1/3rd of a unit of Mercury. The Mixer settings to make Vermillion from Mercury and Blue Vitriol are 1 hour and slow*". There are six such steps that can be combined in various ways, as there are a number of possible recipes for making gold. The *single* versions of the task present this information in one ER. The *multiple* versions present this information in four separate ERs, which are not available co-presently. The content of the ERs is:

- *Chemicals* - the costs of chemicals
- *Apparatus* - the cost of apparatus
- *Formula* - how chemicals must be combined to create new chemicals
- *Settings* – the apparatus settings.

Consequently, for planning solutions information must be integrated from all the ERs. Then when operating the factory, only certain ERs needed to considered at each stage (*e.g.* when starting the fusion generator only the *settings* is needed).

Apparatus

The Alchemist's Factory was created in Macromedia Director. Six different versions that varied in how instructions were represented were created. These instructions are available before starting the factory and on request once inside the factory, but are not visible when someone is interacting with the factory.

Procedure

The Alchemist's Factory first gives a general introduction to the task and shows a sample single step with operations similar to that of the "gold" task. Participants were provided the appropriate form of instructions labeled with explanatory text. The experimenter demonstrated how to use the factory to follow these instructions. Finally, they were told that their aim should be to make gold in the cheapest way. Participants were then told to work independently with the factory until they were successful. If participants made an error such as overheating a chemical, they were informed of their mistake at the end of that step, and given the opportunity to repeat the process. After making gold, participants were asked to teach their apprentice to operate the factory by recording as much of the process as possible. Participants were given pen and paper and told they could use any format they chose for recording this information.

Dependent Variables

Efficiency was assessed by the amount of time that participants spent studying instructions. This was also computed prior to participants making their first move as indication of initial time spent planning. The effectiveness of ERs was determined by calculating the number of errors made in operating the factory, which solution participants chose and the money left after the gold was made. Learning was assessed by examining how many items the participants recalled correctly during the teachback. Given that participants who choose more complex solutions have a greater opportunity to recall more items, this was also coded as a percentage of the maximum they could have recalled given their solution. If it was not possible to determine what step they were trying to record because they had recalled insufficient correct details, the simplest solution was chosen.

Results

To examine the influence of information and computational properties of ERs on problem-solving, a [2 by 3] between groups MANOVA was computed (see Table 1). Analysis revealed a single main effect of complexity ($F_{1,78}$ = 130.11, MSE = 151.79, p<0.001); unsurprisingly the total of instruction requests was much higher when the instructions were presented in MERs although not four times as high. There were significant main effects of form on time spent studying instructions overall and prior to making a first move ($F_{2,78}$ = 5.17, MSE = 29921, p<0.01 & $F_{2,78}$ = 6.71, MSE = 8731, p<0.002). Post hoc comparisons showed that text was associated with significantly greater instruction times than either tables or diagrams (q= 3.6 p<0.05 & q = 4.1 p<0.05) overall and just with tables prior to operating the factory (q= 5.18 p<0.01). There was also a significant main effect of form on recall. Both for total number of items recalled ($F_{2,78}$ = 3.85, MSE = 69.50, p<0.025) where text was associated with significantly greater recall than tables (q = 4.47 p<0.01) and the percentage of items recalled ($F_{2,78}$ = 4.35, MSE = 33.13, p<0.02) where text was associated with significantly greater percentage recall than either tables or diagrams (q = 4.24 p<0.05 & q = 4.24, p<0.05).

Finally, there were interactions between form and complexity for the amount of time spent studying instructions overall (Figure 1) and prior to operating factory ($F_{2,78}$ = 4.01, MSE = 29921, p<0.025 & $F_{2,78}$ = 4.37, MSE = 8731, p<0.02). Simple main effects analysis showed that distribution of information influenced text but no other form of ERs ($F_{1,78}$ = 6.52, MSE = 29921, p<0.02 & $F_{1,78}$ = 7.52, MSE = 8731, p<0.01) and that the computational form of ERs was only significant for single ERs ($F_{2,78}$ = 9.0, MSE = 29921, p<0.001 & $F_{2,78}$ = 9.0, MSE = 8731, p<0.001). Tukey's post-hoc comparisons showed that single text led to significantly greater instruction times than both single diagram or single table overall (q = 5.20, p<0.01 & q = 5.19, p<0.01) and for prior to operation (q = 4.48, p<0.01 & q = 5.71, p<0.001).

Table 1: The influence of representational complexity and form on problem-solving and learning.

	Multiple			Single		
	Diagrams N = 14	Tables N = 14	Texts N = 14	Diagram N = 14	Table N = 14	Text N = 14
In. Times	426.7	392.1	424.9	351.4	351.8	591.9
St. Dev.	(134.1)	(166.0)	(178.1)	(166.6)	(163.3)	(218.7)
In. Times (pre)	193.0	121.0	161.3	148.4	117.5	260.2
St. Dev.	(82.7)	(49.3)	(71.3)	(86.5)	(54.4)	(166.1)
No. Errors	1.7	2.0	1.4	2.1	1.1	0.9
St. Dev.	(1.8)	(2.1)	(1.8)	(3.0)	(1.1)	(0.9)
Money Left	391.1	399.0	395.1	393.3	388.1	396.2
St. Dev.	(18.8)	(11.5)	(13.7)	(21.4)	(26.3)	(13.0)
Items Recalled	20.6	19.3	27.6	22.5	27.7	20.7
St. Dev.	(10.6)	(8.8)	(4.4)	(8.9)	(8.1)	(8.9)
Percentage Recalled	41.8%	37.3%	51.2%	56.6%	37.1%	49.0%
St. Dev.	(19.2%)	(19.2%)	(13.8%)	(18.1%)	(21.9%)	(15.8%)

Key: In. Times = time studying instructions in seconds; In. Times (pre) = time studying instructions prior to making a first move. No. Errors = mistakes in operation; Money Left = Cost of the solution; Items Recalled = items correctly recalled after the task; Percentage Recalled = percentage of items correctly recalled given the participant's solution.

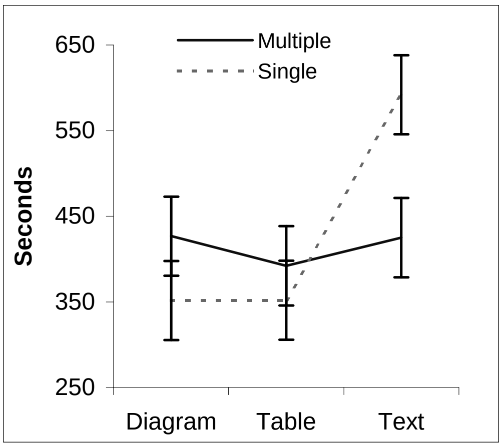

Figure 1: The time spent studying instructions by complexity and format.

Pearson correlation co-efficients were calculated for these dependent variables to determine if time studying instructions was related to performance. Instruction time correlated with number of errors but and negatively with the amount money left. However, there was no correlation between instruction time and what participants remembered about the task.

Table 2: Pearson correlation co-efficients

	2	3	4	5	6
1. In. Times	0.67**	0.34**	-0.21*	0.02	-0.04
2. In. Times (pre)		0.01	0.17	0.09	0.09
3. No. Error			-0.49**	0.03	0.05
4. Money Left				0.07	0.10
5. No. Recalled					0.85**
6. % Recalled					

* = p<0.05, ** = p<0.01, (two tailed test).

To determine if the form or complexity of ER influenced whether participants found the best solution to the problem (*i.e.* the cheapest solution), the number of participants who identified this ideal solution (irrespective of whether they made any errors in following this solution) was calculated by complexity (Table 3 Panel A) and form (Table 3 Panel B)

Table 3: No. of optimum solutions by (a) complexity and (b) form

	Optimum Solution	Other Solutions
Multiple	10	32
Single	18	24
Diagram	8	20
Table	6	22
Text	14	14

Chi squares analysis showed that there were more optimum solutions in the single conditions (X^2 = 3.42, df = 1= p<0.032 (one sided)). There was also a trend for the form of ER to affect whether the optimal solution was found (X^2 = 5.57, df = 2, p<0.062 (two sided)).

Finally, we examined if the way instructions were represented influenced the way participants recalled the information. Three main types of form were evident in participants' records – text, diagrams (either as a flow chart or the presented diagram), and a formula ER that presented the solution in pseudo chemical notation (Figure 2). Only one subject produced a table ER (in the diagrams condition). The form of instruction influenced the way that participants wrote their teachbacks (Fisher's exact test = 12.71, p <0.05). Three one way chi square analyses showed that that were equal number of formula representations in all three conditions but that use of text (X^2 = 7.47, df = 2, p<0.025) and diagrams (X^2 = 5.92, df = 2, p<0.05) was affected by condition

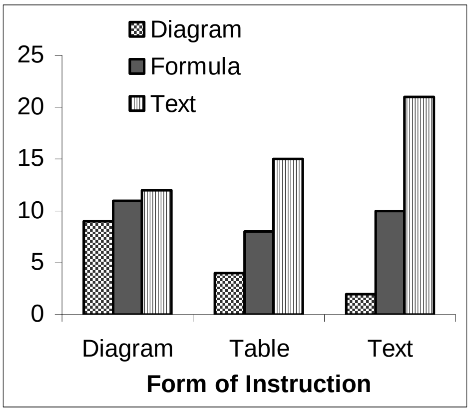

Figure 2: Form of teachbacks by instructional form.

Irrespective of the number of ERs studied, participants tended to integrate all the information into a singe ER (Fisher's exact = 2.40, p = ns) (Table 4).

Table 4. No of representations in teachbacks by representational complexity.

	One	Two	Three
Multiple	29	10	2
Single	33	7	0

Discussion

Participants in all conditions made the same number of errors in completing the task. They also produced roughly equivalent solutions in that the average number of silver pieces spent to create gold (which is a result of the cost of solution and any errors made) was equal. However, the number of participants who found the best solution for the problem was influenced by the way that the instructions were represented. As predicted, a single complete ER did lead to better performance. Presenting information in one ER seemed to allow easier comparison of information, which helps participants identify the best solution. There was also a trend for ER form to influence whether the best solution was found. Participants given text found the optimum solution more often. Overall, more people in the single text condition discovered the optimal solution than in any other condition (9/14 in ST compared to a maximum of 5/14 in the other conditions). The single text ER was also associated with significantly greater reading times prior to action and overall. We propose that finding the optimal solution was not 'easier' in this condition but instead, in line with O'Hara & Payne (1998) and Svendsen (1991), increasing cost of operation encouraged people to engage in more planning behavior. In this case increased planning was associated with increased performance.

After successfully completing the task, participants were asked to record as much of their solution as possible so as to teach their apprentice how to make gold. Consequently, it was possible to assess a simple form of learning – how much information participants had successfully encoded during the task. The form of instruction influenced the amount of information people recalled about the task. Participants in the text condition remembered more about the task than people in the table or diagrams conditions. This was true both for amount of information recalled and as a percentage of the maximum information that could have been recalled given their solution.

Two alternative hypotheses were identified for the influence of ERs on learning. One possibility was that those ERs which make the task easier by reducing working memory demands whilst problem-solving would facilitate learning. The second hypothesis was that ERs, which are difficult to search, will lead to greater learning as participants will aim to minimize the time spent searching by memorizing information. The results do not unequivocally support either hypothesis. Recall in both text conditions was nearly identical – yet the single text had been associated with increased time whereas the multiple text led to no more reading time than the other MERs. In fact there was no relationship between time spent studying instructions and what was subsequently remembered about the task.

One possibility is a stimulus-response compatibility effect. Although participants could use any form of ER, a high number chose to use text. Overall, nearly 60% of people use either text by itself or combined it with a pseudo-chemical ER. In the text condition, 88% of participants used a text form compared to the diagrams condition (50%) or table condition (62.5%).

Overall, there is little evidence that problem solving or learning was enhanced by using MERs, adding further evidence for the negative effects of MERs. This may because this task requires participants to integrate information from separate ERs in order to successfully plan solutions. That they did so is evident from the fact that the majority of teachback ERs are in a single integrated form. If tasks require less coordination of information then potentially MERs may be more beneficial. Furthermore, all the ERs in each condition were of the same form (i.e. all tables or all diagrams or all text). Once information is distributed, it can be represented in different forms. For example, it may be the case that information primarily useful for planning should be presented textually, whilst that aimed at efficient operation of the device in tabular form.

Conclusions

The results of this experiment suggests there is a complex relationship between computational and information properties of ERs and effects on problem-solving and learning between these factors in this domain. The optimum form of ER depends on whether the task was to find the single best solution, to find a satisfactory solution in a time-effective manner or to memorize the most about how to complete the task.

These results suggest that if the aim is to encourage people to find the optimum solution to a problem, then performance will be facilitated if the instructions are presented either in text rather than diagrams or tables and/or in one complete ER rather than in MERs. As the most successful ER was single text, it would suggest that this may occur, apparently paradoxically, because of the difficulty of working with this form of ER. However, finding this solution occurred at the expense of increasing time spent reading instructions. Hence, if the goal is to encourage quick task completion or if solutions are roughly equivalent, then the worst ER to select may be a textual ER that presents a great deal of information. In this case performance will be facilitated by either using MERs or a single diagram or table.

Acknowledgments

This research was supported by the Engineering and Physical Research Council. The authors would like to thank David Clarke for helpful comments on a draft of this paper and J.K. Rowling for inspiring the task.

References

Ainsworth, S. E. (1999). The functions of multiple representations. *Computers & Education, 33*, 131-152.

Ainsworth, S. E., Bibby, P., & Wood, D. (2002). Examining the effects of multiple representational systems in learning primary mathematics. *Journal of the Learning Sciences, 11*, 25-61.

Ainsworth, S. E., Bibby, P. A., & Wood, D. J. (1997). *Evaluating principles for multi-representational learning environments.* Paper presented at the 7th EARLI Conference, Athens.

Bibby, P. A., & Payne, S. J. (1993). Internalization and the use specificity of device knowledge. *Human-Computer Interaction, 8*, 25-56.

Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology, 62*, 233-246.

Cheng, P. C., Lowe, R. K., & Scaife, M. (2001). Cognitive science approaches to understanding diagrammatic representations. *Artificial Intelligence Review, 15*, 79-94.

Cox, R., & Brna, P. (1995). Supporting the use of external representations in problem solving: The need for flexible learning environments. *Journal of Artificial Intelligence in Education, 6*, 239-302.

Gilmore, D. J., & Green, T. R. G. (1984). Comprehension and recall of miniature programs. *Int. Journal of Man-Machine Studies, 21*, 31-48.

Kalyuga, S., Chandler, P., & Sweller, J. (1999). Managing split-attention and redundancy in multimedia instruction. *Applied Cognitive Psychology, 13*, 351-371.

Kozma, R., Chin, E., Russell, J., & Marx, N. (2000). The roles of representations and tools in the chemistry laboratory and their implications for chemistry learning. *Journal of the Learning Sciences, 9*, 105-143.

Larkin, J. H., & Simon, H. A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65-99.

Mayer, R. E., & Sims, V. K. (1994). For whom is a picture worth 1000 words - Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 86*, 389-401.

Meyer, J., Shinar, D., & Leiser, D. (1997). Multiple factors that determine performance with tables and graphs. *Human Factors, 39*, 268-286.

O'Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive Psychology, 35*, 34-70.

Svendsen. (1991). The influence of interface style on problem solving. *International Journal of Man-Machine Studies, 35*, 379-397.

Tabachneck, H. J. M., Koedinger, K. R., & Nathan, M. J. (1994). Towards a theoretical account of strategy use and sense making in mathematical problem solving. *16th Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: LEA.

Tabachneck-Schijf, H. J. M., & Simon, H. A. (1998). Alternative representations of instructional material. In D. Peterson (Ed.), *Forms of representation*. Exeter: Intellect Books.