Mental Effort and Errors in Bracket Expansion Tasks

Paul Ayres
University of New South Wales
<p.ayres@unsw.edu.au>

Previous research has shown that when students expand brackets, errors tend to cluster around key positions due to increased working memory load. This study found that by neutralising the effects of position, the occurrence of error clusters was reduced. Furthermore, a self-rating mental effort instrument was employed which found a positive correlation between errors and mental effort. This instrument also detected subtle variations in mental effort between groups of varying mathematical ability.

The cause of mathematical errors can be quite complex, as errors can be caused by limitations in working memory (Kintsch & Greeno, 1985) as well as by lack of knowledge. Considerable research has been conducted into the nature and cause of mathematical errors. Hitch (1978) found that the solution to a mental calculation, such as 347 + 189, was highly dependent upon the problem solver's ability to hold and process information. If any of the initial information or interim subtotals were forgotten, errors would occur. Consequently, Hitch argued that mental arithmetic errors were caused by decay in the storage of problem information. More recent research by Ashcraft, Donley, Halas and Vakali (1992), and Logie, Gilhooly and Wynn (1994), has confirmed the link between calculation errors and loss of information. Furthermore, Campbell and Charness (1990) found that when students were mentally required to square large numbers (such as 74) by using a calculation algorithm, errors had a higher chance of occurring at "heavy traffic stages in the squaring task" (p. 887). Campbell and Charness argued that these stages corresponded with the high demands of keeping track of subgoals and their results. In addition, Ashcraft et al. (1992) reported that the retrieval of basic addition facts could also be affected by working memory demands. An increase in problem difficulty corresponded to less accuracy in recalling number facts. This last finding is of some significance as it suggested that working memory load affects the retrieval of information as well as its storage.

Much of the research into the relationship between working memory and errors has been conducted in the domain of mental arithmetic. However, evidence of this relationship has also been found in other areas of mathematics such as arithmetic word problems (Fayol, Abdi & Gombert, 1987), geometry (Ayres & Sweller, 1990; Ayres, 1993) and algebra (Ayres, In press). Fayol et al. employed word problems and found that the way that the problem texts were organised led to significant performance differences. Changing the order of the wording of the problems, for example, led to more errors and a reduced performance. Fayol et al. argued that certain text structures forced the participants to problem solve in a manner (bottom up) which overloaded working memory and forced errors. Similarly, Ayres and Sweller (1990) identified a specific error profile in two-move geometry experiments which required the compulsory calculation of a subgoal first before the ultimate goal could be reached. It was discovered that students made more errors in the calculation of the subgoal stage than on the goal stage. The design of these geometry experiments was such that all problems were counterbalanced. In other words, the application of a theorem such as the "angle sum of a triangle is 180 degrees" occurred an equal number of times as a subgoal and as a goal. Consequently, if all cognitive demands were equal, it was expected that errors connected to the triangle property would be equally distributed over both stages. This was not the case, as more errors were reported at the subgoal stage. Consequently, it was argued that errors were not necessarily made because the geometry theorems were poorly learnt, but because the subgoal stage demanded more working memory resources and forced errors.

Error clusters was also identified by Ayres (In press;) on bracket expansion tasks of the type: -3 (-4 - 5x) - 2 (-3x - 4). These tasks require the following four operations to be completed: -3 * -4 (Operation 1), -3 * -5x (Operation 2), -2 * -3x (Operation 3), and -2 * -4 (Operation 4). It was discovered that more errors were made during the expansion of the second bracket compared with the first bracket, and more errors were made during the second operation (Operations 2 and 4) compared with the first operation (Operations 1 and 3) within each bracket. Nearly all errors made involved the manipulation of signs. The clustering of errors indicated that some students could multiply -3 and -4 together correctly when they appeared during the first or third operation, but made errors (-12) when they appeared during the second or fourth operation. To explain this result, Ayres argued that more decision-making processes are made on the second operation within each bracket and during the expansion of the second bracket due to the dual-role of signs which link together brackets and operations. Consequently, as a result of increased decision making, working memory load was not equally distributed over the four operations. At points where the load was heaviest, information was either lost from working memory or incorrectly recalled from long-term memory and errors resulted.

To investigate the working memory load explanation further, Ayres (In press) completed an additional experiment using a dual-task methodology (see Ashcraft, 1995). Prior to completing a specific operation within a bracket problem, students were given a random sequence of letters to remember, which they were required to recall after completion of the computation. It was found that recall varied across the four operations. In particular, recall connected to the fourth operation was the poorest, suggesting that this operation caused the heaviest working memory load. Whereas, this result was consistent with a working memory load model, two anomalies were identified. Firstly, variations in recall were only found on operations that included an x-term. For operations that were purely numerical, no variation in recall was found. This was surprising, as it was found in earlier experiments that error clustering would occur regardless of the inclusion a pro-numeral. Secondly, the students made few calculation errors. This was also unexpected as students were required to simultaneously remember a random sequence of letters and complete a calculation. To explain these anomalies, Ayres suggested that because students were only required to calculate one operation per problem (instead of four), the overall working memory load could have been reduced.

The findings described in the previous paragraph are interesting, because if asking students to complete only one operation per bracket problem reduces working memory load and error rates, then this strategy may have the potential to be an effective instructional technique. Cognitive load theorists such as Sweller argue that instructional techniques are most effective when cognitive load is reduced (see Sweller, 1999). Consequently, this study was designed to investigate further the effects of requiring students to calculate only one operation on bracket expansion tasks that normally require four. More specifically it attempts to examine under this condition:

- whether error clusters appear;
- how students rate their own mental loads per calculation;
- whether error rates and self-rating mental loads vary according to mathematical ability.

Method

Participants

Sixty two students from Grade 8 of a Sydney high school participated in this study. On the basis of Grade 7 mathematics results the students had been streamed into three groups on their entry into Grade 8. The group of highest general mathematical ability is referred to as Group 1 (n =23), the next group as Group 2 (n=22) and the final group as Group 3 (n=17). All students had prior experience of expanding brackets of the type used in this study.

Materials and Procedure

A set of eight problems, each containing two brackets, was designed (see Table 1) and recorded on a work sheet. For each problem, one specific operation was targeted as the designated calculation for that problem. These operations were flagged for the students' benefit by drawing an arrow downward from the relevant term within the bracket. Below the arrowhead was drawn a dotted line indicating where the student was required to complete the given computation. For example, the fourth operation (-3 * -7) was the targeted calculation in Problem-1 (see Table 1) and consequently an arrow was drawn downward from the "-7" in the bracket. This indicated to the student that they were required to complete this calculation only. Overall, eight calculations were required. To detect error clusters without bias, the problem set consisted of a number of counterbalanced pairs. These pairs featured the calculations "-4*3" (operations 1 and 2 in problems 2 and 8), "-3*-7" (operations 3 and 4 in problems 6 and 1, "-3*-8x" (operations 1 and 2 in problems 5 and 3), and "-2*-9x" (operations 3 and 4 in problems 4 and 7). This design ensured that a calculation including a pro-numeral (x) and without a pro-numeral occurred in all four operational positions.

Table 1
The Problem Set Used in the Study

Problems	Targeted Operation	Code
1. 5 (3x - 4) - 3 (6x - 7)	4th	Op. 4
2. -4 (3 + 6x) + 3 (2x - 4)	1st	Op. 1
3 -3 (2 - 8x) + 3 (8x - 7)	2nd	Op. 2x
4. -4 (3x - 1) - 2 (9x - 5)	3rd	Op. 3x
5. -3 (-8x + 2) + 3 (2x - 4)	1st	Op. 1x
6. 5 (3x - 4) - 3(-7 + 6x)	3rd	Op. 3
7. -3 (2 - 8x) - 2 (-5 + 9x)	4th	Op. 4x
8. -4 (6x + 3) - 2 (4x + 5)	2nd	Op. 2

To measure mental load, a self-rating technique developed by Paas and van Merriënboer (1994) was employed. Using a 7-point Likert Scale (ranging from 1- very easy to 7- very difficult) students were asked to self-rate the mental effort involved in each calculation. Students were asked to do this immediately after the completion of each calculation. A designated space was provided on the work sheets for each problem. Before commencing the problem set, a practice example was given by the researcher. During this phase instructions were given and students were allowed to ask questions. Calculators were not allowed and sufficient time was given for students to complete the problem set. It is worth noting that this technique has been used successfully over the last six years in measuring the effectiveness of instructions (see Sweller, 1999). However, as far as the author is aware, it has never been applied in the context suggested in this study.

Results and Discussion

Analysis of Errors

The number of errors made by each student was recorded according to their positional operation and whether they included a x-term or not (see Table 2). A total of 60 sign errors were made. Very few arithmetical errors were made (10%) and were not included in the analysis of the data as they were not the focus of this study. The mean number of sign errors made were 0.3 (SD = 1.4), 1.1 (SD = 1.3) and 1.7 (SD = 1.7) for Groups 1, 2 and 3 respectively. A one-way ANOVA revealed significant between group differences: $F(2, 59) = 6.77$, $p < 0.01$. A post-hoc Tukey-HSD test revealed that the highest ability group (Group 1) made significantly fewer errors than the lowest ability group (Group 3), as might be expected. Group 1 also made fewer errors than Group 2, and Group 2 made fewer errors than Group 3, although these results were not significantly different.

To examine differences between the operations themselves, a number of paired t-tests were completed. No significant differences were found between Operation 1 and Operation 2 ($t = 1.00$), Operation 3 and 4 ($t = 1.00$), and Operation 3x and Operation 4x ($t = 0.33$). However, a significant difference was found between Operations 1x and 2x, ($t(61) = 4.8$, $p < 0.01$), indicating that students made more errors multiplying -3 and - 8x together when it appeared as the second operation rather than the first. Furthermore, the total number of errors made in the first bracket was not significantly different to the total number of errors made in the second bracket: $t(61) = 1.49$, $p = 0.14$). As no differences were found between the two brackets and Operations 3 and 4 (with or without x's) the results are not consistent with the Ayres (In press) study which found large differences on these comparisons when all operations were completed for each question rather than just one. Nevertheless, the significant difference found between Operations 1x and 2x, suggests that Operation 2x may have provoked an increase in working memory load at this point. It is also worth noting that operations with x caused significantly more errors than operations without x under a paired t-test: $t(61) = 4.18$, $p < 0.01$.

Table 2

Frequency of Errors made on each Operation by Group

	Op. 1 -4*3	Op.1x# -3*-8x	Op. 2 -4*3	Op. 2x -3*-8x	Op. 3 -3*-7	Op. 3x -2*9x	Op. 4 -3*-7	Op. 4x -2*9x
Group 1	0	0	1	3	0	1	0	1
Group 2	1	0	2	7	4	5	0	6
Group 3	0	1	0	9	4	5	4	6
Totals	1	1	3	19	8	11	4	13

Note. x indicates that a pro-numeral was present in the operation.

Analysis of Self-rated Mental Effort

Self-rating mental effort scores were recorded for each individual operation (see Table 3). A Cronbach alpha run on this data gave an alpha coefficient of 0.90 indicating that the instrument had high internal consistency. Generally students did not rate these tasks as difficult. A maximum score of 2.2 on any operation indicated that the students rated the tasks as easy to fairly easy. However, a MANOVA (Hotelling's Trace) using the eight operations as dependent variables found a significant difference between the groups: T = 1.98, p = 0.02. Uni-variate tests (see Table 3) revealed that operations 1x, 2, 3 and 4x had p-values of less than 0.01 and therefore contributed to the overall group differences. In these cases, the high ability group (Group 1) exhibited significantly lower scores than the other two groups. This was consistent over the other four operations as well, except for Operation 4 (-3*-7), where Group 2 rated their mental effort lower than the other groups.

Table 3

Self-rating Mental Effort for each Operation

	Op. 1 -4*3	Op.1x -3*-8x	Op. 2 -4*3	Op. 2x -3*-8x	Op. 3 -3*-7	Op. 3x -2*9x	Op. 4 -3*-7	Op. 4x -2*9x
Group 1	1.4 (0.82	1.4 (0.50)	1.3 (0.47)	1.8 (0.74)	1.3 (0.88)	1.4 (0.59)	1.8 (1.23)	1.4 (0.58)
Group 2	1.8 (0.81)	2.1 (0.75)	2.1 (1.02)	2.2 (0.81)	2.1 (0.75)	2.1 (0.90)	1.6 (0.91)	2.2 (1.10)
Group 3	1.9 (0.86)	2.3 (1.16)	1.9 (1.03)	2.1 (1.03)	2.2 (1.48)	1.7 (0.85)	1.8 (1.07)	2.2 (1.03)
Combined	1.7 (0.81)	1.9 (0.89)	1.8 (0.92)	2.0 (0.86)	1.8 (1.10)	1.7 (0.82)	1.7 (1.07)	1.9 (0.95)
F values	1.75	7.13**	5.24**	1.56	4.92**	3.67*	0.46	4.94*

Note. Standard deviations recorded in parentheses; *p < 0.05; **p< 0.01.

To investigate the differences in self-rating mental load within the problem-set, a number of matched t-tests were run on the counterbalanced pairs (see Table 4).

Table 4

Paired t-test Values on Paired Groupings for Mental Effort Measures

	Group 1 t values	Group 2 t values	Group 3 t values	Combined t values
Op 1 v Op 2	0.70	1.67	0.39	1.01
Op 1x v Op 2x	2.86**	0.90	1.33	1.31
Op 3 v Op 4	2.02^	2.13*	1.07	0.60
Op 3x v Op 4x	0.16	0.66	2.58*	1.25
Bracket 1 v Bracket 2	0.41	1.06	1.05	0.88
Numerical v Algebraic	0.56	3.58**	0.55	2.54**

Note. **p < 0.01, *p < 0.05, ^p = 0.06.

Although a number of significant differences were identified on these comparisons, there was little consistency between the groups. For example, in the comparison between operations 3 and 4 (no x's), the top group rated Operation 4 (1.8) more difficult than Operation 3 (1.3); whereas Group 2 rated Operation 3 (2.1) more highly than Operation 4 (1.6). The combined group data produced only one significant difference; viz, operations with algebraic terms were rated more difficult than purely numerical operations. Nevertheless, this result was consistent with the error data, which indicated that more errors were made on the operations that included an algebraic term. The error data also revealed that more errors were made during Operation 2x rather than 1x. However, only Group 1, who made few errors, ranked Operation 2x more difficult that Operation 1x. Despite the variations between the groups, there was a correlation of 0.65 ($p = 0.08$) between the error rate and the mean mental effort over the eight operations. Although this result fails to reach significance at the 95% level, it is fairly close and may indicate a real relationship between mental effort and the occurrence of errors.

General Discussion

Examination of the error profiles revealed that errors were not generally clustered around key points such as Operations 2 and 4 which tends to happen when students calculate all four operations (see Ayres, In press). This study therefore provides some evidence to support the hypothesis that asking students to complete one calculation only per question, reduces the overall working memory load. However, the data did not totally support this argument because one error cluster was observed when students made more errors on Operation 2x than 1x. It is worth noting that " -3*-8x" includes two negatives and a x term. Studies have shown that many students have great difficulty understanding the basic concepts of algebra (see Küchemann, 1981; Herscovics & Linchevski 1995; MacGregor & Stacey, 1996) and negative numbers (see Gallardo, 1994; Herscovics & Linchevski, 1995). It is therefore feasible that the manipulation of an algebraic expression with two negative numbers is fairly complex and may have exerted a heavier load on working memory than the other combinations. This increased load may have re-acted more directly with the bracket task to cause the error cluster. As the results are not conclusive, it is too early to assess the potential of a single calculation technique as an instructional method. Future studies might explore this idea further by directly comparing the single calculation method with the traditional full bracket expansion method.

The use of the self-rating mental effort instrument proved reliable. The instrument demonstrated a high internal consistency and was sensitive enough to identify differences between groups. For example, the group with the highest mathematical ability rated the overall task less demanding than the other two groups, as might be expected. A positive correlation (0.65) was found between errors and mental effort, suggesting that, as mental effort increases, errors are more likely to occur. Generally, calculations that included pro-numerals were rated as requiring more mental effort and tended to promote more errors. As self-rating scales of mental effort were found to be useful in this study, they may have a greater role to play in the learning and teaching of mathematics.

Finally, the use of errors in the teaching of mathematics has long been valued by mathematics educators (see Borasi, 1994; Ashlock, 1986). Some researchers (Brown & Burton, 1978; Resnick, Nesher, Leonard, Magone, Omason, & Peled, 1989) argue that error analysis is an important diagnostic tool in identifying common misconceptions and wrongly applied strategies. This study on errors, like others, has uncovered some useful information about student performance and perceptions. It is therefore argued that research into errors should be continued, as it provides a rich source of knowledge about the processes and influences involved in mathematical problem solving.

References

Ashcraft, M.H., Donley, R.D., Halas, M.A., & Vakali, M. (1992). Working memory, automaticity, and problem difficulty. In J.I.D. Campbell (Ed.), *The nature and origins of mathematical skills (Advances in psychology "91),* (pp. 301-329), Amsterdam: Elsevier.

Ashcraft, M. H. (1995). Cognitive psychology and simple arithmetic: A review and summary of new directions. *Mathematical Cognition*, *1*(1), 3-34.

Ashlock, R.B. (1986). *Error patterns in computation: A semi-programmed approach, 4th edition*. Columbus, Ohio: Charles E. Merrill Publishing Company.

Ayres, P. (In press). Systematic mathematical errors and cognitive load. *Contemporary Educational Psychology.*

Ayres, P. (1993). Why goal-free problems can facilitate learning. *Contemporary Educational Psychology, 18,* 376-381.

Ayres, P. & Sweller, J. (1990). Locus of difficulty in multi-stage mathematics problems. *American Journal of Psychology*, *103*(2), 167-193.

Borasi, R. (1994). Capitalising on errors as "springboards for enquiry": A teaching experiment. *Journal for Research in Mathematics Education*, *25*(2), 166-208.

Brown, A.L. & Burton, R.R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science, 2*, 155-192.

Campbell, J.I.D. & Charness, N. (1990). Age-related declines in working-memory skills: Evidence from a complex calculation task. *Developmental Psychology, 26,* 879-888.

Fayol, M., Abdi, H. & Gombert, J.E. (1987). Arithmetic problems formulation and working memory load. *Cognition and Instruction, 4*, 187-202.

Gallardo, A. (1994). Negative numbers in algebra: The use of a teaching model. In J.P. Da Pointe & J. F. Matos (Eds.). *Proceedings of the XIX Annual Conference for the Psychology of Mathematics Education*., Lisbon, Portugal, Vol. 2, 376-383.

Herscovics, N., & Linchevski, L. (1995). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics*, *27*(1), 59-78.

Hitch, G. J. (1978). The role of short-term working memory in mental arithmetic. *Cognitive Psychology, 10*, 302-323.

Kintsch, W. & Greeno, J.G. (1985). Understanding and solving word arithmetic problems. *Psychological Review, 92*, 109-129.

Küchemann,, D. (1981). Algebra. In K. Hart (Ed.), *Children's understanding of mathematics: 11-16*. Murray: London, 102-119.

Logie, R.H., Gilhooly, K.J & Wynn, V. (1994). Counting on working memory in arithmetic problem solving. *Memory and Cognition*, 22, 395-410.

MacGregor, M. & Stacey, K. (1996). Progress in learning algebra: Temporary and persistent difficulties *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education*., Valencia, Spain, Vol. 3, 297-304.

Paas, F. G. W. C., & van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology. 86*(1), 122-133.

Resnick, L.B., Nesher, P., Leonard, F., Magone, M., Omanson, S. & Peled, I. (1989). Conceptual bases of arithmetic errors: The case of decimal fractions. *Journal for Research in Mathematics Education*, 20(1), 8-27.

Sweller, J. (1999) *Instructional Design*. Melbourne: ACER.