Patterns, Language and Algebra: A Longitudinal Study

Ted Redden
University of New England

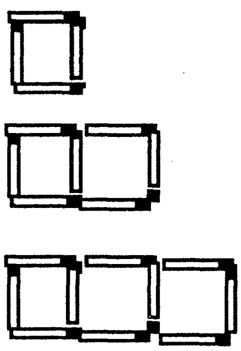
This paper reports on the some findings of a longitudinal study into children's understanding of number patterns by considering the changing use of natural language and symbolic notation over time. The study had its roots in an earlier study that provided a classification system for the responses. The two findings reported here suggest that responses to pattern stimulus items change over time and that the use of natural language in the highest category seems to be a necessary precursor to the emergence of algebraic notation.

Introduction

Following the popularisation of using number patterns as a contextual vehicle for the initial development of algebraic ideas (Mason 1985; Pegg and Redden 1990a; NSW Syllabus 1989; Romburg 1989; Australian Education Council, 1991) some research output has begun to report the efficacy of the approach (Pegg and Redden 1990b; Arzarello 1991; MacGregor and Stacey 1993). This research has used a variety of instruments and strategies to identify response categories to pattern stimulus items. The next phase of this research might investigate the dynamics of children's development in this area of cognition. This paper is based on some findings from a survey of 1435 children (Redden 1994a; Redden 1994b) and investigates some further issues in the context of a longitudinal study.

An analysis of the survey data (Redden 1994a) has identified an hypothesised sequence of development in the natural language used by children to describe number patterns in response to both geometrical and numeric stimulus items. An example of the items used is represented in Figure 1.

Here are some chains of Matches



(a) What is the next term in the pattern?

(b) Describe a general rule for the pattern in natural language

(c) Calculate the value of an uncountable term (eg. n=80)

(d) Write their rule in the symbolic notation of mathematics

Figure 1
A number pattern Stimulus Item

In brief, there were four major response categories identified for question (b) above which are listed here with a brief descriptor of the category and a sample response.

1. **Inappropriate** response. This reflected no attempt to answer the question or a failure to understand any aspect of the question. e.g., "It depends."
2. **One example**. This class of response gave the value for a specific example rather than a general description. e.g., "10 squares needs 31 matches."
3. **Successive** description. Respondents only made use of the dependent variable. e.g., "You start at four add three every time."

4. **Function.** This group of respondents described a relation between a dependent and independent variable. e.g., "The number of squares times three plus one gives the number of matches."

This sequence of development was validated by a number of procedures which included investigating the modal response category for each age cohort, comparing the response categories with categories identified in other studies (Stacey 1989; Ursini 1990; Meira 1990), and by using the SOLO Taxonomy (Biggs and Collis 1982) as a theoretical framework to analyse the complexity of the response. Further, the categories of symbolic notation were found to be strongly associated with the categories of natural language. The categories for classifying the symbolic notation were reduced to four for the purposes of further analysis. They were:

1 **No attempt.** No attempt was made to use symbolic notation.
2. **Operation Symbols.** Attempts at symbol use were restricted to those commonly used in arithmetic. e.g., "3 x 4+1".
3. **Arbitrary** use of letters. This category included a range of responses including replacing every number with a letter (a+b=c) or trying to incorporate some form of repetitive notation e.g., "a+3+3+3+3.......... forever".
4. **Algebra.** Successful use of algebraic notation. e.g., "y=4+3(x-1)".

In fact, the strength of the association between natural language and symbolic notation, and its existence across a number of pattern items, suggested that the ability to use natural language to describe number patterns may be a necessary precursor to being able to provide adequate algebraic descriptions of the relationships inherent in the number patterns.

To further investigate these and other issues a longitudinal study was undertaken over a two year period. Two of several research questions of the longitudinal study are discussed in this paper.

1. Does children's ability to express generality using natural language change over time?
2. Does the symbolic language of algebra follow the natural language development?

Following a brief outline of the research method, an analysis in relation to each of the research questions is presented. The paper concludes with a short discussion of some implications of the findings.

Method

Twenty six children from years five six and seven were selected as a stratified sample of the 1435 children included in an earlier survey. The stratification reflected both age and ability. The children of the sample were interviewed five times at six monthly intervals over the two year period. These are referred to as studies 1 to 5. During the thirty minute interviews they were asked a number of questions about three pattern stimulus items of varying complexity. The discussion here will focus on the responses to the items at the middle level of complexity. A number of items at this level of complexity were constructed of which the item in Figure 1 is typical.

The interviews were audio-taped and later transcribed by the researcher for later analysis in conjunction with the written responses of the children. The interview structure was developed from that used by Booth (1984). The planned questions were supplemented by additional probes to clarify meanings and understandings when ambiguity existed.

The suitability of a longitudinal design for this study is detailed elsewhere (Menard 1991, Cohen and Manion 1989). However, there are a number of difficulties associated with such a design such as institutional change (some children changed schools), sample mortality (five children left the study) and the measurement effect of repeated interviews. Cohen and Manion (1989) argued that repeated interviewing often induces the subjects to respond in a manner different from that which would be provided in a more natural setting. In this context the researcher may be providing cues as to which type of responses are considered desirable. However, in this study it is not what induces the change that is of interest but rather the nature of the change. There are a number of influences that may induce the change, including maturation of the subjects, experiences in mathematics classes at school, and participation in this study. The way in which these contribute to the change in children's responses is beyond the scope of this discussion.

Analysis

Overview of Change

This part addresses the question of:

> Does children's ability to express generality using natural language change over time?

In determining whether or not there was a general change in children's ability to describe number patterns, attention was to be focused on the natural language responses. The Rasch model has been used here to assist in the identification of changes over time in students' ability to respond to the stimulus items. This was facilitated by an analysis of the item difficulty delta values reported by the item analysis statistic, "Stat-Tau", option of Quest (Adams &Khoo 1993, p.34).

Traditionally the Rasch model has been used to analyse students' responses at a single point in time on a number of items of variable difficulty. The Rasch model is capable of reporting δ_i values for each item. The δ_i values indicate the difficulty of that item in Study *i* compared to the other items in the set (Adams and Khoo 1993). If a set of identical items is presented to children over a period of time then the item difficulty (δ_i) values reflect changes in the children's ability to respond to the items. Hence, it is hypothesised that an item in Study 1 will have a higher difficulty value than the corresponding stimulus item in Study 2 and so on until Study 5. That is,

$$\delta_1 > \delta_2 > \delta_3 > \delta_4 > \delta_5$$

Once the values of δ_i are established their sequence can be simply observed; however, some test of significance of the differences needs to be applied. A number of alternatives in structuring such a test are available. The existence of a difference between the five delta values could be investigated, (i) by considering the null hypothesis that:

$$H_0: \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5,$$

(ii) by investigating stepwise differences using the null hypotheses:

$$H_0: \delta_1 = \delta_2, H_0: \delta_2 = \delta_3, H_0: \delta_3 = \delta_4, H_0: \delta_4 = \delta_5$$

or (iii) by investigating the significance or otherwise of the total change over the two-year period. The null hypothesis for this investigation would be:

$$H_0: \delta_1 = \delta_5$$

It was decided to use the first and third methods for investigating significant differences since they would provide answers to the questions:

> Does any significant difference exist? and,
>
> Was there a significant difference over the two-year period of the study?

Method (ii) was seen to have less potential since the repeated testing would result in increasing the chance of a Type I error, and since the SOLO Taxonomy predicts some instability in responses during development from one level to the next it would be possible for reversals in delta values to appear. These reversals would be most likely in adjacent studies with little development time between them.

The test for homogeneity of effect sizes (Hedges and Olkin 1985) was used for investigating the significance or otherwise of differences among a set of item difficulties. It is clear that there is a general shift from the lower levels of response to the higher levels over the two-year period (see Table 1). There is also a ceiling effect on many of the children's responses as they were already responding at the highest level at the beginning of the study or they achieved responses at the highest level during the study. Quest only uses cases and items that do not have a perfect score. Some of the infit statistics indicated a number of reversals, which were to be expected for two reasons. The first is the volatility of responses predicted by SOLO as transition between cognitive structures occurs. Secondly, the items lose power to discriminate as the ceiling is approached and a reversal is therefore more likely to occur. The fit statistics are reported elsewhere (Redden 1995).

Table 1 reports the frequency data for the natural language responses. A consistent pattern of growth is evident across the two years of the study. The ranks of the item difficulty values (δ) are consistent with the predicted pattern, i.e.,

$$\delta_1 > \delta_2 > \delta_3 > \delta_4 > \delta_5.$$

Further, on investigating the homogeneity of effect sizes, using the Q value described by Hedges and Olkin (1985), it was found that the differences are significant at the 0.01 confidence level. (Q=20.599, df =4). Hence the null hypothesis that

$$H_0: \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5$$

is rejected. Additionally, the null hypothesis of equality of delta values between Study 1 and Study 5 was rejected. (Q=15.738, df=1 and P<0.01).

Table 1
Frequency of Responses and Item Difficulty Values

(n=21)	IA	1EG	SUCC	FUNC	Difficulty (δ) SD	Rank of Difficulty δ
Study 1	8	5	3	5	1.14 .29	1
Study 2	3	4	5	9	.09 .31	2
Study 3	2	0	11	8	.08 .37	3
Study 4	2	0	6	13	-.49 .37	4
Study 5	1	0	6	14	-.82 .40	5

From the discussion in this part there is strong evidence in support of the proposition that:

An improvement in the children's ability to express generality over the two years of the study can be detected?

While there is clear evidence of general growth, what can be said about the actual path taken by the children in the sample? The suggested hierarchy of responses reported earlier was:

Inappropriate → **One example** → **Successive** → **Function**

Of the 84 transitions made between categories ten were classified as reversals since they failed to follow the predicted sequence by responding at a level below that of a previous response. However, all reversals with the exception of two were later compensated for. Thus it is argued that the predicted hierarchy reflected the general growth pattern of the children in the study and that the reversals reflected the instability of response patterns as the cognitive structures are changing.

Relationship between Natural and Symbolic Language.

Associated with each natural language response the children were asked to provide a symbolic language response. It has been reported elsewhere (Redden 1994b) that there was significant association between the categories of natural language and the categories of symbolic language. In particular these pairs of associated categories are: **(Inappropriate, No attempt), (One example, Operation symbols), (Successive, Arbitrary use of Letters)** and **(Function, Algebra).**

In order to investigate the question:

Does the symbolic language of algebra follow the natural language development

individual profiles for each child were developed which mapped their responses to the request for both natural and symbolic language against time. Two such profiles are presented here as Figure 2.

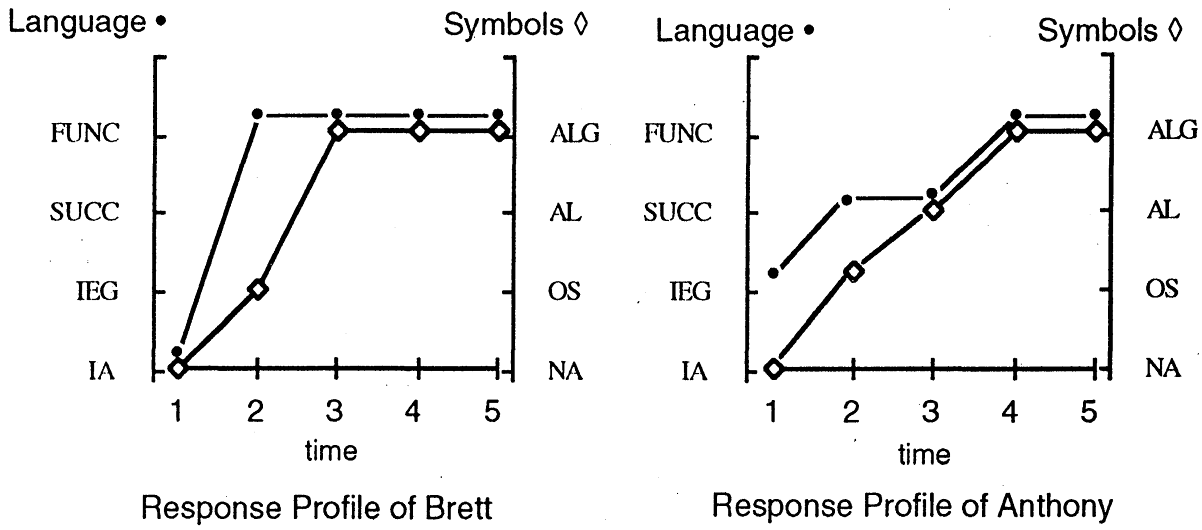
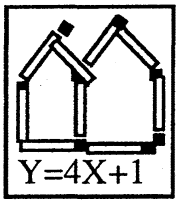


Figure 2

In both the profiles in Figure 2 the higher level of natural language appeared before or at the same time as the associated level of symbolic language. Of the 315 pairs of responses recorded in the study all but six followed this pattern. In 92 cases the level of symbolic response was below the level of natural language response, while in 217 pairs the levels matched as the earlier study predicted

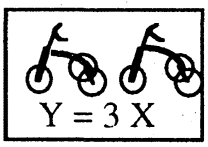
While the proposition that the symbolic language of algebra follows the natural language development is overwhelmingly supported by the frequency data the protocols provided insight into two issues related to this issue. The first is the issue of how do children provide accurate algebraic descriptions of patterns without functional relationships being first expressed in natural language. It will be recalled (see Figure 1) that between asking for a pattern description in natural language and symbolic language the children were asked to apply their rule to an uncountable example. It was this question that forced one child to reconceptualise his rule as the following protocol shows.



Researcher: How did you do Part (A)?
Anthony: For every house starting off with 1 house you need 5 matches and to add another house you need 4 for every one then I counted up those there and added 4 more
Researcher: Added 4 more? (Those there are the number for an extra house) So you have counted matched for 3 houses and added on 4 to get 17.
Anthony: Yes.
Researcher: Part (B). You have used a different rule?
Anthony: Starting with one house add 4 each time......... (goes to Part (C))........... 37 x4+1=149.
(then returns to Part (B) and changes description).......... Times the number of houses by 4 and add 1.
Researcher: Why are you doing that?
Anthony: I did Part (C) first then on a scrap of paper I got the same answer that I had (to Part (A))
Researcher: You got 17?
Anthony: Yes I had a go then I thought up another way and got a new rule
Researcher: So you did Part (C) then went back and did Part (B)?
Anthony: Yes.
Researcher: Part (C) forced you to think about it in another way?
Anthony: Yes.

It is clear that Anthony was forced to reconceptualise his pattern description by the need to calculate an uncountable example. If children are to be taught to express generality it may not be enough to merely provide the opportunity to express generality. They need to be given the opportunity to see the value of expressing generality.

The second issue that arose from the protocols was that of the permanence or otherwise of the relationship between natural language and symbolic language. There was limited evidence of children becoming so competent in their use of algebraic symbols that their search strategies for pattern descriptions took place using that symbolic language. Two children, Peter and Anthony, began to use algebraic notation instead of natural language. However, this practice appeared only towards the end of the longitudinal study. Anthony appeared to use algebraic language because of its convenience, while Peter seemed to conduct his search for the rule in symbolic language. In Study 5, Anthony initially asked for permission to use symbolic language.



Researcher: Part (A)?
Anthony: 3 x 4 = 12
Researcher: Part (B)?
Anthony: Can I just use Symbols?.... 3 x T = N
Researcher: What does that mean?
Anthony: Three times T equals N.....T is the number of triangles and N is the number of wheels.
Researcher: Part (C)?
Anthony: 3 x 55 = 165
Researcher: Part (D)?
Anthony: (as for Part (B)).

Researcher: If you had to write it in English what would you say?
Anthony: The number of tricycles times three gives you the number of wheels.

As Peter's confidence grew more reliance was placed on the symbolic notation.

□	1	2
△	2	5

Y=3X-1

Researcher: Part (B)?
Peter: There it is times 2 , then times 2 + 1, times 2+ 2, Times 2+ 3......Times 2+ 10 +4...Can I write it out....hey....2 times 4 is 8..+4 no plus 3....oh....so 2x2 +1...Yes I've got it....(writes (5x2) + (5-1)=14)....so its Xx2 +(X-1)=y.....yes that's it now to do it in an English sentence. The computer multiplies the top number by two and adds one less than the top number to its answer.

By Study 5 he was writing the pattern descriptions in symbolic notation and could not see the point in repeating the description in natural language. In other words, algebraic notation had become his preferred means of communication.

Conclusion

This paper has reported the results of a longitudinal study that attempted to probe the nature of change in children's response patterns using both natural language and symbolic notation. It was also found that the natural language descriptions of number patterns seem to be a necessary prerequisite for the emergence of algebraic notation as a means of describing the generality of number patterns. It may well be that the necessity for articulating accurate natural language descriptions is a transitory phenomenon.

While some children were describing patterns at the highest level at the beginning of the study, others did not achieve the highest level of response during the two years of investigation. It can be seen that the concept of expressing generality in the form that facilitates algebraic notation is not automatically available to all children who may be exposed to the Year 7 and 8 mathematics syllabus in N.S.W.

The other focus of this longitudinal study was to consider the anomalies identified within the data when it is compared to the theoretical relationships identified. These anomalies included children being able to provide correct answers to uncountable examples, and provide algebraic notation without providing functional descriptions in natural language. It was found that in most cases children are forced to reconceptualise the pattern description when confronted with a problem they cannot answer with their initial response. It would seem that in some cases it was not enough to ask for a generalisation, but the need for, and power of a generalisation needs to be demonstrated.

The study pointed to the elusive nature of children's cognition. In particular the potential for a lack of congruence between children's externalised responses and their 'hypothetical cognitive structure'(Biggs and Collis 1982, p. 22). The variability in the response patterns indicated a number of cases of mismatch to be identified and clarified. This process was facilitated by two design features. The first was the multiple data components collected for each stimulus item, which allowed apparent inconsistencies among responses to each component to be analysed more deeply. The second feature was the interview environment, where a variety of clues as to the child's true intention could be gleaned. This interview environment can be seen as a dynamic interaction between subject, the researcher and the stimulus item. This is contrasted with the more static environment (from the researcher's perspective) of survey data.

References

Adams, R. J., & Khoo Siek-Toon (1993). *Quest: The Interactive Test Analysis System.* Melbourne: Australian Council for Educational Research.

Arzarello, F. (1991). Procedural and relational aspects of algebraic thinking. In F. Furinghetti (Ed.), *PME 15* (Vol.1, pp. 80 - 87). Assisi: International Group for the Psychology of Mathematics Education.

Australian Education Council (1990). *A national statement on mathematics for Australian schools.* Carlton: Curriculum Corporation.

Biggs, J. B., & Collis, K. F. (1982). *Evaluating the quality of learning.* New York: Academic Press.

Board of Secondary Education, New South Wales. (1988). *Mathematics syllabus years 7-8.* Sydney: Board of Secondary Education.

Booth, L. R. (1984). *Algebra: Children's strategies and errors. A report of the Strategies and Errors in Secondary Mathematics Project.* Windsor: NFER-Nelson.

Cohen, L., & Manion, L. (1989). *Research methods in education* (3rd ed.). London: Routledge.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis.* Orlando: Academic Press.

MacGregor, M., & Stacey, K. (1993). Seeing a pattern and writing a rule. In I. Hirabayashi (Ed.), *PME XVII*, 1 (pp. 181-188). Tsukuba: Program Committee of the 17th PME Conference.

Mason, J., Graham, A., Pimm, D., & Gowar, N. (1985). *Routes to roots of algebra.* Milton Keynes: Open University Press.

Meira, L. L. (1990). Developing knowledge of functions through manipulation of a physical device. In G. Booker, P. Cobb, & T. N. de Mendicuti (Eds.), *PME 14*, 2 (pp. 101-108). Mexico: Program Committee of the 14th PME Conference.

Mènard, S. (1991). *Longitudinal research.* Newbury Park: Sage Publications.

Pegg, J. E., & Redden, E. (1990a). Procedures for, and experiences in, introducing algebra in New South Wales. *Mathematics Teacher*, *84*(5), 386-391.

Pegg, J. E., & Redden, E. (1990b). Student experiences in developing algebraic symbolismusing number patterns. Paper presented to the annual conference of the Mathematical Research Group of Australasia. Hobart, Australia.

Redden, E. (1994a) Providing an overview of qualitative data sets: Log-Linear modelling. *MERGA 17* (pp. 523-530). Lismore. Mathematics Education Research Group of Australasia.

Redden, E. (1994b) Alternative pathways in the transition from arithmetic thinking to algebraic thinking. In Pedro da Ponte, J. & Matos, J.F. (Eds.), *Proceedins of the Eighteenth Conference International Group for the Psychology of Mathematics Education.* Lisbon: International Group for the Psychology of Mathematics Education..

Redden, E. (1995) *A longitudinal Investigation into children's understanding of number patterns and the consequent emergence of algebraic concepts.* Unpublished Doctoral Dissertation. Armidale. University of New England.

Romberg, T. A. (Ed.). (1989). *Curriculum and evaluation standards for school mathematics.* Reston: National Council of Teachers of Mathematics.

Stacey, K. (1989). Finding and using patterns in linear generalising problems. *Educational Studies in Mathematics*, *20*, 147-164.

Ursini, S. (1990). Generalisation processes in elementary algebra: Interpretation and symbolisation. In G. Booker, P. Cobb, & T. N. de Mendicuti (Ed.), *P.M.E. 14*, 2 (pp. 149-156). Mexico: Program Committee of the 14th PME Conference.