DOCUMENT RESUME

ED 411 602		EA 028 646
AUTHOR	Allison, Patricia A.; Demaerschalk, Dawn; Allison, Derek J.	
TITLE	Thinking through an Administrative Problem: Processing Differences between Expert, Average, and True Novice Responses.	
PUB DATE	1996-04-00	
NOTE	32p.; Revised version of a paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996).	
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)	
EDRS PRICE	MF01/PC02 Plus Postage.	
DESCRIPTORS	*Classification; *Coding; Cognitive Processes; *Data Analysis; Foreign Countries; *Matrices; Principals; Problem Solving; Qualitative Research; Questionnaires; *Research Methodology	

ABSTRACT

The fundamental question in creating coding categories for an open-ended questionnaire is how to transform a complete transcript into manageable pieces of data. This paper describes the methodology involved in coding qualitative data derived from an evaluation of the Cognitive Approaches to School Leadership (CASL) program. The first task was to develop a suitable methodology for coding think-aloud data so as to analyze the ways in which respondents processed the problem case. The second major task was to apply the coding system and explore ways in which analysis might be conducted. The final model contained 4 functions and 10 operations, giving a 40-cell matrix. The data indicate that nonprincipals approach the problem on an entirely case-specific basis, and that principals are more likely to operate from a schema that allows them to define the problem faster and more comprehensively. The results provide an indication of a typical problem-processing pattern and its components. The four dominant behaviors used by principals in solving problems included conclusion/reasoning, problem/reasoning, detection/action, and attack/action. Five figures and eight tables are included. (Contains 14 references.) (LMI)

Thinking through an administrative problem: Processing differences between expert, average and true novice responses

Patricia A. Allison, Dawn Demaerschalk and Derek J. Allison
The University of Western Ontario

Revised version of a paper presented as part of the symposium
Clinical assessments of practical performance in school leadership: Findings from novice-expert studies of the elementary principalship.
Session #41.22 at the Annual Meeting of the
American Educational Research Association
New York City, April 1996

This research was supported by the Social Sciences and Humanities Research Council of Canada under grant #410-92-0329

CASL Project Paper #AERA96 4.2, revised August 1996.

Thinking Through an Administrative Problem: Processing differences between expert, average and true novice subjects

Patricia A. Allison, Dawn Demaerschalk & Derek J. Allison

This is one of three analytical papers prepared for a 1996 American Educational Research Association symposium reporting the work of the Cognitive Approaches to School Leadership (CASL) project. The analyses which follow deal only with the think-aloud responses to the case study and specifically the question of how subjects processed the problem presented to them. Another of the papers prepared for the symposium (#3.2) considered differences there are between the case content considered by participants—the topics they considered; this paper addresses the strategies and approaches used in problem processing; *how* they handled the task.

This analysis builds on earlier work. In our first principal problem processing study we found differences among groups which were difficult to explain by simply considering *what* subjects did, and this led us to feel that there might be value in checking out *how* they approached the problem. We already had some intuitive notions but clearly what we were looking for were statistical patterns grounded in data. An earlier influence also contributed to this wish to find a pattern in process. In a project which preceded our first PPP study (Nagy & Allison, 1988), an attempt was made to map the problem-solving processes used by subjects with reference to a generalized model of the topics and issues considered. In the absence of established methods, we invented several analytical techniques and applied them out with varying success. It was obvious that there was considerable potential in such an approach if the methodological hurdles could be cleared. Nagy (1991) has since refined the earlier techniques and we further extended his work (Allison & Allison, 1993).

The first major task at hand for the analysis reported here was to develop a suitable methodology for coding think-aloud data in such a way as to enable us to describe, compare and contrast the thinking processes which our various subjects used to process the problem case. We set ourselves the restriction of developing the coding methodology only after the data were fully collected, so that the system developed would be solidly grounded in those data. The second major task was to apply the coding system and begin to explore ways in which

CASL Project Paper #AERA96 4.2, revised August 1996.

analysis might be conducted. We devoted considerable time to the development of the coding system and to the careful coding of the data as a necessary foundation for future work.

Method

Weber (1985) describes the typical sequence followed in performing an analysis of transcribed protocols as beginning with the first step of deciding on the unit of analysis. The second step is to define some general categories for coding and test them out on a sample of text. Revisions are then made based on the results of this testing process. Several iterations of this testing and revising cycle may be necessary or desirable, depending on the nature of the data and of the research. Just as large units of analysis may be too broad for a particular task and tiny units too small, so coding categories can be so wide that distinctions are lost or so narrow that there are far too many categories to manage. The creation of coding categories, as Simon and Kaplan (1989) explain, is the application of the researcher's theoretical framework so as to make it possible to interpret the data. To a large extent what one chooses to look for in coding a protocol depends on the purpose of the study.

The question of how to break the protocols down into manageable units is discussed more fully in the symposium paper on topical attention analysis (#3.2). Briefly, the fundamental question is how to chunk a complete transcript into manageable pieces for analysis. Weber (1985) identifies six commonly used units of analysis: word, word sense, sentence, theme, paragraph and whole text. What is most appropriate depends largely on the purposes for which the analysis is intended. Individual words may sometimes be appropriate as the unit of analysis; searching for key words, however, presupposes that not everything which was said will be analyzed, and ignores context. At the same time there are problems with using paragraphs or sentences, especially since the protocol is usually a transcription from an audio tape and such divisions are often a result of transcriber decisions. Using either sentences or paragraphs as the unit of analysis is an attempt to capture coherent thoughts, and when dealing with transcriptions, in which neither sentences nor paragraphs are entirely reliable, it is, as Weber points out, quite acceptable to try to capture discrete thoughts as the unit of analysis. After some deliberation, we decided to use "thought units" [TUs] as our unit of analysis.

The first step in developing a coding system was to look for extant systems which might possibly be adopted or adapted, or at the very least provide a starting point. The early experience described above (that is, the lack of extant models) had changed only slightly in the intervening years, with a few notable exceptions. Beginning with our own earlier work (Nagy, 1991; Allison & Allison, 1993), and other recent publications (Leithwood & Stager, 1989; Leithwood & Steinbach, 1995), and following up on references and other leads, we were able to begin to form a general picture of what should be included. We found a few detailed explanations of models but only three coding systems with sufficient detail to distinguish the categories used.

1. Ericsson (1975) used a coding system which distinguished four types of statements, each representing internal information of importance to the subject, pulling these statements out of the transcripts but not coding all verbalizations. The four categories were *intentions*, *cognitions*, *planning* and *evaluations*.

2. Goor (1974) proposed a seven-category coding system for analyzing think-aloud protocols from problem solving tasks. His categories, which were designed to be complete and mutually exclusive, grouped sets of activities in terms of the type and status of information upon which they were based. The seven categories were as follows:
    1 Surveying given information;
    2 Generating new information or hypotheses;
    3 Developing or working on a hypothesis;
    4 Unsuccessful solution;
    5 Changing the conditions of the problem;
    6 Self-reference or self-criticism; and
    7 Silence.

3. Smith (1971) created a three point classification system in which each statement was given three scores, one for each of three categories: content (symbolic, figural, semantic, behavioural); operation (cognition, memory, production, evaluation); and product (units, classes, regulations, systems, transformations, implications).

The warning given with the last system was that although some information should be readily encodable, inferences about the content and operation categories would often be problematic and require attention to the context of preceding protocol segments. Simon and Kaplan (1989) make particular mention of the necessity to consider context in order to preserve potentially crucial information. As everyone who has ever spoken to a news reporter knows, statements taken out of context are often ambiguous. Take for example a statement such as "I need to try and get the big picture": if this statement stands entirely alone, with no connected thoughts around it, it could be a general orientation statement, or even potentially a metacognitive thought; it could be part of a sequence of reasoning through various approaches to the problem; it could be an expression of desperation and defeat; if, however, it is preceded by a series of well-specified actions each intended to collect particular information, it becomes an explanation of the preceding activities. The difference may seem small but in terms of understanding the subject's approach to finding a way across the problem space it could be quite significant: a general idea of "getting the big picture" is substantively different from systematically seeking well-described pieces of information in order to build that picture. This question of context emerged as a very influential factor in the development of our coding system.

The development of coding categories was a slow and re-iterative process. We began by taking the categories in the Leithwood model (Allison, 1996, pp. 532—533; Leithwood & Steinbach, 1995): *interpretation* and *goals*, which are subsumed under the broader process of *understanding*; *constraints* and *solution processes*, which are grouped as *solving processes*; and *values* and *mood*, which were understood as influencing the previous categories and forming the general category of *disposition*. We added the category of metacognition, as this is recognized as an important process in the relevant literature and has been cited as a point of contrast in previous novice-expert research. Since this type of analysis is still very new, we were careful to remain open to different possibilities, and thus several more categories were added as we progressed; some were deleted later during the test and revise cycle. In some cases we changed the labels on categories to better describe what belonged in the category; in other instances some single categories were divided and in others two or more were combined.

Following Weber's (1985) suggested sequence, we tested the various incarnations of the system on sample protocols. Members of the research team would independently code sample transcripts, making notes about difficulties encountered. These independent experiences were then discussed, modifications made to the system, and the process repeated. After numerous such re-iterations and re-definitions, we were able to more clearly identify some of the more serious omissions and confusions.

One serious confusion was with regard to the purpose of an operation, which brought us back to the consideration of context. Merely describing an operation as, say, an action, did not capture the purpose of that action, and the purpose could be significant. For example, the specifics of the action might be the calling of a staff meeting. The purpose of the staff meeting might be completely exploratory, to find out more about how the staff feels; it might be the first action in a partially-formed strategy to be guided by what happens; or it might be a forum in which the principal dictates to the staff what the new rules are going to be, having already made up her mind. Only by taking the action in context could the purpose be clear, but without somehow recording the purpose, we were obviously missing a vitally important part of the single operation as well as missing the links between thought units. Indeed, we encountered a similar problem in our first PPP study, when we tried to collapse protocols into action sequences for analysis. While we found that the number of actions proposed in a transcript was a useful discriminator (Allison & Allison, 1991), we eventually abandoned this approach when it became clear that by themselves actions were hopelessly ambiguous without the surrounding context. We concluded, therefore, that we needed a two tier coding system, which assigned to each TU (1) a description of the operation involved and (2) a general purpose to be served by the operation.

Our final model contained four functions and ten operations, giving a forty-cell matrix. The first two functions (detection and conclusion) can be conceptualized as interpretation stages concerned with building the problem space. The third (problem-working) can be understood as seeking to engage an appropriate schema and testing possible paths across the problem space, and the fourth (attack) as planning. The four functions, represented in the horizontal rows of the coding sheet were defined as follows.

1. **The Detection function [DF]**. All TUs in which the subject appears to be engaging in the process of looking for cues and clues about the nature of the problem.

2. **The Conclusion function [CF]**. All TUs in which the subject appears to have reached a conclusion about the problem or some element of the problem. The subject may be recognizing elements of the problem as matching previous experience. In the case of experts, we would expect that this matching would facilitate the retrieval of an appropriate schema and the construction of the problem space.

3. **The Problem-working function [PF]**. All TUs in which the subject appears to be attempting to organize information and/or reason through parts of the problem. The subject may be thinking through possibilities, considering tactics or generally musing on the problem or some aspect of the problem. By definition, any TUs which involved searching for further information or diagnosing elements of the situation could not be included. In the course of working with these categories, we came to use the term "pondering" as a way of understanding this function.

4. **The Attack function [AF]**. All TUs which are related to specified behaviours, apparently intended to tackle the situation. By definition, this function could not include statements which involved seeking information, nor could it include statements which did not have a clear plan of attack. It is quite

possible—indeed it is hard to imagine how it could be otherwise—that the subject did have a detection purpose in mind, or was acting on the basis of a conclusion or diagnosis reached, but did not express this. Given the nature of the think-aloud exercise, there are really only two possibilities: either the subject is not following the instructions very well and missing out some thought processes; or s/he is tackling the problem on the basis of thinking patterns which are retrieved directly from memory. In terms of coding extant data there is little that can or should be done about this; it would certainly be inadvisable to go any further with inferences than we already do. In terms of analysis, however, we may wish to pay particular attention to this feature.

The ten operations, represented in the vertical columns of the coding sheet, were defined as follows.

1. **Action/goal [A/G]**. Any thought unit containing a specific action, a statement about something which would be done, and/or the intended outcome. Originally these were two separate operations but it became evident that the distinction between an action and a goal or outcome in a think-aloud transcript is often quite difficult to make: a subject might make a statement which includes a goal state but at the same time implies a distinct action. For example, one subject wanted to "talk [with] the staff and with Miss MacDonald, setting some common expectations for classroom management in the library." The action, speaking with the staff and the librarian, is not complete without the addition of the goal, which in this case defines what the discussion will be about.

2. **Values [V]**. Any statement which clearly articulates, following Kluckhohn (1951), a conception of the desirable. This is roughly comparable to Leithwood's original *Principles* and later *Values*, and includes both values held regarding the school, (e.g., "I think prep. time is really vital") and those regarding how to proceed with problem solving (e.g., "I think it's important to go slowly here"[1]).

3. **Rhetorical Question [RQ]**. Any question which is asked, not as a request to the researcher, but as part of processing the problem. We found that there were many instances in which a subject asked a question as part of working through the problem. At one point we attempted to code these TUs as detection, making this an operation, but soon found that such a category became overloaded with vastly different types of operations. Questions were sometimes thrown out as part of a detection train of thought, but were just as often voiced when pondering over the problem; they were sometimes mere interjections, apparently

[1] This is a good example of the importance of context. This statement, in isolation, might support several interpretations. It could, for example, be taken as metacognition if it were understood that the subject was thinking about his/her own thinking processes. In fact, this subject was saying that it is important to go slowly and carefully when working with a group of people to change their behaviours.

unconnected to any train of thought. Coding them as a separate operation and categorizing them according to function seemed to work best.

4. **Metacognition [MC].** A thought unit in which the subject appears to be actively monitoring his/her own thought processes or reflecting on his/her own approach to the problem.

5. **Constraint [C].** Any statement regarding an anticipated barrier or obstacle seen as impeding a solution to the problem. This corresponds exactly to Leithwood's similarly named category.

6. **Example [E].** Any statement referring to, describing or summarizing a previously experienced situation, regardless of whether it relates to the case or not.

7. **Case Material Verbatim [CMV].** Any unit in which a piece of the original case material is being re-read directly from the printed copy. Because we wished to code every thought unit, this category was necessary to accommodate those subjects who went back to the case and read pieces of it out loud. We could have chosen to merely ignore them, but we felt that this could not be justified, especially if there might be a pattern in this behaviour.

8. **Case Material Processed [CMP].** Statements in which material from the case is paraphrased or manipulated in some way. Some subjects referred back to specific elements in the case in a slightly different way, having in some way processed the information given. There seemed to be an important difference between re-reading passages of the case and interpreting what was given, and between reasoning through the case in a general sort of way and referring to specific information given in the case.

9. **Self-referent [SR].** Any statement in which the subject refers to him/herself. These statements usually take the form of a personalization; that is the subject talks about how s/he would feel in a similar circumstance. Some subjects punctuated their think-aloud with comments about their own personalities (e.g., "I'm really a people person").

10. **Reasoning [PSR].** Statements in which the subject is explaining his/her analysis of the problem or aspects of the problem, including providing the reasons for a particular course of action or for a diagnosis or decision made. By definition, this does not include statements which could be coded in any of the other nine columns; that is, statements which included, for example, a rhetorical question or an example. This category was originally three different categories: diagnosis, reasoning and problem-solving. This proved to be unworkable, however. Diagnosis, which was far too large and diverse a category, became, in essence, the Conclusion function. Problem-solving, which could include some discrete operations such as

providing examples, expressing values or processing information given, became the Problem-working function.

Our intention was to code every thought unit somewhere, and to make this possible we needed a place to deposit leftovers. A final category of (11) **Superfluous [S]** was therefore created to provide a space for all those thought units which could best be described as off task. This included expletives, incomplete thoughts and meaningless verbalizations of various types. We were very strict in the use of this category and tried to make sure that we threw away only those TUs which were truly superfluous and completely unsalvageable. The addition of this final category made it possible to code every thought unit in each transcript.

	A/G	V	RQ	MC	C	E	CMV	CMP	SR	PSR
DF										
CF										
PF										
AF										
S										

Coding Key

Functions	Operations	
DF = Detection function	A/G = Action/goal	E = Example
CF = Conclusion function	V = Value	CMV = Case Material Verbatim
PF = Problem function	RQ = Rhetorical Question	CMP = Case Material Processed
AF = Attack function	MC = Metacognition	SR = Self referent
(S = Superfluous)	C = Constraint	PSR = Reasoning

Figure 1
Example of blank coding matrix with coding key

The final system was mapped onto a coding matrix, which was then used to record the codings for each transcript. Figure 1 displays a blank coding matrix, reduced in size. The placement of the operations and functions on the matrix was totally arbitrary, having evolved along with changes in the categories. In retrospect there might have been some point in deliberating over the final placement of rows and columns, and we may revisit this issue at some future date. Figures 2a and 2b provide examples of thought units coded in each of the cells (except, of

	A/G	V	RQ	MC	C
DF	I would use [the librarian] as a valuable source of information to move ahead with this	I think that making the library a more attractive place] would be very important.	Is that the teacher's fault or the librarian's fault?	I'm trying to keep the full perspective here.	But the principal's presence upsets things.
CF	So it should be decided that teachers should use this as a reward and not just dumping problem students into the library.	I mean, it's true. She shouldn't expect the librarian the baby-sit her students.			I don't know anything about this school.
PF	So perhaps inviting other staff members on to a Partners in Action committee, who she might feel comfortable working with.	Then again, teachers aren't supposed to be making complaints about other teachers to the principal.	Why would she submit after the deadline?	I have to separate these two people from my decision-making here.	I really couldn't tell you how this could be solved without pitting one side against the other.
AF	I'd transfer Miss MacDonald immediately.	Well, I guess it's always advisable to proceed with caution.	Is this something that's probably going to take me all year to deal with?		But not to the point where we're dealing with bitter battles between individuals.
S	I'm not so sure.....; I'm sort of rambling here.....				

Figure 2a
Examples of coded thought units

course, those in which no TUs were coded). These quotations are for illustrative purposes only; in actual coding we recorded only the TU numbers in the cells, as illustrated in Figures 3a and 3b which appear later in the paper.

It should be noted here that all cases were assigned random numbers so that the coders (the first and second authors) had no overt indication of the identity of the subjects. Furthermore, neither coder had been involved in collecting the data and would not, therefore find any of the transcripts familiar. The coding process was thus conducted under largely blind conditions. We coded transcripts by recording the TU numbers in the appropriate cell of the coding matrix, rather than just tallying, thus making it possible to directly check inter-rater reliability and to pull out descriptive quotations in subsequent analysis. We strongly recommend this approach to others attempting similar analyses; being able to locate coded thought units on demand vastly enriches the analysis. Our actual process for coding was to have the two raters code the transcripts independently and then meet to compare results. Inter-rater reliability, on comparison, was in the region of .75 to .80. All discrepancies were

	E	CMV	CMP	SR	PSR
DF	But in our particular board, where we have a librarian half time in my school so it's very difficult to schedule in and have your classroom teacher book PiA.	Staff thought nothing was wrong with the library and how it was operating.	Yet down here the teachers on staff seem to feel that the library is operating fine.	Myself, don't have a clue what Partners in Action is.	Try to get the big picture here.
CF	I think that the resource centre, at least in my home school, was not utilized to its full extent.	She's been here for fifteen years.	Pat is the new principal and he has this problem with the librarian.	I'd be nervous and restless too probably.	It's quite a responsibility for the library teacher to take on if several groups are descending on her at once.
PF	This is a very real problem for me right now because I have a librarian who's been there for 18 years and is a highly skilled teacher but the library is very much her domain.	It seems that his first problem was that the library quarters were small and there appeared to be few modern or new books likely to meet the needs of older students.	Miss MacDonald is the school librarian; she knows how the library works and she's also a teacher.	I wouldn't want my teacher hanging over my shoulder when I was in the library.	Probably because the previous principal was ill, the staff has maybe had to operate on their own quite a bit.
AF	So we were able to recruit some teachers to make the school turn a little bit in a different direction.		Sounds like there are some staff people who would like to do Partners in Action activities.		And then you try to work on it a little step at a time.
S	I think I'm done.				

Figure 2b

Examples of coded thought units (continued)

then debated to consensus and a final, agreed upon, coding for every thought unit in each case was recorded. This was obviously a very slow process, involving as it did three passes through every transcript. Short-cutting at this stage would, we felt, have been a mistake; since the process of coding transcribed think-aloud protocols is inherently one of inference and conjecture, it is vital that extra care be taken to ensure that each coding decision is made with scrupulous attention to contextual factors, to wording, and to meaning. Whether the extra care results in more accuracy is probably open to debate given the nature of the exercise but at the very least we, as researchers, need to know that we have not been slapdash in the interpretation of our data.

Findings

General Patterns

When all cases were coded, we created a simple database for preliminary quantitative analysis, in which we recorded the raw frequencies of thought units coded in each cell of the coding matrix. Future analyses may take us in different directions; we may, for example, try coding each thought unit consecutively so that it is possible to preserve the sequence of thoughts. Initially, however, since there was such a wide discrepancy in the total number of thought units per transcript, this method of recording data was rejected in favour of simple frequencies for each cell in the matrix. The variation of the total number of thought units still presented some problems: raw frequencies could not be compared directly since they would give no indication of the relative proportions of thoughts devoted to different components. We decided, therefore, that it would be more useful for direct comparisons if we converted the number in each cell into a percentage of the total number of thought units, thus providing a measure of the relative amounts of thought devoted to the different functions and operations. The raw frequencies were retained for some analyses, and in addition new variables were created which converted each raw frequency into a percentage of the total TUs for the transcript as a whole.

Table 1 shows the distribution of the coded thought units for all subjects. The first column provides a measure of overall *usage*, that is, the percentage of the total sample for whom thought units were recorded in this particular cell. Thus, for example, in the first line of the second row from the bottom of the table, we see that 38% of all subjects had TUs coded in the Problem/Action cell, and 75% in the Conclusion/Reasoning cell. The second column is the mean usage (by which the cells are rank ordered), and the third column is the range of usage. Thus, for example, 53% of the subjects had TUs coded under Conclusion/Case Material Processed; the mean usage was 3.7% of the total number of TUs per transcript; the minimum usage was none at all and the maximum usage was 18.8% of the total thought units in a transcript.

The coding matrix contains 40 cells in total. A perfectly even distribution would place approximately 2.5% of the total TUs in each cell. This being so, any cell with more than 2.5% could be considered to be a more common operation and any cell with less than 2.5% a less common operation. For convenience, we have divided the distribution into five categories which correspond to breaks in the frequency distributions: those cells with no

Table 1
Rank ordered percentage usage, means and ranges for coded TUs

Usage	Mean	Range	Function/Operation
0	0	0	Conclusion - Rhetorical Question
0	0	0	Conclusion - Metacognitive
0	0	0	Attack - Metacognitive
0	0	0	Attack - Case Material Verbatim
0	0	0	Attack - Self referent
2%	0.03	0 - 1.6	Attack - Case Material Processed
4%	0.06	0 - 1.9	Conclusion - Case Material Verbatim
2%	0.06	0 - 3.0	Attack - Constraint
2%	0.08	0 - 4.0	Detection - Values
4%	0.08	0 - 2.4	Detection - Example
4%	0.09	0 - 3.7	Conclusion - Constraint
4%	0.11	0 - 3.4	Detection - Metacognitive
4%	0.11	0 - 3.8	Detection - Constraint
9%	0.21	0 - 4.8	Detection - Case Material Verbatim
9%	0.21	0 - 3.0	Detection - Self referent
4%	0.22	0 - 7.4	Problem - Constraint
7%	0.23	0 - 5.6	Attack - Rhetorical Question
6%	0.26	0 - 6.3	Problem - Metacognitive
7%	0.26	0 - 6.1	Conclusion - Self referent
7%	0.27	0 - 6.1	Problem - Self referent
7%	0.34	0 - 11.1	Conclusion - Action
11%	0.36	0 - 6.5	Problem - Value
7%	0.39	0 - 9.5	Problem - Rhetorical Question
4%	0.53	0 - 18.5	Attack - Example
13%	0.55	0 - 6.4	Attack - Value
2%	0.63	0 - 33.3	Conclusion - Example
11%	0.75	0 - 16.0	Conclusion - Value
9%	0.84	0 - 24.1	Problem - Case Material Verbatim
19%	1.2	0 - 24.4	Problem - Example
24%	1.5	0 - 17.5	Detection - Case Material Processed
26%	1.8	0 - 21.5	Detection - Rhetorical Question
38%	2.6	0 - 20.6	Problem - Action
45%	2.9	0 - 32.5	Problem - Case Material Processed
53%	3.7	0 - 18.8	Conclusion - Case Material Processed
75%	6.2	0 - 17.7	Attack - Reasoning
77%	6.5	0 - 22.9	Detection - Reasoning
75%	11.9	0 - 55.6	Conclusion - Reasoning
89%	12.5	0 - 57.1	Attack - Action
87%	12.7	0 - 51.4	Detection - Action
87%	15.2	0 - 37.1	Problem - Reasoning

coded thought units at all; those with a mean usage of less than 1%, which comprises the largest group; those with a mean usage of more than 1% but less than 2.5%; those with a mean usage of more than 2.5%; and those with more than 10%. The latter two categories, which include a total of nine cells, must be regarded as the most common operations.

There were five cells in which no thought units were coded at all, these being Rhetorical Question and Metacognition under the Conclusion function and Self-referent, Metacognition and Case Material Verbatim under the Attack function. It is quite logical that there would be nothing which could reasonably be coded as a Rhetorical Question conveying a conclusion, since these would seem to be mutually exclusive. The remaining combinations were, it would seem, simply extremely rare and disappeared in some functions altogether. None of these empty cells, therefore, caused any concern.

There were a further 23 cells in which the mean percentage of coded thought units fell below 1%. The range of percentages in these cells, as shown in Table 1, is highly variable. In every case, the lowest frequency was 0, and in many instances the majority of cases were 0 or very low. For example, under Attack/Constraint, where only one subject had any TUs coded, the actual percentage of the overall transcript represented by those thought units was the only 3%. In others, a single subject may have accounted for the majority of the recorded instances, thus skewing the mean considerably. For example, under Conclusion/Example, only one subject had TUs coded, but that subject devoted a full third of his/her thought units to describing examples which supported conclusions reached, which was a unique pattern. In contrast, under Conclusion/Value, 11% of the subjects (i.e., 6) had TUs coded but the mean usage is only 0.75%, indicating that even when this particular combination occurred it was not much used.

Three cells produced mean percentages between 1% and 2.5%: Detection/Case Material Processed; Problem/Example; and Detection/Rhetorical Question. Five cells produced mean usages above 2.5%, making them quite significant: Problem/Action; Problem/Case Material Processed; Conclusion/Case Material Processed; Detection/Reasoning; and Attack/Reasoning.

A further four cells produced mean percentages above 10%, which given the size of the matrix must be considered high. Two of these cells are Action cells, one under the Detection function and one under the Attack function. The remaining two are Reasoning cells, one under the Conclusion function and one under the Problem

Table 2
Rank ordered percentage usage, means and ranges for coded functions and operations

Usage	Mean	Range	Coding category
			Horizontal: Functions
91%	24.3	0 - 57.5	Problem
96%	23.3	0 - 88.6	Detection
94%	20.1	0 - 70.0	Attack
87%	17.9	0 - 100	Conclusion
			Vertical: Operations
9%	0.38	0 - 6.3	Metacognition
13%	0.49	0 - 7.4	Constraints
17%	0.74	0 - 15.2	Self-referent
21%	1.1	0 - 24.1	Case Material Verbatim
30%	1.7	0 - 16.0	Values
30%	2.4	0 - 21.1	Rhetorical Question
26%	2.5	0 - 33.3	Example
83%	8.2	0 - 32.5	Case Material Processed
100%	28.1	8.6 - 78.6	Action
100%	39.9	7.1 - 71.4	Reasoning
94%	14.5	0 - 42.4	Superfluous

function. These, then, are the most common operations: reasoning about the problem; reasoning to support conclusions drawn; specifying actions to seek missing information; and taking direct action. The inter-relationship of these components will be discussed at length later.

When the vertical (Operations) and horizontal (Functions) components of the coding matrix are totalled, as shown in Table 2, a more distinct pattern appears. On the vertical scales, 3 columns were used in coding less than 20% of all cases and provided a mean usage of less than 1%: Metacognition, Constraints, and Self referent, from which we can conclude that these are rare operations. Four more columns were used in coding between 20% and 30% of the subjects but provided a very low mean usage, the highest being 2.5%: Case Material Verbatim, Values, Rhetorical Question, and Example. The range in percentage usages in these operations varies from 16.0% to 33.3%. Following these relatively low usage operations, there is a dramatic increase in usage. Some thought units were coded as Case Material Processed for 83% of all subjects; the highest individual usage was 32.5%, which is quite high, but the mean usage was 8.2%, which suggests that although the greater majority of subjects were coded as having used this operation, not many of them used it a great deal.

Every subject was coded as performing some Actions and every subject was coded as using Reasoning. The minimum usage recorded for each of these operations was slightly less than 10%, the maximum over 70% and the means 28.1% and 39.9% respectively, indicating not just ubiquitous usage but also consistently intense usage. Going back to Table 1, it is evident that the most common types of Action thought units were coded under either the Detection or the Attack function, while the most common types of Reasoning units were coded under either the Conclusion or the Problem-working function.

On the horizontal scales, the four functions were very evenly distributed. In every function at least 87% of the subjects had some thought units coded. The lowest percentage usage in every case was 0 but the highs ranged from 57.5% to 100%—one subject's thought units being coded exclusively in the Conclusion function,[2] but this was quite anomalous, as the mean usage of 17.9% indicates. The remaining mean percentages range from 20.1% to 24.3%, indicating a fairly even distribution of usage in most cases.

The Superfluous category was used in coding 94% of the transcripts, with a mean usage of 13.8%. One subject was coded as making such statements 42.4% of the time, which is obviously anomalous and served to drive up the mean.

In summary, the typical pattern of response to the case would have been coded as dividing attention fairly evenly across the four functions, making a few remarks diagnosing elements of the problem, a number of statements indicating what further information was going to be necessary in working on the problem, a number of statements regarding intended ways of attacking the problem, and a number of statements essentially pondering over the problem and its elements. The typical response would probably not have included any statements which could have been coded as consciously attending to thinking patterns (Metacognition), or any which considered constraints which might inhibit working on the problem. It would probably not include any self-referenced statements and the subject was unlikely to spend time re-reading the case. The typical response might have included a single value statement related to the problem; it might have included a rhetorical question, most likely as part of the quest to identify what further information was necessary; it might have included a reference to an example from experience which seemed to the subject to be related. It most probably would have included some

[2] The subject whose thought units were all coded in the Conclusion function was a true naive and, not surprisingly, her response was rated extremely low by the jury.

considerations of how the information presented in the case could be interpreted (Case Material Processed). It would definitely have included a substantial number of statements coded as intended specific actions. most of them aimed at acquiring further information. or beginning to attack the situation. It would have included a substantial number of statements coded as reasoning through the problem. most of them either generally considering alternatives or explaining diagnoses reached.

This typical pattern was derived from pooling the data of all subjects, males and females, those with experience in the principalship and those without, those with aspirations in the education system and principals in the High. Medium and Low imputed expertise categories. The next step was to investigate how the pattern varied when some of these distinctions were taken into account. Analyses were conducted which compared subjects according to the following distinctions: (1) gender; (2) principals versus non-principals; (3) principals only, grouped according to imputed expertise. A full explanation of how the expertise groupings were determined is included in the design paper prepared for this symposium. The immediate problem for us, of course, was to determine what kind of analysis would be appropriate. In any attempt to reduce language-based data to numerical data for quantitative analysis there is an unavoidable loss of meaning. The quantitative analyses we wished to perform were intended to identify patterns in the data which could then be further analyzed in whatever way seemed appropriate. Since we had no predictions regarding the patterns which might emerge, it was important to stay as close to the original data as possible and avoid further loss of meaning by grouping items together on the basis of numerical logic only. We chose. therefore. to perform very simple analyses, consisting primarily of simple frequencies and t-tests. The results presented here are deliberately rough. since we do not wish to claim statistical significances which may be meaningless in linguistic terms. Our intent was to search for statistical and numerical patterns which might be interpretable in some way which would lead us back to the data again. In future work. we hope to be able to achieve greater statistical clarity but we are not willing to embark on more rigorous statistical measures until we are much clearer about what we are working with.

Comparisons

Tables 3a and 3b present all of the simple comparative data across the identified groups. The three horizontal divisions identify functions, operations and cells which seem to be particularly important. The first set

Table 3a
Percentage usage and means for Gender and participant Groups

Coded Functions & Operations	Male N=26		Female N=28		Principals N=30		Novice (B.Ed). N=10		Naive N=14	
Functions:										
Detection	*96%*	23.8	96%	22.8	**100%**	**26.9**	**100%**	**15.6**	86%	21.1
Conclusion	*86%*	15.7	89%	29.8	**83%**	**12.0**	100%	18.6	**86%**	**29.4**
Problem	*88%*	24.8	93%	19.8	93%	23.7	**100%**	**33.9**	**79%**	**18.7**
Action	*96%*	21.5	93%	18.8	93%	22.7	100%	16.9	93%	16.8
Operations:										
Actions	100%	30.1	100%	26.4	**100%**	**31.7**	**100%**	**21.9**	100%	25.2
Values	28%	1.9	32%	1.5	28%	1.5	40%	1.2	29%	2.7
R Questions	24%	1.9	36%	2.8	31%	2.0	**10%**	**.29**	**43%**	**4.6**
Metacognition	16%	.65	4%	.13	14%	.47	0	0	7%	.45
Constraints	12%	.36	14%	.60	14%	.35	10%	.43	14%	.80
Examples	36%	3.4	18%	1.7	**41%**	**3.3**	**10%**	**.21**	7%	2.4
Case Verbatim	16%	1.2	25%	1.0	10%	1.0	30%	.69	36%	1.6
Case Processed	72%	6.9	93%	9.3	83%	7.3	100%	12.5	71%	7.1
Self-referent	16%	.55	18%	.91	17%	.52	0	0	29%	1.7
Reasoning	100%	38.6	100%	40.9	**100%**	**37.2**	**100%**	**47.9**	100%	39.6
Cells:										
C - CM Processed					52%	2.9	80%	6.5	-	-
C - Reasoning					72%	8.5	-	-	79%	19.5
P - Reasoning					90%	6.5	100%	22.2	71%	11.0
A - Reasoning							90%	8.4	64%	3.8
D - Action										
P - Example										

of columns in Table 3a provides the comparisons for males and females. The second set of columns compares the three categories of principals, novices (student teachers) and naives. The first set in Table 3b collapses these to principals and non-principals. The final set of three columns in Table 3b compares principals in the Low, Medium and High imputed expertise categories. In each column, the *italicized numbers* are usage percentages (i.e., what proportion of the defined group were coded as using this operation), and the normal print number is the mean usage for that group. Significant comparisons (i.e., those with a probability of .01 or less on a t-test of group means) are **bolded** with the exception of the bottom row (Cells) of each Table, in which all comparisons shown were statistically significant.

Gender. When subjects were divided by gender and the coded results compared, no differences between the groups were found. In fact, not only are there no significant differences, there is a high level of correspondence between the two groups, despite the fact that there is a preponderance of males in the principal groups and a preponderance of females in the non-principal group. There are, however, a few interesting differences in overall patterns. Female subjects were coded less often as making metacognition statements, or using examples, but were coded more often as interpreting elements of the case. Other than these variations in what are essentially minor instances, which may in any case be attributable to role rather than gender, males and females do not appear to have processed the presented problem in different ways.

Principals vs. Others. There were some distinct differences in the way principals and non-principals were coded in their use of the functions and operations, although sometimes the differences occurred between principals and a particular category of non-principals. It must be remembered in these comparisons that the number of subjects in each of the non-principal categories is quite small (i.e., 10 student teacher novices and 14 pre-novice naives).

1. Principals were coded significantly more often in detection activities than were student teachers, but not significantly more than those in the naive category. The implication seems to be that principals were more aware of what extra information they needed than were neophytes to the education system, but that competent and intelligent outsiders were also more aware of needing much more information. The question which arises immediately here is whether there is a substantive difference between the type of detection carried out by principals and that carried out by outsiders. In order to test this possibility (albeit rather superficially), we drew samples from each group in this, and two other apparently pertinent areas, for a direct comparison of the text. Table 4a contains the examples for comparison. It appears that there is indeed a qualitative difference in the type of information sought in detection. Non-principals appear to be thinking in terms of specific components to the problem, such as the book budget, and yet do not seem to

Table 4a
Examples of TUs coded in the Detection Function for Principals & Non-Principals

Principals	Non-Principals
• Look at the historical background to see if there was something amiss there. • Let's identify what we want to get as a staff out of the library. • See what we're doing well in each division; see the areas where we could look to improve while maintaining those things that we do very well. • I would ask Miss MacDonald to explain to me how she enacted Partners in Action in the library.	• I would want to talk to the students about it as well. • Check the book budget. • The principal has to get to know those teachers. • I would ask her honestly what she sees as the problem with the library.

Table 3b
Percentage usage and means for Principals and Non-principals and for imputed expertise categories for Principals

	Participant Group				Imputed Expertise					
Coded Functions & Operations	Principals N=30		Non-principals N=24		Low N=5		Medium N=19		High N=6	
Functions:										
Detection	**100%**	**26.9**	92%	18.8	**100%**	**14.5**	**100%**	**31.7**	100%	22.9
Conclusion	**83%**	**12.0**	**92%**	**24.9**	80%	11.2	83%	12.3	83%	11.9
Problem	93%	23.7	87%	25.0	100%	30.9	89%	22.8	100%	20.2
Action	93%	22.7	96%	16.8	100%	35.0	89%	18.3	100%	25.9
Operations:										
Actions	**100%**	**31.7**	**100%**	**23.8**	100%	31.3	100%	29.9	100%	37.4
Values	28%	1.5	33%	2.1	40%	1.7	22%	1.4	33%	1.3
R Questions	31%	2.0	29%	2.8	20%	1.4	33%	2.3	33%	1.7
Metacognition	14%	.47	4%	.26	0	0	22%	.76	0	0
Constraints	14%	.35	12%	.65	0	0	17%	.49	17%	.24
Examples	**41%**	**3.3**	8%	1.5	**60%**	**9.8**	**39%**	**1.7**	**33%**	**2.7**
Case Verbatim	10%	1.0	33%	1.2	0	0	11%	1.6	17%	.17
Case Processed	83%	7.3	83%	9.3	80%	5.7	89%	8.1	67%	6.2
Self-referent	17%	.52	17%	.99	0	0	17%	.65	33%	.57
Reasoning	**100%**	**37.2**	100%	43.1	80%	41.5	100%	38.2	83%	30.7
Cells:										
C- CM Processed	-	-	-	-	-	-	-	-		
C - Reasoning	72%	8.5	79%	16.2	-	-	-	-		
P - Reasoning	-	-	-	-	-	-	-	-		
A - Reasoning	-	-	-	-	-	-	-	-		
D - Action	93%	15.3	79%	9.5	100%	6.9	100%	17.7		
P - Example	10%	.21	4%	.09						

have much idea of what they are looking for, as for example in wishing to talk to students to find out what they think. Principals appear to have a clearer idea of what they are looking for, such as finding out exactly how Partners in Action is being implemented, while at the same time thinking in broader terms, such as the overall goals of the library and how the rest of the program in the school is operating.

2. Principals were coded significantly less often in conclusion activities than were non-principals in general, and this appears to be more so in comparison with true naives than with student teachers. This is especially so, as the bottom panel of Table 3b shows, in relation to explanations supporting conclusions or diagnoses (C - Reasoning). This seems to imply that principals are less likely to move to precipitate conclusions and diagnoses, presumably because they have a better understanding of the possible complexities of the situation, and to spend less time explaining those conclusions and diagnoses which they do make. Again, we undertook a qualitative check of the types of conclusions drawn by subjects in

the two groups, with Table 4b presenting comparative examples of thought units coded for each group under the Conclusion function. The non-principals seem to be diagnosing small elements of the situation, such as concluding that the library is out of date or that the problem is fundamentally just a personality clash. Principals, on the other hand, seem more likely to be thinking in much broader terms about the problem and how they will have to work on it, rather than picking out specific elements from the overall context and drawing conclusions which would almost certainly restrict their subsequent thinking about the problem.

Table 4b
Examples of TUs coded in the Conclusion Function for Principals & Non-Principals

Principals	Non-Principals
• I'm assuming that she is going to be responsive to my support and my willingness to indicate that this program is a vital program in the school. • It is definitely a year or two year process before we're going to begin to see some results. • The library problem seems to be a symptom of a much larger problem.	• I think the first problem is that the library is out of date. • It seems to be some type of personal thing. • Perhaps because of the length of time she's been there, she's become too accustomed to the way things have been going on in the past. • Miss MacDonald doesn't seem too happy being a librarian.

3. Student teachers were coded much more often in Problem thoughts than were principals, but naives were coded much less often, cancelling out the difference when the two groups are combined as non-principals. This is particularly evident in the Problem/Reasoning cell, where the mean usage for student teachers was 22.2% and for principals was only 6.5%. Again, the difference may be qualitative: that is, the relationship between student teacher thinking and principal thinking may be a true comparison of similar operations, reflecting the fact that as total novices the student teachers have more difficulty matching the case to their less well developed schemata, whereas the naives are not a true comparison because they are working from a completely different perspective. A systematic qualitative analysis of the relevant passages of the transcripts would be necessary to explore this possibility.

4. Although the difference in mean usage is not significant, principals were coded more often in attack strategies than were non-principals. If the actual content of the attack strategies were to be compared, this difference might be significant, and probably meaningful.

5. Principals were coded in actions significantly more often than were non-principals, especially student teachers, and most particularly in detection actions. This seems to imply that principals have more of an idea of what they need to do and find out than do non-principals, indicating that they are drawing on and attempting to apply richer, more comprehensive problem relevant schemata.

6. Naives were coded as asking rhetorical questions far more often than anyone else. This is probably a reflection of their relative lack of contextual knowledge, and/or difficulties in knowing how to get to grips with the situation. A particularly revealing TU in this connection is the following utterance by a naive: "What is the purpose of some of those roles?".

7. Principals were more often coded as using examples from experience than were non-principals, especially in relation to working on the problem. This is intuitively sensible: principals are obviously going to have more relevant examples from which to choose than are non-principals. It is also consistent with our overall conceptual framework to expect principals to be more likely to access relevant memories indexed through larger, richer schemata. Again, we would expect the examples used by principals to be more immediately relevant than those used by non-principals, as is illustrated in the two examples shown in Table 4c.

Table 4c
Examples of TUs coded as Examples for Principals & Non-Principals

Principal	Non-Principal
• This is a very real problem for me right now because I have a librarian who's been there for 18 years.	• The resource centre, at least in my home town, was not utilized to its full extent, and a lot of the teachers did pass off their students on the librarian in there. They just kind of guided them into the library and they disappeared.

8. Principals were coded less often in general reasoning than were student teachers[3], but not significantly less often than were naives, or than non-principals when the two were grouped together. We might speculate that the type of reasoning engaged in by naives is qualitatively different from that which principals engage in although it is roughly comparable quantitatively; principals might be finding it easier to access and test the components of relevant schemata while naives have fewer ideas of where to start. Student teachers, who have some socialization to the context, might be spending more time in reasoning because it is more difficult for them to access and test problem relevant components in their relatively impoverished schemata, which would nonetheless be expected to be more developed than those of naives. A careful analysis of the content of statements coded in this way would be necessary before such a conclusion could be reached.

9. Student teachers were coded as reviewing what they interpreted from the case material more often than were principals. This may corroborate the suggestion made in the previous note, that student teachers

[3] It is worth reminding ourselves here that the numbers we are working with are percentages, which means that there is a finite total possible. Any increase in one type of activity will of necessity be at the expense of a different type.

have more difficulty mapping the problem onto their partially-developed schema than do principals. but they are ahead of naives who, in effect. have very sketchy problem-relevant schemata on which to draw.

Summary

In summary, we can draw the following conclusions:

- Principals appear to have a better idea of what extra information they need and how to get it.
- Principals appear to be more aware of the possible complexities and ramifications of the situation.
- Principals move more often into both actions and attack strategies, spending less time reasoning through the problem.
- Principals are more likely to draw on domain-relevant examples from their own experience.
- Student teachers are definitely novices in relation to principals, but they may be different from the naives in significant ways.

These factors taken together seem to support the idea that whereas non-principals are approaching the case on an entirely case-specific basis, principals are more likely to be operating from a schema which has already provided places in which to store the case information, and from which they are able to proceed. In short, principals defined the problem space faster and more comprehensively and were able to identify approaches for crossing the problem space faster.

A simple correlation of the four functions seems to corroborate these findings. Among principals, there is a significant negative correlation between the Attack function data and the Detection function data; that is to say the principals, in general, spend extra time in action at the expense of detection - and vice versa. This could mean that, in conformity with our theoretical framework, principals who are able to readily access relevant schematic knowledge don't need to spend as much time in detection because they already know what to expect and what to do; on the other hand it could mean that principals with well developed schema have a better idea about what extra information they need. In contrast, among non-principals there is a significant negative correlation between Conclusion and Detection data; that is to say the novice and naive subjects tended to spend extra time in conclusion to the detriment of detection and vice versa.

Having determined that there are considerable differences in the ways in which principals and non-principals approach processing the problem case, the next question is how the approach of principals differed across the imputed expertise categories. Using the measures discussed in the design paper, we divided principals into three categories of Low, Medium and High imputed expertise. It is important to remember that all of the principals in the High category had, by definition, received high ratings from the peer jury on the quality of their response to the case study.

Imputed expertise

The extreme right hand columns of Table 3b as presented earlier give the comparisons of the three groups. The first impression is that there is actually not much difference at all. There is, in fact, only one instance in which a significant difference between principals in the Low and Medium categories remains significant when compared with those in the High category. Of course, we must immediately concede that both the small numbers in some categories and the potential frailty or inappropriateness of the analysis may be to blame. However, since the intention was to look for patterns to guide future analysis, we must take a closer look at what we do have.

1. The single robust significant difference occurs in relation to the use of examples. Principals classified into the Low imputed expertise category used examples more often than those in either the Medium or High categories. A closer look at the specifics of these codings revealed principals in the High group were only ever coded as using examples under the Problem function, whereas principals in both the Low and Medium categories used them in Detection and Attack as well. One way we might interpret this finding is that more expert principals were fabricating a customized solution to the problem, whereas weaker principals were conducting what we came to think of as a "lazy search," that is they tended to search for a ready-made solution stored in memory rather than seeking to build one specially for this problem. If our conceptual framework holds true, and pondering the problem is to be understood as searching the memory for appropriate matches, the use of examples in this process might, depending on the actual content of the expressed thought, be more appropriate than using them to support diagnoses, which might imply trying to force the current situation to fit a previous situation in order to find a ready-made solution.[4] This would also be an indication of a relatively impoverished and/or poorly organized problem relevant schemata, since more richly endowed, more highly differentiated and well-indexed schematic knowledge would

[4] In this respect, we recalled a subject in an earlier preliminary study (Nagy & Allison, 1989), who, after reading the case began by saying, "Oh, yes. I had a case just like this a few years ago and this is what I did." He then went on to describe the entire situation, which was not, in fact, very similar to the given case other than being located in the library, and never referred back to any specifics of the Miss MacDonald case. Again, in our first PPP study, several highly experienced principals retrieved and applied complicated solutions from memory that had little to do with the specifics of the presented case.

theoretically both allow and encourage a more thorough search for solution ingredients rather than the inappropriate imposition of a previously adopted response.

2. Principals in the Low imputed expertise group were coded far less frequently in detection related TUs, especially those specifying detection strategies, than were those in the Medium category, and (not significantly) those in the High category. Specifically, Low group principals did not refer to interpretations of the case material at all in their detection thinking, whereas principals in both the Medium and High groups occasionally did. The implication here may be another indication of lazy searching or poorly-developed schema, in that weaker principals gave the impression of "fishing" whereas principals with claims to higher imputed expertise did not. Further analysis of the specific thought units involved might illuminate this.

Since these comparisons among expertise groupings were so unrevealing, we tried other approaches to analysis to see if there were patterns elsewhere. One approach was to tabulate each individual case and directly compare the distribution of coded thoughts, looking for a difference in the distribution pattern. No meaningful patterns emerged. We did find something, however, when we returned to the raw frequencies instead of percentages. First of all, the principals in the Low imputed expertise group had been coded as having between 19 and 41 TUs in total, drawn from between 295 and 705 words; while the means for principals in the Medium group was 458 words, coded into 29 TUs. In contrast, principals classified in the High imputed expertise category were coded as having between 40 and 99 thought units, drawn from between 556 and 1656 words; the means being 919 words, coded into 61 TUs. Put simply, principals in the High imputed expertise group had on average twice as much to say as those in the Low group. When means of the raw frequencies for the individual cells in the coding matrix were calculated, this general finding of twice as much to say was found to have a few interesting distributions. Table 5 gives these statistics for principals in the High and Low imputed expertise groups. Those in the High imputed expertise category were coded with roughly twice as many TUs on average in most cells. When we focus attention on the four which were identified earlier as the most important cells, the pattern is a little different. On average, principals in the High expertise group had about twice as many attack strategies (12.2 as opposed to 6.4), and about twice as many TUs explaining diagnoses (6.0 as opposed to 2.6), almost four times as many action plans for detection (7.8 as opposed to 2.0) and roughly similar numbers of thoughts generally pondering the problem (6.3 as opposed to 5.2). Thus, whereas principals in both High and Low imputed expertise

Table 5
Means of raw frequencies for High and Low imputed expertise principals

High Imputed Expertise Principals (N=6):

	Operations										Row
Functions:	A/G	V	RQ	MC	C	E	CMV	CMP	SR	PSR	Totals:
Detection	7.8		0.8				0.2	1.2	0.2	1.5	11.7
Conclusion	0.2				0.2			1.7		6	8.1
Problem working	2.8	0.5				1.3		0.5	0.2	6.3	11.6
Attack	12.2	0.2						0.2		4.7	17.3
Column Totals:	23	0.7	0.8		0.2	1.3	0.2	3.6	0.4	18.5	48.7

Low Imputed Expertise Principals (N=5):

	Operations										Row
Functions:	A/G	V	RQ	MC	C	E	CMV	CMP	SR	PSR	Totals:
Detection	2		0.4			0.2				1.6	4.2
Conclusion								0.6		2.6	3.2
Problem working	0.8					2.2		0.8		5.2	9
Attack	6.4	0.4				1				2.2	10
Column Totals:	9.2	0.4	0.4			3.4		1.4		11.6	26.4

categories devoted roughly the same amount of attention to musing on the problem and searching memory for matches, the more expert principals had far more ideas about where to go next, particularly in taking action to complete the building of the problem space. In short, the thinking of the more expert principals was much more effective than that of the weaker principals in both identifying the possible components of the problem space and the ways in which to begin traversing the space, suggest that the schemata they accessed were more complex and relevant.

We must recognize, of course, that the judged quality of the response to the case is related to the length and fluency of the response.[5] Since our expertise ratings are based on other measures of performance as well as the judged quality of the case response, we would have to consider whether verbal fluency also influences those measures. We can certainly postulate that the more expert a principal is, the more thoughts she or he was able to bring to this kind of on-the-spot response, and thus we might conclude that the richer the schematic memory

[5] This is a phenomenon which we have observed before in unrelated research. See Nagy and Allison (1989) and Allison and Allison (1991).

accessed. At the same time, verbal fluency itself may well be an indicator of cognitive complexity, and thus a potential predictor of expertise.

Discussion

We can draw some reasonably solid conclusions about the way in which, using our analytical framework, our subjects approached the processing of the case problem. Four particular types of operation dominated. In later analyses we may wish to collapse some of the lesser used categories into these four dominant cells, but, bearing in mind the earlier caveat about getting too far away from the language of the original data, it would be inadvisable to do so on merely numerical grounds. If on further deliberation we feel that it would be useful to collapse the less frequently used cells into the dominant cells, each thought unit would have to be re-read and re-coded into the appropriate cell on the basis of our emerging understanding about what each represents.

The four dominant behaviours noted were Conclusion/Reasoning, Problem/Reasoning, Detection/Action and Attack/Action. If we attempt to view these operations through the screens provided by our overall conceptual framework of cognitive applications, which is explained in detail in the first paper in the symposium set, we find an interesting fit.

Conclusion/Reasoning involves interpreting the information given in such as way as to begin mapping the problem space. It is not a first level interpretation (which would probably have been coded under Case Material Processed), nor is it moving immediately to conclusions and acting on them (which would probably have been coded as Actions). It could be thought of rather as a process of testing first level interpretations against memory fragments for coherence and determining which slots in a retrieved schema can be filled immediately. Problem/Reasoning is a freer and less structured accessing of memory in the process of mapping the problem space. Detection/Action involves employing clearly-defined search strategies to find missing or desirable information, or to assign default values for the time being. This can be thought of as filling empty slots in the schema which is being used to map the problem space. Finally, Attack/Action is the plan of attack which will be used to begin traversing the problem space. Thought units coded into this last cell were, by definition, without preceding or succeeding traces of thinking which linked them directly to Detection, Conclusion or Problem functions. This lack of linkage implies, as noted earlier, that the subject is either not following the think-aloud instructions or is not aware of the source of the thought, thus retrieving the content directly from memory. Action oriented ideas which spring into short-term memory in this way, without apparent linkage to something else in short-term memory, must have been drawn from long-term memory. In terms of our framework, they must have been drawn from a schema in such a way that it is essentially automatic.

The sequence in which we explained these four components above seems to make sense: first finding pieces of interpretation which call up an appropriate schema to organize initial information, then pondering on the emergent problem space, then deciding how to fill in the missing pieces and finally planning an attack. In practice such a sequence did not occur in such a neat and tidy fashion: our subjects did not perform one operation then another in sequence, but jumped back and forth between them. This is not to say that thoughts were random or haphazard. Most subjects were coded with a sequence of thought units in one cell, followed by a sequence in another, often re-iterated. Figure 3a is a copy of a coding sheet from the transcript of how one of the principals in the High imputed expertise group thought through the problem, with linked sequences of TUs bracketed together. This principal began the working part of the think-aloud with three thought units (#s 3, 4 & 5—TUs #1 & #2 being coded in the Superfluous category) which were coded in the Detection/Action cell, followed by four coded in the Problem/Example cell, then two in the Detection/Reasoning cell and so on. This subject was coded as

	A/G	V	RQ	MC	C	E	CMV	CMP	SR	PSR
Detection	(3,4,5) (12) (16,17) (20) (32)									(10,11) (33)
Conclusion										
Problem Working	(24,25) (30)	(34,35,36)				(6,7,8,9)				(13,14,15) (18,19) (23) (31) (37,38,39,40) (43)
Action	(44,45)									(21) (26,27,28)
Superfluous	1,2,22,29,41,42,46									

Figure 3a
Sample coding sheet from a principal in the High imputed expertise group

articulating thoughts in the Detection/Action cell on five separate occasions as he or she worked through the problem, and six times in the Problem/Reasoning cell. The four TUs coded as examples occurred early in the thinking sequence, with no subsequent TUs being coded on this operation. Three TUs were coded as values statements towards the end of the sequence as the subject prepared to commit to a final plan of action. The essence of that plan was compactly and coherently stated in two TUs right at the end of the sequence before the final Superfluous comment signifying that the subject had finished.

	A/G	V	RQ	MC	C	E	CMV	CMP	SR	PSR
Detection	(9)									(3) (10,11)
Conclusion										
Problem Working								(13,14)		(4,5,6,7,8)
Action	(12) (18)	(2)								(15,16,17)
Superfluous	1,19									

Figure 3b
Sample coding sheet from a principal in the Low imputed expertise group

In contrast, Table 3b reproduces the coding sheet for one of the weak principals. The most immediate difference is in the overall number of TUs and the lack of re-iterative patterns. Obviously these two elements are intimately connected but there is an evident difference in the complexity of thinking brought to bear on the problem. These examples are typical: designated expert principals expressed more discrete thoughts in more complex patterns than did weak principals. Obviously these patterns deserve further study.

Future directions

Our results so far have given us an indication of a typical problem processing pattern and its component parts, and have suggested where differences between subjects with domain-relevant schema and those without, and differences between more expert and weak principals might lie. We plan to pursue three major avenues in future analysis.

1. The numerical data have shown us where patterns in the data might lie, but we must not lose sight of the fact that our data are verbal protocols. We must return to the actual data to look for meaning in some of the findings discussed above, to determine how well the numerical analyses predict differences in

meaning, and whether, in those instances in which we expected to find differences but did not, the differences reside in meaning rather than quantities.

2. Following on the same line of inquiry, we intend to combine the raw data from the process analysis described in this paper with those from the topical attention analysis described in a companion paper in this symposium set, to gain a more complete picture of the differences between principals and non-principals and between principals with different apparent levels of expertise.

3. The coding matrix proved very useful in bringing order to the data so that we could achieve the results we have. It is necessary now for us to take what we have learned and re-visit that analytical framework, re-modelling it to better serve the next phase of analysis. A re-modelling which takes into account the four main processes observed (i.e., information searching activities, diagnostic interpretative thinking, pondering, and launching a plan of attack), and recognizes the interactive nature of these, should be valuable in following up the question of sequence and how this reflects schema structure and perhaps cognitive complexity in problem processing.

List of Papers Comprising the 1996 AERA CASL Symposium Set

Author(s)	Titles
Allison, D. J.	Problem processing and the principalship: Theoretical foundations and the expertise issue.
Allison, D. J. & Morfitt, G.	Problem processing and the principalship: Design, methods and procedures.
Morfitt, G., Demaerschalk, D., & Allison, D. J.	Paying attention: Content considered by experts and others when solving a case problem.
Allison, P. A., Demaerschalk, D., & Allison, D. J.	Thinking through an administrative problem: Differences between expert, average and true novice responses.
Allison, D. J., Morfitt, G., & Demaerschalk, D.	Cognitive complexity and expertise: Relationships between measures of cognitive complexity and abstraction and responses to a case problem.
Allison, P. A., & Allison, D. J.	Cognitive approaches to school leadership: Some implications for the selection, training and development of principals.

Contact address: The CASL Project
Faculty of Education
The University of Western Ontario
London, Ontario, N6G 1G7
Canada.

CASL Project Paper #AERA96 4.2, revised August 1996.

References

Allison, D. J. (1996). Problem finding, classification and interpretation: In search of a theory of administrative problem processing. In K. Leithwood, J. Chapman, D. Corson, P. Hallinger & A. Hart (Eds.), *International handbook of educational leadership and administration* (Part 1, pp. 461—534). Dordrecht, Netherlands: Kluwer Academic Publishers.

Allison, D. J. & Allison, P. A. (1993). Both ends of a telescope: Experience and expertise in principal problem solving. *Educational Administration Quarterly, 29*(3), 302-322.

Allison, P. A., & Allison, D. J. (1991). *Experience and expertise in administrative problem solving*. ERIC Document # ED 333 583. Paper originally presented to the Annual Conference of the American Educational Research Council, Chicago, April, 1991).

Ericsson, K. A. (1975). Instruction to verbalize as a means to study problem solving processes with the Eight Puzzle: A preliminary study (No. 458). *Reports from the Department of Psychology*. Stockholm: University of Stockholm.

Goor, A. (1974). Problem solving processes of creative and non-creative students (Doctoral dissertation, University of North Carolina, 1974). *Dissertation Abstracts International*, 1974, *35*, 3517A. (University Microfilms No. 74-26878).

Kluckhohn, C. (1951). Values and value orientations in the theory of action: An exploration in definition and classification, in T. Parsons & E. Shills (Eds.), *Toward a general theory of action* (pp. 398-433). Cambridge, Mass.: Harvard University Press.

Leithwood, K. A. & Stager, M. (1989). Expertise in principals' problem solving. *Educational Administration Quarterly, 25*(2), 126-161.

Leithwood, K. A. & Steinbach, R. (1995). *Expert problem solving: Evidence from school and district leaders.* Albany, NY: State University of New York Press.

Nagy, P. (1991). *Using schema to map ill-structured problem solving: An application to school administration.* Paper presented at the Annual Meeting of the American Educational Research Association, Chicago.

Nagy, P. & Allison, P. A. (1989) The role of verbal fluency in the assessment of leadership potential. Paper presented at the VIIth World Congress of Comparative Education, Montreal.

Nagy, P. & Allison, P. A. (1988). *School-level decision making: A cognitive perspective*. Paper presented at the Annual Meeting of the Canadian Society for the Study of Education, Windsor.

Simon, H. A. & Kaplan, C. A. (1989). Foundations of cognitive science. In M. I. Posner (Ed.), *Foundations of cognitive science*, (pp. 1-47). Cambridge, Mass.: MIT Press.

Smith, C. O. (1971). *The structure of intellect processes analyses system. A technique for the investigation and quantification of problem solving processes.* Unpublished doctoral dissertation, University of Houston.

Weber, R. P. (1990). *Basic content analysis.* (2nd edition). Newbury Park, CA: Sage Publications Inc.

EA 028646

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Thinking through an administrative problem: Processing differences between expert, average and true novice responses

Author(s): Allison, Patricia A., Dawn Demaerschalk and Allison, Derek J.

Corporate Source: University of Western Ontario

Publication Date: April 1996

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

[✓] Check here **For Level 1 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

The sample sticker shown below will be affixed to all Level 2 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2

[ ] Check here **For Level 2 Release:** Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here→ please

Signature: [signature]

Printed Name/Position/Title: Patricia Allison

Organization/Address: University of Western Ontario
Faculty of Education
1137 Western Rd.
LONDON, Ont. N6G 1G7

Telephone: (519) 661-3182

FAX: (519) 661-3833

E-Mail Address: pallison@edu.uwo.ca

Date: Oct. 21/96 July 30/97