DOCUMENT RESUME

ED 412 604 EA 028 655

AUTHOR Allison, Derek J.; Morfitt, Grace; Demaerschalk, Dawn
TITLE Cognitive Complexity and Expertise: Relationships between External and Internal Measures of Cognitive Complexity and Abstraction, and Responses to a Case Problem.
SPONS AGENCY Social Sciences and Humanities Research Council of Canada, Ottawa (Ontario).
PUB DATE 1996-04-00
NOTE 29p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996). For a related paper, see EA 028 656.
CONTRACT 410-92-0329
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Ability; *Cognitive Measurement; *Cognitive Processes; *Educational Experience; Elementary Education; Foreign Countries; *Principals; Problem Solving; Productive Thinking; *Schemata (Cognition)
IDENTIFIERS *Cognitive Approaches to School Leadership Project

ABSTRACT

On-the-job experience is commonly regarded as the prime prerequisite for proficiency. This paper presents findings of a study that compared the ways in which a group of practicing elementary principals (n=31) and aspiring principals (n=25) thought their way through a case study. The study was conducted as part of the Cognitive Approaches to School Leadership (CASL) Project. The paper considers relationships between respondents' responses to the case problem and their cognitive abilities and styles, particularly cognitive complexity and levels of abstract thought. The theoretical frame adopted for the CASL Project associates the development of expertise in the principalship with the development of richer, more complex and integrated domain-relevant schemata. Four measures of cognitive complexity were used: Stamp's (1978) Symbol Card Task (SCT); Schroder's Paragraph Completion Test (PCT); a set of questionnaire items developed by Sashkin (1990) to estimate Jaques' level of work capacity through self-reports of time-span; and Fiedler's (1967) Least Preferred Co-worker (LPC) scale. The main finding is that domain knowledge, rather than general cognitive style or native capacity, better predicts and explains judged quality of response to the case problem. The cognitive capacities taped by the SCT appear to have little relevance. In addition, the significant age difference between the principals and novice participants may be masking or otherwise distorting the data. Other facets of the cognitive complexity complex, especially those tapped by the PCT, appear to be more relevant. Five figures and four tables are included. (Contains 53 references.) (LMI)

Cognitive complexity and expertise: Relationships between external and internal measures of cognitive complexity and abstraction, and responses to a case problem

Derek J. Allison, Grace Morfitt & Dawn Demaerschalk

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

The research reported in this paper was supported by the Social Sciences and Humanities Research Council of Canada under grant #410-92-0329

CASL Project Paper #AERA96 5.2. revised August 1996.

Cognitive complexity and expertise: Relationships between external and internal measures of cognitive complexity and abstraction and responses to a case problem.

Derek J. Allison, Grace Morfitt & Dawn Demaerschalk

This is this third of three analytical papers prepared for a 1996 AERA symposium reporting the conduct and some findings from our Cognitive Approaches to School Leadership [CASL] Project. The two accompanying analytical papers reported results from studies of how a group of practising principals and true novice subjects thought their way through a case study. The theoretical frame for the CASL Project and the design yielding the data discussed in the symposium were presented in two additional papers. This paper considers relationships between responses to the case problem and cognitive abilities and style, especially cognitive complexity and levels of abstract thought.

On-the-job experience is commonly regarded as the prime prerequisite for proficiency. By this token it is held that one cannot become a proficient practitioner without first hand experience, preferably a lot of it, in the domain in question. This is not news for professors and practitioners of educational administration, steeped as they are in the hoary old debates over theory and practice. What may be news is that contemporary cognitive science theory both acknowledges and offers an explanation for the central importance of sustained experience in the development of expertise. On this view, the rich, interconnected, cross-indexed, knowledge schemata believed to underlie domain expertise can only be developed through direct, extended, and intense interaction with the tasks and problems characteristically encountered in the domain. Only extended practice in working on the problems of practice will provide the learning opportunities for developing awareness of the cues and clues needed to recognize problem types and for building the productions, scripts and other chunks of proceduralized knowledge which are considered to provide the basis for automated reactions to frequently encountered situations, and the ingredients for understanding and planning responses to less frequently encountered problems or discovered opportunities. This is not to discount the importance of other forms of knowledge, especially so-called declarative knowledge acquired through reading, talking and participating in graduate seminars and the like. Not only does such knowledge have a potentially crucial contribution to make in constructing conceptual scaffolds which may help organize and

CASL Project Paper #AERA96 5.2, revised August 1996.

integrate lessons from experience, it also has an intrinsically important role to play in building schemata used to interpret the world.

But the most interesting and important point about experience is that does not automatically bestow expertise. Simple observation is sufficient to demonstrate that some principals appear to benefit from time in office more than do others. One principal (or teacher or other domain practitioner) with ten or so years of experience may be reputationally (even visibly) more competent than another with as much or more experience in role. A lack of expertise, or even a lack of basic competence, is of course always open to special explanation on the grounds of personal or contextual difficulties—the individual in question has developed an illness, suffered a domestic tragedy, or has been placed in an impossible situation. Yet while such explanations may sometimes be true and at others merely convenient, the basic claim that some practitioners appear to learn from experience better than others cannot be readily denied or disproved. The Leithwood and Stager (1989; Leithwood & Steinbach, 1995) study that pioneered cognitive approaches to studying school leadership provides some interesting data in this respect. Whereas the six designated expert principals in that study had accumulated an average of 15 years of experience, the 16 principals comprising their average expertise group had slightly more experience—a mean of 17 years in office. As discussed more fully in the design paper prepared for this symposium, our first PPP study provided us with an even clearer case in point. In that study we recruited subjects solely on the basis of experience, expecting more experienced principals to provide more expert responses to a case problem. As it turned out, three different juries—including one composed of reputationally expert principals—judged the responses from the most experienced principals (20 or more years in role) to be generally poorer than those given by some principals with much less experience. Mary Kennedy (1987) drew attention to the underlying point in her valuable analysis of various models of expertise and its development when she reminded readers that "experience can only contribute to expertise if practitioners are capable of learning from it" (p. 148). Wagner (1993) made the point more explicitly during his discussion of recent studies investigating the practical or tacit knowledge used by business managers (e.g. Wagner & Sternberg, 1985; 1990; Sternberg & Wagner, 1988). After noting that the measures used in that research differentiated between managers with more or less experience, he cautions that "not everyone appears to acquire tacit knowledge from the experience at the same rate, and some appear to acquire very little" (Wagner, 1993, p. 98).

What may account for this difference? While aspects of the role environment itself may at times impede learning (and at other times presumably enhance the process), it would seem that the major source of variation will be rooted in individual differences, presumably cognitive differences. Even so, variations in learning from work experience do not appear to be adequately captured, still less predicted, by IQ measures. Schon's (1983) well known deliberate action model points to possible variations in the ability to both interpret complex action situations and then to distil useful knowledge through reflection. Wagner and Sternberg's (1985; 1990) theory of tacit knowledge or "street smarts" builds from this view to postulate a form of practical intelligence that enables some individuals to learn more rapidly and completely from complex, open-ended, fluid work environments, with this practical intelligence being viewed as distinct from the type of intelligence measured by IQ tests. Gardner's (1983) theory of multiple intelligences proposes a more elaborate scheme which includes a variety of non-IQ based abilities and talents. A recent review of expert performance across diverse domains by Ericsson and Charness (1994) also points to the lack of a clear relationship with IQ. While they note that expert performers often have relatively high IQs, they conclude that "IQ does not reliably discriminate the best adult performers from less accomplished adult performers in the same domain" (p. 730). The reference to adult performance appears important on several counts. IQ measures were originally developed to predict children's success in school, and IQ tests developed for adult populations typically rely heavily on tests of linguistic, geometric, mathematical, and general reasoning, measures that remain rooted in scholastic knowledge domains and formats. As Wagner and Sternberg point out in discussing their street smarts theory, the cognitive factors or styles that appear to be associated with adults' abilities to learn from their experiences may have little to do with proficiency in school-like knowledge. Jaques (1986) drove this point home when he pointed out that "by 18 years of age one reaches full maturity in whatever it is that IQ tests rate" (p. 377). Even allowing for a substantial margin of error in this claim, it seems clear that factors other than those measured by IQ tests must play important parts in mediating how people learn from their work experiences.

Schema theory and cognitive complexity

The theoretical frame adopted for our CASL Project associates the development of expertise in the principalship with the development of richer, more complex and integrated domain relevant schemata. A schema is

understood as a cluster of situation and problem relevant knowledge held in long term memory which stores information about the characteristics of specific tasks and problems together with pre-formed problem spaces, typical goals, constraints, and solution procedures. Taylor and Crocker's (1981) definition portrays schemata as being "hierarchically organized with more abstract and general information at the top and categories of more specific information nested within the general categories" (p. 92). Experts' domain specific schemata are thought to embody many more interconnected elements and sub-assemblies than those of non-experts, allowing access to more embedded sub-schemata as well as more hierarchical levels of abstraction (Rumelhart & Norman, 1983). "Experts," according to Reimann and Chi (1989, p. 172) "have developed several layers of this hierarchy, whereas only the first level ... seems to be developed for the very beginning novices." This offers a potentially productive way of conceptualizing differences between the consequences of more and less productive experiential learning. Greater expertise, by definition, will be associated with the development of more levels of abstraction in schematic knowledge; lesser levels of expertise with fewer levels of abstraction. The potential to learn from experience—to form richer, more differentiated schemata—may thus be related to some broad ability to generate and work at greater levels of abstraction. Yet experts' schemata are also considered to be characterized by higher levels of integration and greater amounts of problem specific knowledge. These characteristics suggest that successful learning from experience will also be associated with the ability to recognize fine distinctions and make less than obvious connections between problem elements. In an earlier paper we sought to capture the essence of these seemingly paradoxical abilities through the notion of being able to examine domain problems and situations through both ends of a conceptual telescope—to draw on higher levels of abstraction to place things in context and see the bigger picture, while also being able to notice and take account of fine details embedded in situations (Allison & Allison, 1993).

A more detailed and sophisticated account of such processes is provided by contemporary theories of cognitive complexity. These theories stem from earlier work by Piaget (1952), Kelly (1955) and Bieri (1955). Bieri (1955) introduced the construct of cognitive complexity—simplicity, outlining his belief that the cognitively complex person has more dimensions of social judgement and versatility available than does the cognitively simple individual. Harvey, Hunt and Schroder (1961) related cognitive complexity to four developmental stages of abstract thinking, ranging from "concrete" to "abstract," a progression which O'Connor (1971) sought to measure through

his Abstract Orientation Scale [AOS]. Schroder, Driver and Streufert (1967) extended the general theory by recognizing the interaction of differentiation, integration and discrimination processes in cognitively complex individuals. More recently, Streufert (Streufert & Nogami, 1989; Streufert & Swazey, 1989) has developed a more complete application of this conceptualization to the management domain. The three key processes are generally understood in the following ways: differentiation refers to the attribution or recognition of multiple dimensions to a person, object, task, situation or other object of attention; integration speaks to interrelationships drawn or noticed between differentiated dimensions; discrimination has to do with gradation—shades of grey—within dimensions. Cognitive complexity is thus associated with recognizing greater and more finely graduated differences within objects and situations heeded and making more connections between them. Enriched schemata of the kind considered to be formed and used by experts can be readily seen as providing a memory deposit for an expert's ability to do just this when recognizing and working through domain problems. But the formation of such enriched schemata presumably requires some initial capacity to differentiate, discriminate and integrate from experience. Such a more general capacity is usually referred to in the literature as a cognitive style Whether cognitive complexity is best understood as a style—a pervasive, relatively consistent approach to viewing and interpreting the world—or as a learned ability embedded in schematic memory and keyed to a particular domain is an open question in the literature (Streufert & Nogami, 1989; Messnick, 1984). For the research reported here we chose to regard cognitive complexity as operating primarily as a style, while remaining open to the possibility that the ability to differentiate and integrate within a domain, such as school administration, may also be fostered or enhanced through training and other developmental activities.

Much of the theorizing and research regarding cognitive complexity as briefly reviewed above originated in North America. A related theory evolved independently on the other side of the Atlantic from early studies undertaken as part of the Glacier Metals Project in the United Kingdom (Brown & Jaques, 1965). This theory maintains that there are qualitative, disjunctive breaks in human capacity to do work which are manifest in successively more abstract levels of thinking linked to progressively longer time horizons. As originally developed by Elliott Jaques (1951; 1956; 1961; 1964; 1976, 1978) this theory was grounded in the time-span of discretion associated with the completion of tasks and responsibilities assigned to organizational roles. Jobs where the required work is largely prescribed and requires little discretion typically consist of relatively concrete tasks which

have short completion times, whereas jobs demanding the exercise of greater discretion typically require the completion of less well-structured, more abstract tasks extending over longer time spans. The key element of the theory postulates regular discontinuities in the time-spans associated with levels of abstraction in work. These disjunctions are claimed to occur at the three month, one year, two year, five year, ten year and twenty year time span levels, with these levels corresponding to qualitative breaks in the hierarchy of responsibility in formal organizations. Independent support for the principle of discontinuity underlying this theory has come from laboratory studies of adult problem solving by Isaac and his associates which identified multi-modal distributions in the response data which were theorized as corresponding to qualitatively different levels of abstraction in the way subjects responded to the presented problems (Isaac & O'Connor, 1969/1978; Gibson & Isaac, 1978).

Table 1

Jaques' levels of conceptual abstraction and time spans of discretion in work

Strata / Levels	Time Span [months]	Level of Conceptual Abstraction (Jaques, 1976 & 1978)	Cognitive States in SST Theory (Jaques, 1986)
VII	20 yrs [240]	?	**Extrapolative** development of whole systems
VI	10 yrs [120]	Institution Creating	**Definition** of whole system on large scale
V	5 yrs [60]	Intuitive Theory	**Shaping** of whole systems from within
IV	2 yrs [24]	Conceptual Modelling	**Transformation** of concrete systems
III	1 yr. [12]	Imaginal Scanning	**Extrapolation** in concrete systems
II	3m. [3]	Imaginal Concrete	**Definition** of tasks
I	< 3 months	Perceptual-motor Concrete	**Shaping** of concrete objects

More recent discussions of what Jaques (1986) has come to call Stratified Systems Theory [SST] extend the fundamental principles and provide extensive applications to organizational design and administrative practice, linking competence in hierarchically superordinate administrative roles to the development of intellectual capacities that are readily equated with cognitive complexity, or what Jaques terms cognitive power. As rendered in SST, the key time-span strata and associated levels of cognitive complexity and abstraction are as shown and named in the right hand column of Table 1. Based on an extensive analysis of comparative time-spans and earnings data, Jaques proposes that the capacity for work, and thus the associated levels of abstraction and related time-spans, matures throughout adulthood in general conformity with a set of growth curves (Jaques, 1989 Figure 4). This theory predicts, for example, that individuals with the cognitive capacity to work effectively at level 1 by

age 25 have the potential to effectively perform and gain satisfaction from level 2 work from about age 30 onwards. The potential growth curves associated with the satisfactory performance of higher levels of work at relatively young ages have steeper trajectories so that someone able to work at, say, level 3 by age 25 is regarded as having the potential to handle level 4 work at around age 35 - 45, with more rapidly maturing individuals in this group having the potential to work at level 5 tasks from their mid-forties onward. The precise ages concerned are less important for us than the principle underlying these growth curves, which would appear to map differences in the ability of individuals to learn from work experience, differences which would seem to be grounded in cognitive capacities associated with levels of abstract thinking.

Application to the study

Our conceptual framework treats the school principalship as a distinct knowledge and action domain. In conformity with the general theory of cognitive problem processing informing the study, this framework assumes that while prospective principals will acquire domain relevant knowledge prior to their first appointment, direct, first-hand interaction with domain tasks and problems will be required for the development of the enriched schemata associated with expertise. Given the discussion in previous pages, the development of such schemata is likely to be enhanced or inhibited by each principal's capacity for cognitive complexity. More cognitively complex principals—or in Jaques' terms those capable of functioning at higher levels of work—would thus be expected to learn more rapidly from their experiences, developing schemata incorporating more levels of abstraction and greater integration which, in turn, should enable them to demonstrate higher levels of expertise in their work and when responding to the case problem presented to subjects in our second PPP study. In contrast, less cognitively complex principals would be expected to benefit less from experience in role and thus demonstrate lower levels of expertise. Given the ubiquity of schools and the administrative process in contemporary society, then it also seems reasonable to expect that some aspects of the work domain of school leadership and administration will be comprehensible to non-principals, including true pre-novices who are not preparing for the principalship or have not been socialized to the adult work culture of schools. Within limits, then, greater cognitive complexity might also enable domain novices to think and reason coherently about the case problem, although their responses would

be expected to lack the technical detail and overall quality of response expected from practising principals, especially those who have developed greater expertise in the domain.

We attempted to explore the general ideas underlying these expectations in the analysis of data from our first PPP study as reported in the "telescope" paper mentioned earlier (Allison & Allison, 1993). That analysis was limited by the lack of a direct measure of level of abstraction which forced us to rely on estimates derived from the broadness of the goals pursued in the think aloud transcripts, which we roughly equated to Jaques' levels of abstraction. We sought to remedy this in the study reported in this symposium by incorporating a variety of independent measures of cognitive complexity.

Measures

Four measures of cognitive complexity discussed in the literature were employed: (1) Stamp's (1978) Symbol Card Task [SCT]; (2) Schroder's Paragraph Completion Test [PCT]; (3) a set of questionnaire items developed by Sashkin (1990) to estimate Jaques' level of work capacity through self-reports of time-span; and (4) Fiedler's (1967) Least Preferred Co-worker [LPC] scale. In addition, we also asked subjects about the overall goal they were attempting to achieve in their think aloud response to the case study and how long they thought it would likely take to achieve that goal in the context of the case. These last measures were collected in order to repeat the analysis undertaken in the earlier study using direct rather than imputed data.

The Symbol Card Task. This was developed by Stamp (1978) to provide an opportunity to observe how individuals exercise judgement and discretion when working on a problem, and thus estimate their "current level of capability" (Stamp, 1988, p. 11) within Jaques' levels of work framework. The task is a more complex and open version of a card-sorting activity originally used by Isaac and O'Connor (1969/1978), which was derived in part from Bruner's (1966) experimental work on concept formation. As refined by Stamp, subjects are asked to place cards bearing symbols to match a set of four display cards, using feedback from the task administrator in an attempt to discover the rule governing correct placement. Each of the 162 cards in the pack given to subjects contains a unique combination of symbols, which vary by shape, size, colour and number, each factor having three states (i.e. circle, triangle, square; small, medium, large, etc.). Three of the display cards contain an appropriate combination of symbols, the fourth being blank. As the subject places cards to match the display cards the

administrator confirms whether or not the placement matches the rule, except when cards are placed under the blank card, when no feedback is given. This adds a desirable element of ambiguity to the task. Stamp views the SCT as an opportunity to observe how subjects approach and exercise discretion when completing—or attempting to complete—a piece of work. Within our conceptual framework it can be viewed as a domain free, partly unstructured problem processing task. Knowledge embedded in domain relevant schemata—especially those developed for a social action area such as school leadership—will presumably have little if any application to finding a path through the problem space presented by the SCT. Moreover, the ambiguity surrounding the blank display card and the uncertainty regarding the nature as well as the specifics of the rule that subjects are seeking to discover makes the problem space less than well structured. Attempts at actively constructing a path through that space challenge participants to discriminate the salient dimensions in the task, differentiate between initially ambiguous feedback, and integrate the information collected. As such we viewed the SCT as providing a test of subjects' native capacity for cognitive complexity.

Performance on the SCT is scored against a set of behaviours, tactics and patterns of response identified by Stamp (1978, 1991) as being characteristic of each of Jaques' levels of work and abstraction. To enhance our application of these interpretive indicators and to refine administration procedures we first practised on half-a-dozen unsuspecting colleagues and then engaged in a more carefully conducted pilot study involving 14 principals participating in a residential professional development session. As used in this study, one member of the research team (usually the first author) administered the task while another (usually the second author) observed and took notes. The same placement rule (a two factor rule) was used for all subjects and the cards were resorted after trials to ensure each subject encountered the same sequence of cards as they worked on the problem. As soon as possible after completion of the task the administrator and observer compared the level to which they had independently assigned the subject. Initial ratings allowed for differentiation within levels, so that a subject might have been assessed as 3.4 by the observer and 3.8 by the administrator. The Pearson correlation between observer and administrator's initial ratings for the 56 novice and principal subjects was .841. Discrepancies between ratings were discussed to consensus in all cases. In addition to assigning subjects to Jaqusian levels of work/abstraction, we also took note of the time taken to complete the task, the number of cards used, and how many cards were placed before a subject placed his or her first card under the blank display card (the discard pile—and a

theoretically important variable). Number of cards used was the performance variable employed in the original Isaac and O'Connor (1969) study. After completion of the task, subjects were debriefed using a standard protocol, during which they were asked to describe how they had attacked the problem, the parallels they saw between the SCT and their regular work, and whether or not they enjoyed the task.

Schroder's Paragraph Completion Test. This measure evolved from Harvey's (Harvey *et al.* 1961) "This I believe" test. As modified by Schroder and Streufert (1963) and Schroder (1971) subjects are asked to write paragraph length responses to stem words or phrases. Six stems were used in our application, these being (1) Laws, (2) Human nature, (3) Self-interest, (4) Freedom, (5) The exploration of outer space, and (6) Leadership. Subject responses are scored for differentiation, discrimination, and integration in accord with guidelines given by Schroder (1971) in his *Integration Index* and by Baker-Brown *et al.* (1992), scores ranging from 1 (low integration) to 7 (high integration). High levels of reliability are reported when paragraphs are scored by trained analysts. Scoring in this study was done by the second and third authors who were not formally trained in the procedure. Consequently, our inter-rater reliability was not as satisfactory as we would wish, ranging from 59% to 83% initial agreement between coders, all discrepancies being discussed to consensus. Because we did not decide to add the PCT to our instrument battery until after the study was underway, we only have data from a total of 27 subjects, (15 principals and 12 true novices). Cronbach's alpha for the ratings on the 6 stems used in this application was .84. Scores from the six stems were summed to provide the single PCT measure reported in this paper.

Time Span Scales. We had originally intended to use O'Connor's Abstract Orientation Scale as a pencil and paper measure of cognitive complexity, but difficulties in obtaining a copy of the scale and the dated air of some items persuaded us to substitute a set of items taken (with permission) from Sashkin's (1990) Leader Behavior Questionnaire. In retrospect we regret forsaking the additional measure, but at the time the decision was made we wanted to avoid overburdening subjects with additional tasks. As described by Sashkin and Burke (1990, p. 313), the five item *Long-term Leadership* scale in the LBQ was designed to tap elements of Jaques' time-span of discretion construct, although they acknowledge that "this scale must be treated as experimental at this time" (p. 313). In reviewing these items we concluded that there was some conceptual overlap between two of the LBQ items, and decided to substitute an item of our own (TS5 in Table 2) for one of these two (item 38 in Sashkin 1990

[LBQ, Third Edition]). We also made some minor adjustments in the wording of several items. The wording of the five items finally used in our study is shown in Table 2. Cronbach's alpha for the 54 responses to the five items in this study was calculated as being a depressing .178. When item TS5 is removed from the matrix, alpha rises slightly to the still unsatisfactory level of .289. Given these unacceptable reliability statistics the five items were not combined to form a single scale and relevant statistics for each discrete item are reported in the analyses that follow. It is important to stress that because of the adjustments made to the items as used in this study, the reliability statistics and other data reported here cannot be generalized to the *Long-term Leadership* scale as used in the LBQ itself.

Least Preferred Co-worker Scale. Originally developed by Fiedler (1958; 1967) as a measure of leadership style, the LPC scale has attained a certain notoriety in the literature—while psychometric data appear to establish that the scale is measuring something, there is a distinct lack of agreement as to what this may be. Foa, Mitchell and Fiedler (1971) linked LPC scores to cognitive complexity and more recent versions of Fiedler's theory explicitly incorporate a "cognitive resources" component (Fiedler & Garcia, 1987). That the scale may measure or at least reflect cognitive complexity gains face validity from similarities between the semantic differential response scales comprising the LPC and similar scales used in Bieri's early work. Indeed, Bieri (1961, p. 370) specifically noted that the LPC may be a measure of cognitive complexity. Research evidence is mixed, suggesting that low LPC scores (often interpreted as being associated with Task Orientation [TO]) are likely indicative of low cognitive complexity, but that high scores (often associated with Relationship Orientation [RO]) do not necessarily equate with high levels of cognitive complexity (Arnett, 1978; Evans & Dermer, 1974; Mitchell, 1971; Vecchio, 1979). We included the LPC in our measures with a view to exploring relationships with the other measures adopted. Cronbach's alpha for the data obtained in this study was .905, with the overall mean score of 59.05 (s.d. 20.38), matching Posthuma's (1970) normative mean of 59.0.

Internal measures. In an earlier study we used imputed goals derived from reviews of think aloud transcripts as an indicator of the level of abstraction at which subjects were processing the case problem (Allison & Allison, 1993). This guiding objective or implicit goal was rarely articulated clearly by subjects during their analysis of the case in either the first or second rounds of data collection. In the second round study which provided the data discussed in this symposium subjects were asked a set of standard debriefing questions after they had

completed the think aloud session. One question in the debriefing protocol asked subjects to articulate the main goal they were attempting to achieve in their response to the case. Answers were subsequently analyzed and classified into the same increasingly abstract categories used for the imputed goals in the earlier study. These categories were constructed to capture the principle of increasing abstraction underlying Jaques' levels of work as shown in Table 1. Declared goals which focused on attempting to "fix" what subjects took to be the immediately presented problem or problems, with proposed actions typically being directed at adjusting or shaping concrete concerns, such as attending to the paperwork necessary to facilitate the librarian's transfer, improving physical aspects of the library, or adjusting library or school schedules, were classified in level 1, Physical goals category. Subjects classified as pursuing goals within the Personnel category (level 2) evidenced a preoccupation with personal needs of characters in the case (including the principal or the subject in the role of the principal) or the school staff in general. Level 3 goals were those that were primarily directed toward improving the library Program, while the highest abstraction level 4 category was reserved for goals involving Transformation of the school as a whole. Subjects' declared goals were reviewed independently by the second and third authors and then compared. Where more than one goal had been stated, the one falling into a higher level on the abstraction scale was coded and included in the database. There were only three cases of disagreement between the independently coded declared goals, which were readily resolved through discussion.

During the think aloud debriefing subjects were also asked how long they thought it would probably take to achieve their articulated goals within the circumstances described in the case. These time-to-goal responses were viewed as providing time-span estimates grounded in subjects' understanding of and response to the case, and thus as a domain referenced estimate of level of abstraction.

Results

Table 2 tabulates correlations between the measures of cognitive complexity and abstraction described above. In a number of ways the relationships in the Table are disappointing, revealing little in the way of coherence between measures. The strongest correlations are between variables within the Symbol Card Task, which relate in expected ways. Thus, the reasonably strong negative relationships between assigned level and number of cards used, time taken, and number of cards placed before the discard pile was used (before a card was

Table 2

Pearson Correlations between Measures of Cognitive Complexity and Abstraction

(Decimal points omitted. **Bold** indicates $p. < .05$)

	Symbol Card Task			PCT	LPC	Internal		Questionnaire Items				
	Cards	Time	Dis.	Sum	Score	Level	Time	TS1	TS2	TS3	TS4	TS5
N=	56	56	56	27	56	55[1]	55	55[2]	55	55	55	55
SCT Level (N=56)	**-605**	**-549**	**-598**	113	032	**-266**	141	208	177	131	-090	069
Number of Cards in **SCT**	—	**693**	**301**	190	159	199	186	-117	-086	022	096	-171
Time Taken for **SCT**	—	—	**314**	007	156	174	150	043	141	-62	179	-218
SCT cards before discard	—	—	—	-105	192	**271**	001	**-344**	-172	087	102	202
Paragraph Completion Test (Sum)				027	112	**466**		357	058	**459**	079	082
Least Preferred Co-worker Score				020	-078			-125	025	126	031	-028
Internal measures												
Level of Goal Abstraction						**512**		112	054	047	012	054
Time to Goal (Months)						—		**291**	101	235	057	017
Questionnaire items:												
TS1 I have plans in mind [for my school][3] that extend over several years.									**341**	089	019	-046
TS2 I often consider how a specific action plan I have developed might be extended to benefit the larger organization.										091	-125	175
TS3 I focus on clear-cut, short-term goals rather than less clear, longer range aims.											142	010
TS4 It is difficult to clearly explain my long range goals to others.												**-331**
TS5 It's frustrating when people I work for can't see the big picture.												—

[1] Although, as described in the design paper, we lost TA responses from one Novice and one Principal, the case debriefing data are available for the Principal, but not the Novice.

[2] One questionnaire from a Novice subject was unusable.

[3] Words in square brackets omitted from the questionnaire completed by Novices.

placed under the blank display card) are all to be expected. In these cases, moreover, the lack of stronger correlations helps illustrate the weight of observational data in assigning candidates to a SCT level. Even so, the lack of substantial and significant correlations with the various SCT measures and the other indices used is both disappointing and intriguing. What is more, two of the three statistically significant correlations are disturbing as the relationships indicated are contrary to what would be anticipated. The most worrisome relationship is the

weak ($r.$ = -.266, $p.$ = .0494) negative correlation between Internal level—that is the level of abstraction associated with the declared goal pursued in the case study think aloud session—and the SCT level. Given that these two measures share the same theory base, then a relationship is to be expected, but it should be in the positive direction rather than the negative. The obvious explanation that springs to mind is that we seriously erred in assessing subjects' level of cognitive capacity in the symbol card task. Yet this interpretation is not upheld by the significant and reasonably substantial correlations between the various internal SCT measures summarized in Table 2. Another possible, and interesting, explanation is that the application of abstract thinking in solving the domain task (the case study think aloud) utilized different cognitive resources from those tapped by the SCT. The most parsimonious explanation for this would seem to be that the pursuit of more abstract goals in the case response was driven in large part by learned schemata (acquired knowledge and skill), whereas the response to the SCT was driven by native capacity of the kind equated with cognitive style, or, in Jaques' terms, cognitive power. This explanation is accommodated by our conceptual framework, although we anticipated greater dependence between learned and innate applications of abstract analysis implied by the correlation in Table 2. A more prosaic explanation would be that the coefficient in question is a meaningless statistical artefact generated as a consequence of attempting to calculate a product moment correlation from interval data. To lay this ghost a Spearman rank order correlation was run between the SCT levels and the goal abstraction levels: *rho* was found to be -.264, substantiating the relationship shown in Table 2. Even so, the associated probability of .0494 leaves the door ajar for a Type 1 error.

The two remaining significant correlations between the SCT and other measures add fuel to the potential paradox. The positive correlation ($r.$ = .271 [$p.$ = .0456]; *rho* = .264) between the internal level of goal abstraction and the number of cards placed before the discard pile was used is also contrary to theoretical expectations: individuals operating at higher levels of abstraction are expected to make use of the discard pile earlier than others when completing the SCT. On the other hand, the significant negative relationship ($r.$ = -.344) between the TS1 questionnaire item and use of the discard pile conforms to theoretical expectations.. The TS1 item (I have plans in mind [for my school] that extend over several years) is perhaps the "cleanest" and most straight-forward of the time-span questionnaire items, and it also correlates positively ($r.$ = .291) with the internal measure of time-to-goal as shown in Table 2, as would be expected. Consequently, the positive relationship between the internal level of

goal abstraction and the use of the discard pile in the SCT is particularly intriguing. When taken together with the negative relationship between internal level and SCT level it would seem to support the possible explanation suggested earlier, the measure of internal level reflecting the application of learned schematic knowledge and the SCT level native capacity.

The two significant and reasonably substantial correlations with the PCT data are more straight-forward. It must be remembered that PCT data are available for only 27 subjects, that is about half of the dataset. Given this, the reasonably substantial positive associations between the PCT score and the time-to-goal measure from the think aloud problem ($r.$ = .466) and the TS3 questionnaire item ($r.$ = .459) are notable. So, too, is the lack of any relationships between the PCT score and the SCT scores. The moderately strong positive correlation between the internal measures of time-to-goal and level of goal abstraction ($r.$ = .512) is to be expected and provides evidence of internal reliability for these measures. The very low to virtually non existent correlations between the five questionnaire items explains the unacceptable reliability statistics noted earlier. With the exception of TS1 and TS2 where there is a weak to moderate positive relationship and TS4 and TS5 where a negative relationship is shown, there was a distinct lack of coherence across responses.

Table 3 shows the distribution of subjects by SCT levels and compares this with other distributions of interest. The numbers of subjects classified at SCT levels 4 and 5—the two together accounting for more than half of the subjects—may seem high, but it should be remembered that subjects were drawn from somewhat elite populations. The Chi Square statistics in Table 3 show there were no statistically significant differences in the SCT distribution associated with subject group (True Novices—Principals), or gender, or age. This is so for both the distributions shown in the Table and when the SCT levels were collapsed into low (levels 2 & 3) and high (levels 4 & 5) categories. the test statistics in the "SCT 2 Chi Square" sub-column reporting test statistics for the collapsed distributions. those in the SCT 4 column being for the distributions as shown in the Table. Nor were any differences apparent across the jury ratings of the think aloud responses, or across the imputed expertise categories for the principals. Subjects who performed on the SCT in such a way as to be classified in high cognitive capacity/levels of work categories were just as likely to have been judged as having given high—or low—quality responses to the case problem as were those who were classified in lower SCT categories. Given that the case problem was a domain anchored problem and the SCT appears as a less than well structured non-domain problem,

Table 3

Distribution of SCT levels by demographic and other Categories

		SCT Level						
		2	3	4	5	Total		
	All Ss	8 (14%)	14 (25%)	26 (48%)	8 (14%)	56	Chi Square	
							SCT 4	SCT 2
Group:	Novices	5 (20%)	7 (28%)	8 (32%)	5 (20%)	25	4.25 (3)	1.44(1)
	Principals	3 (10%)	7 (22%)	18 (58%)	3 (10%)	31	$p.= .235$	$p.= .231$
Gender:	Male	3 (12%)	8 (31%)	12 (46%)	3 (12%)	26	1.16(3)	0.19(1)
	Female	5 (17%)	6 (20%)	14 (47%)	5 (17%)	30	$p.= .763$	$p.= .667$
Age:	< 30 yrs	5 (31%)	4 (25%)	4 (25%)	3 (19%)	16		
	31-40	0	3 (33%)	4 (45%)	2 (22%)	9		
	41-50	2 (17%)	6 (24%)	16 (60%)	2 (8%)	25	9.5(9)	2.71(3)
	> 50 yrs	1 (17%)	1 (17%)	3 (50%)	1 (17%)	6	$p.= .385$	$p.= .439$
Jury Rating:	1 - 2	1 (8%)	3 (23%)	8 (62%)	1 (8%)	13		
	3 - 6	5 (18%)	9 (32%)	9 (32%)	5 (18%)	28	4.33(6)	2.07(2)
	7 - 9	2 (15%)	2 (15%)	7 (54%)	2 (15%)	13	$p.= .633$	$p.= .356$
Imputed Expertise Categories (Principals only)								
	Low	0	2 (40%)	3 (60%)	0	5		
	Medium	2 (10%)	4 (20%)	11 (60%)	2 (10%)	20	2.49(6)	0.19(2)
	High	1 (17%)	1 (17%)	3 (50%)	1 (17%)	6	$p.= .871$	$p.= .911$
LPC Type (Fiedler partition)								
	Low (TO)	6 (23%)	8 (31%)	9 (35%)	3 (25%)	26		
	Medium	0	0	7 (70%)	3 (30%)	10	10.30(6)	8.78(2)
	High (RO)	2 (10%)	6 (30%)	10 (50%)	2 (20%)	20	$p.= .112$	$p.= .012$

the complete lack of any relationship between performance on these two tasks implies a notable lack of dependence between the schematic knowledge associated with domain expertise and cognitive capacity tapped by the SCT. Such a conclusion is further supported by the lack of a relationship between SCT level and imputed expertise levels

for the principal group. In this case, however, the data to hand must be interpreted with caution, given the small numbers in the Low and High imputed expertise categories.

The last panel in Table 3 tabulates an interesting relationship between SCT levels and LPC types. In Fielder's theories and in other work, LPC scores are partitioned into high and low categories, people classified into the high LPC group being regarded as primarily relationship oriented [RO] and those in the low category as task oriented [TO]. There has been some considerable debate, nonetheless, over where the line dividing the two groups should be properly drawn. Whereas some approaches adopt a mean or median split, Fielder has maintained that the RO group should consist of subjects with LPC scores of 64 and above, and the TO group of those with scores of 57 and below. This leaves an intermediate range where Fielder suggests subjects may be either TO or RO. The mean for the LPC scores in this study was 59, falling slap in the middle of the indeterminate range, and also conforming to the normative mean derived by Posthuma (1970) in his review of LPC studies. As shown in Table 3, ten subjects in this study had LPC scores within the indeterminate medium range, and all ten of these people were classified in either SCT level 4 or 5, while subjects classified as TO or RO were distributed across all four SCT categories. When the SCT levels are collapsed into low and high groups, the Chi Square value for the distribution has an associated probability of .012, suggesting a real effect is at work. If mid-range LPC scores are interpreted as indicative of greater discrimination and differentiation, which is eminently plausible, then this effect would seem to be readily accommodated by cognitive complexity theory. If so, this would also help to restore confidence in the SCT as a measure of cognitive capacities related to cognitive complexity.

Further support for the SCT is provided by the distribution shown in Figure 1, which presents a histogram of the number of cards placed by subjects. The histogram bars and the accompanying density trace clearly show a multi-modal distribution which bears a marked similarly to that reported by Isaac and O'Connor (1978, Figure 4.1, p. 45). The dot plot which appears below the histogram is particularly informative, each of the dots in this plot showing the number of cards used by an individual subject, the overprinting indicating multiple subjects with similar card counts. This plot clearly shows a clustering of individual data points into five distinct groups.

Table 4 presents tabulations of data from other cognitive complexity and abstraction measures against the quality of response to the case study ratings obtained from the jury of reputationally expert principals, as explained more fully in the design paper prepared for the symposium. The jury ratings have been partitioned as shown in the

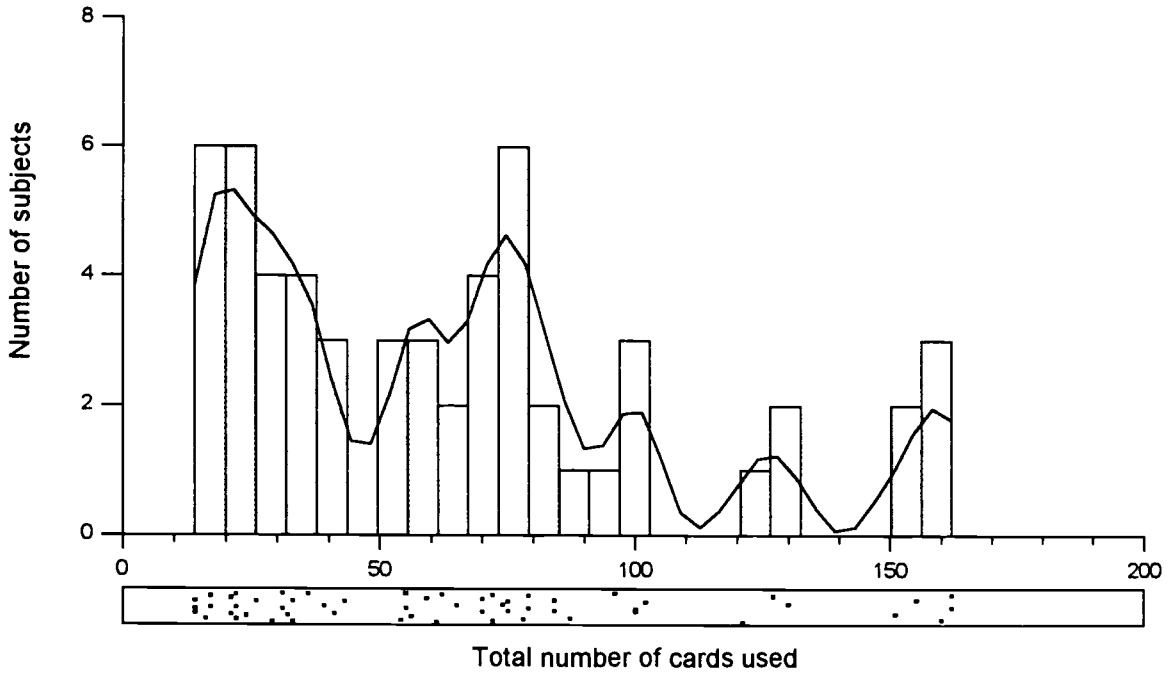

Figure 1
Histogram of number of cards used in SCT with density trace and dot plot

upper panel in the Table, the cut points providing a quartile distribution. The first cross-tabulation clearly demonstrates the operation of domain relevant knowledge, with none of the case responses from the novice group appearing in the upper quartile of the jury ratings.

The next panel shows the distribution of subjects by jury rating quartiles and PCT scores. The total PCT scores ranged from 6 to 24, with a mean of 17.3 and a median of 17. The cut point of 19 shown in the Table was selected from inspection of histograms of the scores which revealed a bi-modal distribution reminiscent of the multi-modal SCT distribution shown in Figure 1. Cutting the PCT distribution at 19 captured the two modes in the distribution while retaining a tolerable numerical balance between the comparison groups. Even so, the 10 subjects constituting the higher scoring group ensures that the expected cell values for the Chi Square comparisons fall below 5. To compensate for this, Table 4 shows Chi Square test statistics for both the distribution shown in the Table and for a cross-tabulation in which the jury ratings are collapsed in to two categories by splitting the distribution at 5. In both cases Chi Square remains significant. In addition a clear trend is evident in the Table: only one person with a high PCT score fell into the lower quartile of the jury ratings, while half of this high PCT

Table 4

Distribution of jury ratings against selected cognitive complexity measures

		Jury Ratings					
		1 — 2	3 — 6	7 — 9	Total		
	All Ss	13 (24%)	28 (52%)	13 (24%)	54	Chi Square	
						As shown	Collapsed
Group:	Novices	11 (46%)	13 (54%)	0	24	18.94(2)	n.a.
	Principals	2 (7%)	15 (50%)	13 (43%)	30	$p. < .0001$	n.a.
PCT Sum:	< = 19	6 (38%)	9 (56%)	1 (6%)	16	7.58(2)	6.52 (1)
	> 19	1 (10%)	4 (40%)	5 (50%)	10	$p. = .028$	$p. = 0.012$
TS1 "I have plans in mind [for my school] that extend over several years":							
	< = 4	8 (25%)	20 (65%)	3 (10%)	31	8.98 (2)	n.a
	> 4	4 (18%)	8 (37%)	10 (46%)	22	$p. = .011$	n.a
TS3 "I focus on clear-cut, short-term goals rather than less clear, long range aims":							
	< = 4	10 (26%)	23 (59%)	6 (15%)	39	6.67 (2)	0.20 (1)
	> 4	2 (14%)	5 (36%)	7 (50%)	14	$p. = .036$	$p. = .198$
Goal Abstraction from Think Aloud Task							
	Concrete	3 (50%)	2 (33%)	1 (17%)	6		
	Personnel	8 (33%)	14 (58%)	2 (9%)	24		
	Program	0	8 (67%)	4 (33%)	12	13.97 (6)	10.02(2)
	Transform	2 (17%)	4 (34%)	6 (50%)	12	$p. = .030$	$p. = .006$
Maximum time to Goal (Months)							
	< = 6	8 (47%)	9 (53%)	0	17		
	6 — 12	4 (19%)	14 (67%)	3 (14%)	21	23.0 (4)	n.a
	> 12	1 (6%)	5 (31%)	10 (63%)	16	$p. = .0001$	n.a

group gave case responses judged as falling into the upper quartile of ratings. The jury, of course, had no knowledge of the PCT scores.

The next two panels in Table 4 cross-tabulate responses to selected questionnaire items against the jury ratings. Questionnaire responses were all made on a five point scale, where a response of five was specified as

being "completely true" with lower numbers indicating the questionnaire item was considered progressively less true. Responses to most of the questionnaire items were considerably skewed to the upper end of the response scale. In consequence, the two comparisons shown in Table 4 divide subjects on the basis of those giving a five (completely true) response and all others. A clear difference is evident for the pattern of responses to the TS1 item, respondents claiming this as being completely true for them tending to have received higher jury ratings. The patterns are less clear in the case of TS3, primarily as a result of the relatively few respondents choosing the completely true category. When the jury rating categories were collapsed to two, the initial significant difference disappears.

The lower two panels relate the two internal measures of goal abstraction and time-to-goal from the think aloud task to the jury ratings. The jurors, of course, did not have access to these data, although some think aloud responses made specific reference to both goals and time lines, and in others a subject's goals could be inferred from his or her proposed actions. The patterns shown in the cross-tabulations support the findings reported in our earlier analysis of data from our first PPP study (Allison & Allison, 1993). Subjects who were judged to have given higher quality responses to the case study tended to have pursued more abstract goals, the pursuit of which they saw as extending over longer time spans. The time span relationship is particularly clear in the data display, none of the 13 subjects in the highest jury ratings category—all of whom were principals—considering they could achieve their goals in less than six months. The difference involved here is more marked when the mean times for the three groups are compared: subjects in the lower quartile of the jury ratings had a mean time-to-goal of 6.4 months, those in the interquartile range, 12.7 months, while those in the upper quartile had a mean time to goal of 24.9 months ($F[1,53] = 11.5$, $p. < .0001$).

Figure 2 illustrates relationships between judged quality of response to the case problem and goal abstraction for the novice and principal groups. For the principals (designated by the circles, which plot mean scores) a steady upward progression in the jury ratings is evident with increasing levels of goal abstraction. As far as our reputationally expert jurors were concerned, pursuit of more abstract goals was associated with a higher quality response. The same is true for the novice group, but with two important contrasts. First, while novices pursuing Personnel and Program goals were awarded progressively higher ratings by the jury, their mean ratings are consistently lower than those of the principals pursuing similar level goals. This would seem to imply that even

when principals were pursing less abstract goals, they were able to draw on and apply domain knowledge for which the jurors gave some credit. Second, there is a marked downturn in the ratings for novices who declared in the debriefing session that they were attempting to pursue Transformational level goals in their case response. Indeed, the mean jury rating for novices pursuing Transformational goals is lower than that for novices pursuing Program level goals, although the standard error around the mean is greater. Moreover, while the plot shows a sharp down turn for the novice group between the program and transformational goal categories, a slight upturn is detectable for the principal group. The implication here would appear to be that the pursuit of transformation level goals requires domain knowledge which the novices were unable to make reference to in their TA responses, branding them as novices in the eyes of the jurors, but which the principals pursuing such goals were able to at least outline, impressing the jurors.

The final two Figures return the discussion to the relationship between experience and the acquisition of domain knowledge with which we began this paper. Both Figure 4 and Figure 5 plot the principals involved in the

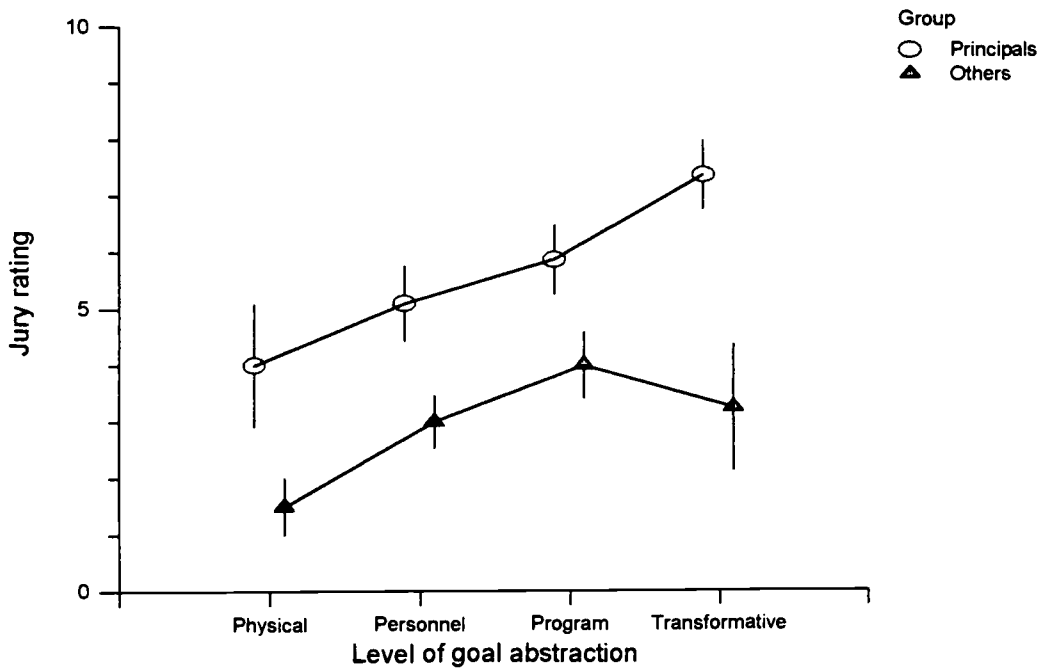

Figure 2
Error bar plot of level of goal abstraction in think aloud task

study against their years of experience-in-role and the jury ratings assigned to their think aloud responses. The horizontal and vertical axes in the two Figures are scaled the same, and thus the two plots are directly comparable. The important difference between the two plots lies in the numbers which plot each principal. In Figure 4 these numbers indicate the SCT Level to which they were assigned, and in Figure 5 they represent the level of goal abstraction they were judged to be pursuing in their think aloud solution. Ignoring the significance of these numbers for a moment, the overall pattern shown in the plots drives home the point made in the beginning of this paper regarding the relationship between experience and expertise: there is no clearly evident positive relationship in the plots between experience in role and jury ratings. This may be partly attributable to a lack of more experienced principals in the sample. Even so, of the six principals with ten or more years of experience in role, the think aloud responses from two fell in the mid range of the jury ratings, two fell in the upper range and two in the lower. Moreover, the plots show that the three principals receiving the highest ratings all had five *or less* years of domain experience. But while in those cases the principals concerned demonstrated high expertise with little experience, fully half of the principals with five or less years of experience received only moderate to weak ratings from the jury. So we return to our original question: why should this be so?

If cognitive capacity of the kind measured by the Symbol Card Task facilitates learning from experience, then those at lower SCT levels with little experience would be expected to have received lower ratings from the jury. As shown in Figure 4, this was so to a limited degree. The two principals at SCT level 3 with only a few years of experience both received low ratings—but so did four principals at SCT level 4 with less than ten years of experience. Then again, a single principal with but a few years of experience but whose case response was awarded a nine rating was classified at SCT level 2. Even so, this is the only truly glaring anomaly, and might be explained by exceptional circumstances. With a few less glaring exceptions, (all of whom have seven or more years of experience) all of the other principals who received moderate to high ratings with relatively little experience were classified at SCT Levels 4 or 5.

Figure 5 has a few anomalies also, chief among which is the principal with five years of experience whose case response received a nine rating, but who declared in the debriefing that he or she was pursing a Personnel (2) level goal. This aside, and ignoring the two less experienced but very high scoring individuals and the two principals with but a few years of experience who received low ratings, there is a tendency in Figure 5 for the

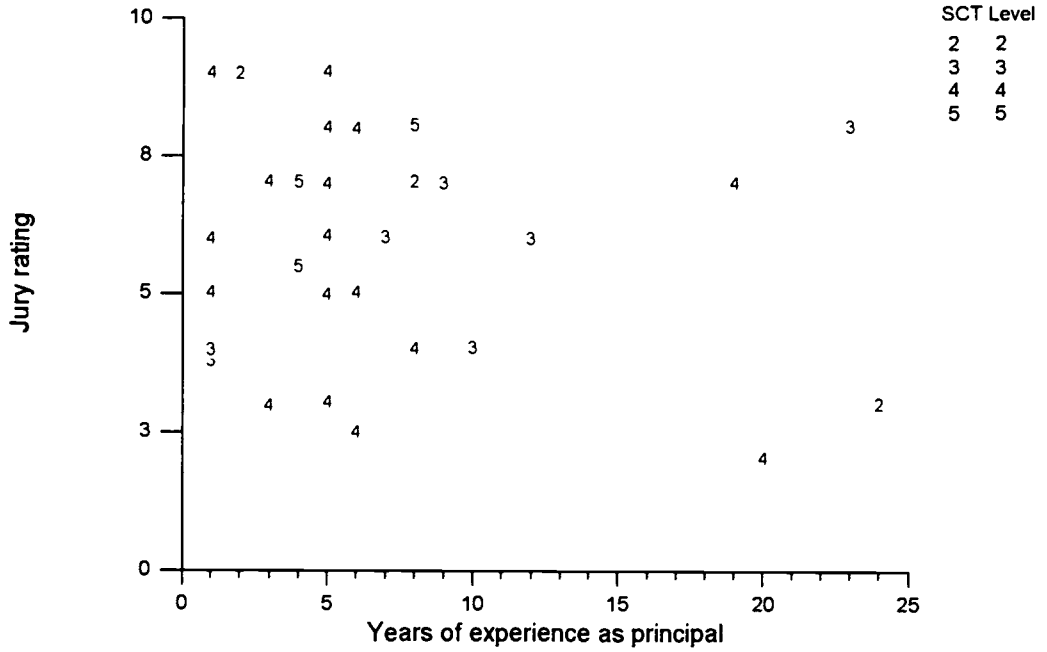

Figure 4
Years of experience in role by jury ratings, showing SCT level

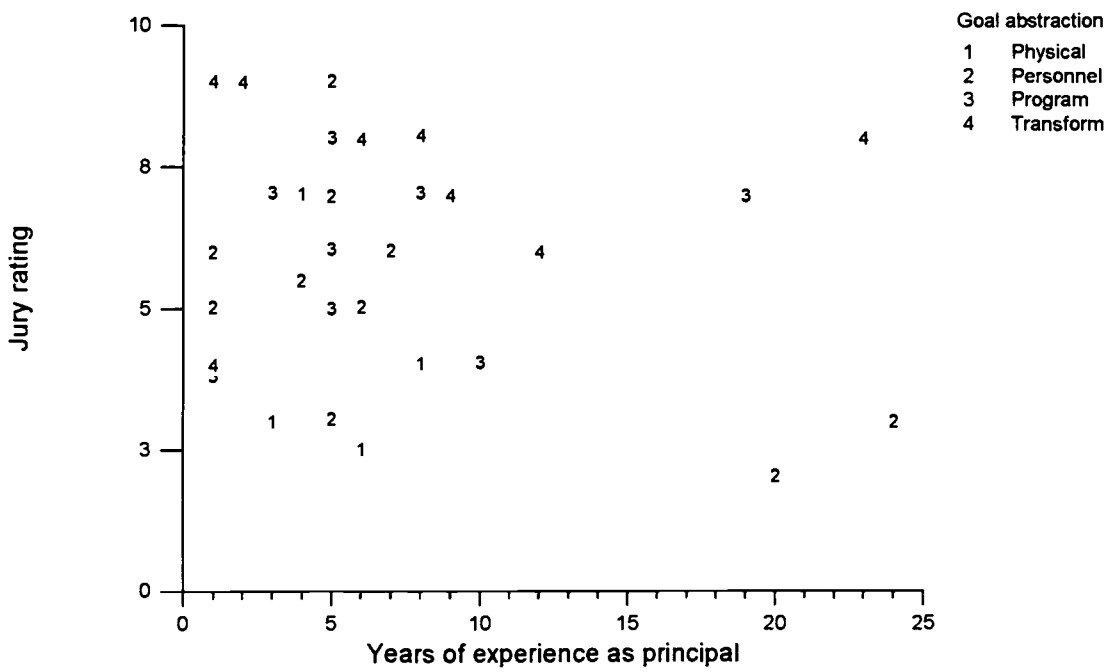

Figure 5
Years of experience in role by jury ratings, showing level of goal abstraction Level

plotted numbers to increase in value moving from the lower left to the upper right. There are, again, a few exceptions, but if this pattern is as suggested, it implies something of a positive correlation between experience and the recognition and pursuit of more inclusive and abstract goals, presumably as a product of enriched schemata embodying increased levels of abstraction. The clear and interesting exception is provided by the two highly experienced principals who received very low ratings for their think aloud responses, and who were both pursuing Personnel level goals. These two individuals, it seems, did not learn as well from their experience as did others. If both had been placed at low SCT Levels, then a potential explanation would be at hand based on the reasoning advanced in the early part of this paper. But as can be seen by comparing across Figures 4 and 5, only one of these principals was placed at SCT Level 2, the other at SCT 4.

Conclusion

The main finding to emerge from the results reported is that domain knowledge, rather than general cognitive style or native capacity, would appear to better predict and explain judged quality of response to the case problem. In the context of findings from previous expert-novice studies, as well as the theoretical framework used in the study, this is not remarkable. Given the social context of the principalship, it is nonetheless a finding of some potential import. It would appear that the elementary principalship may well have substantive grounds for being recognized as an area of expertise in its own right. As far as the development of expertise in this domain is concerned, the findings reported here suggest that the cognitive capacities tapped by the Symbol Card Task may have little relevance. Still, the lack of anticipated relationships (especially the several contradictory correlations) between our SCT data and the other measures employed suggest that some confounding influence may be at work. Given the relationships between age and level of work reported by Jaques, then the significant difference in age between the principals and other participants may be masking or otherwise distorting the data. We plan to investigate this. Even so, the most clear demonstration of differences in Jaqusian levels of work and time span in the present analysis came from domain anchored measures, and not the SCT. Other facets of the cognitive complexity complex, especially those tapped by the PCT, appear to be more relevant. Nonetheless, our main finding is that the task of disentangling the horse of cognitive capacity from the cart of expert schematic knowledge is a lot more difficult than it looks.

References

Allison, D. J., & Allison, P. A. (1993). Both ends of a telescope: Experience and expertise in administrative problem solving. *Educational Administration Quarterly, 29*, 302—321.

Arnett, M. D. (1978). Least preferred co-worker score as a measure of cognitive complexity. *Perceptual and motor skills*, 47, 567—574.

Baker-Brown, G., Ballard, E. J., Bluck, S., de Vries, B., Suedfeld, P., & Tetlock, P. E. (1992). The conceptual/integrative complexity scoring manual. In C. P. Smith (Ed.), *Motivation and personality: Handbook of thematic contact analysis* (pp. 400-418). Cambridge: Cambridge University Press.

Bieri, J. (1955). Cognitive complexity—simplicity and predictive behavior. *Journal of Abnormal and Social Psychology, 51*, 263—268.

Bieri, J. (1961). Complexity—simplicity as a personality variable in cognitive and preferential behavior. In D. W. Fiske & S. R. Maddi (Eds.), *Functions of varied experience*. Homewood IL: Dorsey Press.

Brown, W., & Jaques, E. (1965). *Glacier project papers*. London: Heinemann.

Bruner, J. (1966). *Toward a theory of instruction*. New York: Norton.

Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist, 49*, 725—747.

Evans, M. G., & Dermer, J. (1974). What does the least preferred co-worker scale really measure?: A cognitive interpretation. *Journal of Applied Psychology, 59*, 202—206.

Fiedler, F. E. & Garcia, J. E. (1987). *New approaches to leadership effectiveness: Cognitive resources and organizational performance*. New York: Wiley.

Fiedler, F. E. (1967). A theory of leadership effectiveness. New York: McGraw-Hill.

Fiedler, F. E. (1958). Leader attitudes and group effectiveness. Urbana, IL: University of Illinois Press.

Foa, U. G., Mitchell, T. R., & Fiedler, F. E. (1971). Differentiation matching. *Behavioral Science, 16*, 130-142.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.

Gibson, R. O., & Isaac, D. J. (1978). Truth tables as a formal device in the analysis of human actions. In E. Jaques, R. O. Gibson, D. J. Isaac (Eds.), *Levels of abstraction in logic and human action: A theory of discontinuity in the structure of mathematical logic, psychological behaviour and social organization* (pp. 132—160). London: Heinemann.

Harvey, O. J., Hunt, D., & Schroder, H. M. (1961). *Conceptual systems and personality organization*. New York: Wiley.

Isaac, D. J., & O'Connor, B. M. (1978). Experimental treatment of discontinuity theory of psychological development. In E. Jaques, R. O. Gibson, D. J. Isaac (Eds.), *Levels of abstraction in logic and human action: A theory of discontinuity in the structure of mathematical logic, psychological behaviour and social organization* (pp. 37—97). London: Heinemann.

Isaac, D. J., & O'Connor, B. M. (1969/1978). Experimental treatment of discontinuity theory of psychological development. *Human Relations, 22*, 427—455.

Jaques, E. (1951). *The changing culture of a factory*. New York: Dryden Press.

Jaques, E. (1956). *The measurement of responsibility*. Cambridge, MA: University Press.

Jaques, E. (1964). *Time-span handbook*. London: Heinemann.

Jaques, E. (1976). *A general theory of bureaucracy*. London: Heinemann.

Jaques, E. (1986). The development of intellectual capacity: A discussion of stratified systems theory. *The Journal of Applied Behavioural Science, 22*, 361—383.

Jaques, E. (1989). *Requisite organization: The CEO's guide to creative structure and leadership*. Arlington, VA: Cason Hall.

Kelly, G. S. (1955). *The psychology of personal constructs*, (Vol. 1 A theory of personality). New York: W. Norton.

Kennedy, M. M. (1987). Inexact sciences: Professional education and the development of expertise. In E. Z. Rothkopf (Ed.), *Review of research in Education*, (pp. 133—167). Washington, DC: American Educational Research Association:

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

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

Messick, S. (1984). The nature of cognitive styles: Problems and promise in educational practice. *Educational Psychologist, 19*, 59—74.

Mitchell, T. R. (1971). Leader complexity and leadership style. *Journal of personality and social psychology, 16*, 166-174.

O'Connor, J. (1971). Developmental changes in abstractness and moral reasoning. Unpublished Doctoral Dissertation, George Peabody College for Teachers.

Piaget, J. (1952). *The origins of intelligence in children*. New York: International Universities Press.

Posthuma, A. B. (1970). Normative data on the Least Preferred Co-worker scale and group atmosphere. University of Washington, Seattle. Department of Psychology, Organizational Research Technical Report 70-B.

Reimann, P. & Chi, M. T. H. (1989). Human expertise. In K. J. Gilhooly (Ed.), *Human and machine problem solving*, (pp. 161—191). New York: Plenum Press.

Rummelhart, D. E., & Norman, D. A. (1983). *Representation in memory*. San Diego: University of California.

Sashkin, M. (1990). *Leader Behavior Questionnaire*. (Third Ed.). King of Prussia, PA: Organizational Design and Development.

Sashkin, M., & Burke, W. W. (1990). Understanding and assessing organizational leadership. In K. E. Clark & M. B. Clark (Eds.), *Measures of Leadership* (pp. 297—325). West Orange, NJ: Leadership Library of America.

Schon, D. A. (1983). *The reflective practitioner*. New York. Basic Books.

Schroder, H. M. (1989). Managerial competence : the key to excellence. Dubuque, Iowa : Kendall/Hunt.

Schroder, H. M., & Suedfeld, P. (1971). Personality theory and information processing. New York: Ronald Press.

Schroder, H. M., Driver, M. J., & Streufert, S. (1969). *Human information processing*. New York: Holt, Rinehart and Winston.

Schroder, H. M., & Streufert, S. (1963) The measurement of four systems of personality structure varying in level of abstractness: Sentence completion method. Princeton University, ONR Technical Report 11.

Stamp, G. (1978). Assessment of individual capacity. In E. Jaques, R. O. Gibson, D. J. Isaac (Eds.) *Levels of abstraction in logic and human action: A theory of discontinuity in the structure of mathematical logic, psychological behaviour and social organization* (pp. 251—270). London: Heinemann.

Stamp, G. (1991). The symbol card task. In C. Stamp & J. Baker (Eds.), *Listening to my story has been really interesting: A guide to Career Path Appreciation*, (pp. 64—79). Uxbridge, Middlesex: Brunel Institute of Organisation and Social Studies.

Sternberg, R. J., & Wagner, R. K. (1988). Individual differences in practical knowledge and its acquisition. In P. L. Ackerman, R. J. Sternberg, & R. Glaser (Eds.), *Learning and individual differences* (pp. 255—278). New York: Freeman.

Streufert, S. & Nogami, G. Y. (1989). Cognitive style and complexity: Implications for I/O psychology. In C. I. Cooper, & I. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 93—143). New York: John Wiley and Sons.

Streufert, S. & Swezey, R. J. (1986). *Complexity, managers and organizations*. New York: Academic Press.

Taylor, S. E., & Crocker, J. (1981). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman, M. P. Zanna (Eds.), *Social cognition: The Ontario symposium* (pp. 89—133). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vecchio, R. (1979). A test of the cognitive complexity interpretation of the least preferred co-worker scale. *Educational and Psychological Measurement, 39,* 523—526.

Wagner, R. K. (1993). Practical problem-solving. In P. Hallinger, K. Leithwood & J. Murphy (Eds.). *Cognitive perspectives on school leadership,* (pp. 88—102). New York: Teachers College Press.

Wagner, R. K., & Sternberg, R. J. (1990). Street smarts. In K. E. Clark & M. B. Clark (Eds.), *Measures of leadership* (pp. 493—504). West Orange, NJ: Leadership Library of America.

Wagner, R. K., & Sternberg, R. J. (1985). Practical intelligence in real-world pursuits: The role of tacit knowledge. *Journal of personality and social psychology, 48,* 463—458.

Weiss, H. M., & Adler, S. (1981). Cognitive complexity and the structure of implicit leadership theories. *Journal of Applied Psychology, 66,* 69—78.

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.

email: allison@edu.uwo.ca

CASL Project Paper #AERA96 5.2, revised August 1996.

EA028655

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

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Cognitive complexity and expertise: Relationships between external and internal measures of cognitive complexity and abstraction, and responses to a case problem.

Author(s): Allison, Derek J., Morfitt, Grace and Demaerschalk, Dawn

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.

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

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

[X]

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.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1

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:	Printed Name/Position/Title: Derek J. Allison, Associate Professor	
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: allison@edu.uwo.ca	Date: Oct. 21/96 July 30/97