INCORPORATING COGNITIVE DEVELOPMENTAL CONCEPTS INTO STUDENT EVALUATION OF COLLEGE TEACHING

Presented at the Annual APA Convention
in Los Angeles, California
August, 1981

Tom E Trabin
Hennepin County Mental Health Center

Kenneth Doyle
University of Minnesota Measurement Services Center

INCORPORATING COGNITIVE DEVELOPMENTAL CONCEPTS INTO STUDENT EVALUATION OF COLLEGE TEACHING

Presented at the Annual APA Convention
in Los Angeles, California
August, 1981

Tom E Trabin
Hennepin County Mental Health Center

Kenneth Doyle
University of Minnesota Measurement Services Center

During the past two decades, research designs of outcome studies in educational psychology have increasingly combined what Cronbach (1957) called "the two disciplines of scientific psychology". Correlational research attentive to individual differences became more frequently incorporated into experimental designs that sought to control and manipulate types of treatments. Nearly a thousand such studies were recently reviewed by Cronbach and Snow (1977) with results summarized as modest if not equivocal. Recommendations for future research have included the use of multivariate analysis of several dependent variables (Cronbach and Snow, 1977), the use of more complex designs that account for interactions between a greater number of independent variables (Jenkins, 1977), and more extensive theory development and clarification of these interactions as a guide to more fruitful research (Blommers, 1970; Cronbach and Snow, 1977).

While debate continues on types of teaching methods most facilitative of different outcomes for different types of students, regular evaluation of instruction continues within college settings. Student evaluation of instruction forms are used in many college and university settings to provide feedback to instructors on their teaching strengths and weaknesses, as well as to provide information to adminsitrators making promotion and salary decisions. Although considerable psychometric rigour has been put into the construction and validation of the items in these rating forms, the item pools have not tended to focus upon possible attribute-treatment interactions that might effect how students view their instructors. Typically, some attempts have been made to account for achievement ("what is your G.P.A.?") and motivation ("was this course required for your major?"), but more subtle personality and cognitive style variables rarely seem to be incorporated into existing item pools according to

the available literature.

Some experts in cognitive style research have strongly recommended studying the interaction of cognitive styles and student ratings of instructors (Messick, 1970; Witkin, 1976). Initial studies using traditional measures of cognitive styles have reported some significant interactions between these styles and student perceptions of teaching (Crockett, 1975; Witkin, 1976; Wright and Richardson, 1977; Hacker and Bain, 1978). This suggests potentially fruitful use of cognitive style and cognitive development items in evaluation of instruction questionnaires. However, there are several major difficulties with pursuing this task: 1) the psychometric rigour of some of these measures has been brought into question, 2) most cognitive style and cognitive developmental measures are too cumbersome for routine mass administration and scoring, and not easily transformable into a few select items, and 3) cognitive style and developmental theories and measures are diverse, often loosely defined, and unclear in their interrelationships with eachother. The purpose of this study is to take some initial steps towards reducing these difficulties.

Measures

Cognitive style and cognitive developmental measures were selected on the basis of wide usage, research respectability relative to other measures, and diversity of method. For these reasons it was decided to measure field independence with the Group Embedded Figures Test, cognitive complexity with Bieri's modified version of Kelly's Rep Grid Inventory, and integrative complexity with the Paragraph Completion Test.

Field dependence-independence is a bipolar cognitive style defined as "the extent to which a person is able to deal with a part of a field separately from the field as a whole, or the extent to which he/she is

able to disembed items from an organized context. . . the extent to which he is analytical." (Witkin, 1976, pp. 41-42). A review of accumulated education-related findings across numerous studies (Witkin, Moore, Goodenough and Cox; 1977) suggests that field independent students, in comparison to field dependent students, are 1) less reliant on external references and therefore less likely to require externally defined goals and reinforcements, 2) less attentive to and therefore less able to remember social material, 3) more likely to make use of mediational processes such as analyzing, structuring, abstracting, and general principles, and 4) less likely to have difficulty in accepting the irrelevance of salient attributes in concept learning. The GEFT used in this study to measure field independence is a timed test consisting of 18 complex designs from each of which the subject must disembed a specified geometric figure; lower scores indicate greater field dependence.

Cognitive complexity is also a bipolar cognitive style, which has its roots in Kelly's (1955) theory that the master cognitive motive is the tendency of the organism to move in the direction of better prediction of others' behavior. Kelly's Role Construct Repertory Test (Rep Test) for eliciting and analyzing an individual's personal constructs for social judgment was later modified into a 10 x 10 grid form with constructs already provided (Bieri et al., 1966). The subject rates ten different roles (i.e. self, parent, friend, etc.) on each of the ten constructs using a six-point Lickert-type rating scale. Within each role, all possible pairwise comparisons of ratings are made (45), with a score of 1 given for each comparison in which ratings differed and a score of 0 for each comparison in which ratings were the same. In this way, 450 pairwise comparisons are made for each subject. The concept of cognitive complexity based upon this modified Rep Test was explicated

by Bieri, et al. (1966) as an information processing variable that helps predict how an individual transforms specified behavioral information into social or clinical judgments. It reflects the relative degree of differentiation of a person's system for construing behavior. Cognitive complexity is defined as the capacity to construe social behavior in a multidimensional way. Scoring methods are described in the next section.

In contrast to the relatively stable and bipolar characteristics of field independence and cognitive complexity, integrative complexity is based upon a developmental stage theory. The stages are hierarchically integrated, form an invariant succession in individual development, involve qualitative differences in modes of thinking between stages and an underlying holistic organization of thought at each stage (Witherell, 1978). One way in which development is conceptualized is as a movement from a concrete to an increasingly abstract conceptual system. According to this theory, more concreteness represents minimal differentiation and little or no integration of concepts, while the opposite is true of more abstractness. The greater an individual's abstractness, the greater his or her ability to consider alternatives, to transcend immediacy and to relate facets of the world in terms of their interrelatedness (Harvey, Hunt and Schroder, 1961). The PCT used to measure integrative complexity involves elaborating six sentence stems into paragraphs (i.e. "When I am criticized . . .", "When I am told what to do . . .").

In addition to these three measures, two items were constructed for use in student evaluation of instruction questionnaires that began with the sentence stem: "I learn best when the task is to . . ." and provided various alternatives for the students to rank order. One item provided five alternatives fashioned after Bloom's (1956) taxonomy of the cognitive domain including knowledge of facts, understanding of basic prin-

ciples, application, analysis and synthesis. The other item provided three alternatives fashioned after Perry's (1968) developmental theory including the stages of dualism, multiplicity and commitment. These two items (see Table 1 for description) were included within an evaluation of instruction questionnaire developed from studies by Doyle & Whitely (1974).

The above-mentioned theories share in common the concept of a more highly differentiated sense of self on one end of the continuum. Field dependence-independence theory does not seem to address multidimensionality of social judgment as distinctly as does cognitive complexity, and neither of these two cognitive styles includes the ability to integrate concepts as does integrative complexity. Both field independence and high integrative complexity include the ability to analyze and to abstract. These differences and overlaps between theories make for interrelationships that are difficult to interpret. The rather sparse empirical research along these lines, using measures similar to those used in this study, indicate cognitive complexity to be negatively correlated with integrative complexity (Epting and Wilkins, 1974), and field-dependence-independence to be nonsignificantly correlated with integrative complexity (Stewin, 1976) and with cognitive complexity (Elliott, 1961).

Because the GEFT, PCT, Rep Test and cognitive developmental items involve very different types of tasks (drawing geometric figures, writing paragraphs, rating people, and ranking preferences), we would expect little or no method variance. And since these measures are based on somewhat differing theories as well, we would expect intercorrelations between them to be low. However, since the cognitive developmental items and the PCT are each based upon developmental schemas, we would expect significant correlations between them.

Method and Procedures

Subjects for this study were 497 male and female lower division college and university students from 18 classes in the physical sciences, social sciences and humanities. The measures were administered during three different sessions during the academic quarter.

GEFT

Subjects were instructed to mark an "x" on the page of the GEFT where they were working when four minutes were announced as passed on each section, and to stop drawing when five minutes was announced as passed. This allowed us to investigate the possibly more discriminating time limit of four minutes per section for college students. Since correlations between four- and five-minute scores were all in the high .90s, the traditional five-minute scores were used in the inter-test analyses.

PCT

The PCT was scored by the Ontario Institute for Studies in Education. Interrater reliability for the first forty PCTs was in the high .80s. Summary scores of integrative complexity based upon the PCT were derived from averaging the top three paragraph scores as recommended by Hunt, et al. (1978). This score was then compared with an average of all six paragraph scores, yielding a correlation of .95. This confirmed the adequacy of the traditional scoring method which was then used for the inter-test analyses.

Rep Test

The Rep Test was scored by the traditional method described above of Bieri's "d", with higher scores indicating greater cognitive complexity. Because the psychometric rigour of this scoring method seemed questionable to us, we also scored the Rep Test by a variety of other methods (Trabin, T., Doyle, K., and Wood, P., Note 1) from which two were selected as alternative

measures of cognitive complexity. The first alternative measure is the average of the standard deviations for each column (the ratings for each role on ten constructs); this measure gives the added information of degree of difference in ratings while Bieri's "d" gives only the number of different ratings. The second alternative measure used is the first eigen value derived from a factor analysis of the rows (ratings for each construct across ten roles) for each individual. The greater the first eigen value, the more unidimensional (or cognitively simple) a subject's social judgment; this method would seem to most accurately fit the definition of cognitive complexity as multidimensionality of social judgment. A detailed discussion of the statistical properties within and interrelationships between these and other Modified Rep Test Grid scoring procedures may be found in Trabin, T., Doyle, K. and Wood, P (Note 1).

Cognitive Developmental Items

Subjects were asked to rank their favored way of learning as "1", the next best as "2", etc. on the five alternatives in the Bloom-based item and the three alternatives in the Perry-based item. The alternatives were then weighted with those higher in the developmental hierarchy given lower weights. In this way, if ranking is multiplied by weighting for each alternative and then added together , subjects with the highest scores will be those whose learning preferences require the highest cognitive development. This scoring procedure also yields a broader range of scores suitable for correlational analyses (see table 2).

Inter-Test Analyses

Scores for the GEFT, PCT, three Rep Test measures, and two evaluation questionnaire items on cognitive development were analyzed for systematic relationships within a correlation matrix. Also included in the matrix were data on subjects' sex, GPA and standard deviation of ratings on twelve

instructor evaluation items (see table 3).

Results

Results indicate no systematic relationships between any of the cognitive style or cognitive developmental measures and sex or GPA with the exception of a significant positive correlation between the PCT and GPA. The PCT is the only measure in this study requiring verbal facility, and this facility is probably an important contributing factor to the .249 ($p < .000$) correlation between GPA and PCT.

Both cognitive developmental items showed mild correlations of .15 ($p < .003$) and .14 ($p < .006$) with the standard deviation of instructor ratings; this suggests a relationship between cognitive development and differentiatedness of student evaluations of their instructor across different teaching dimensions. Two of the Rep Test measures (Bieri's "d" and the average standard deviation of ratings for each role) both showed statistically significant correlations of .22 ($p < .000$) with differentiatedness of student evaluation of instructor ratings. These results replicate Wright and Richardson's (1977) findings, and recalls their argument that cognitive complexity is an important response style to be further studied. The GEFT obtained a negative correlation of -.18 ($p < .002$) with differentiatedness of evaluation ratings, suggesting that the more field dependent subjects may be more sensitized to and therefore more able to differentiate between differing instructor behaviors. This is congruent with the theory that field dependent persons are more attentive to and able to remember social stimuli than are field independent persons (Witkin, 1976).

Analyses of relationships between different measures of cognitive

style and cognitive development indicated no systematic relationships between cognitive style and cognitive developmental measures, which supports Loevinger's (1966) assertion that developmental stage and bipolar trait concepts have fundamental differences that prevent easy comparison. The two cognitive style measures (GEFT and Rep Test) were not significantly correlated with eachother, which may be because they measure different traits or because they involve different tasks.

In contrast to the above-mentioned paucity of inter-test relationships, the three cognitive developmental measures (PCT and two cognitive developmental items) were significantly correlated with one another despite the different tasks they involved. The two items correlated .35 with eachother ($p < .000$), and .16 with PCT ($p < .008$). This clearly indicates stronger relationships of cognitive developmental measures with eachother than with cognitive style measures. The results also suggest that the items may provide simpler measures of developmental constructs similar to those measured by the PCT.

Conclusions

This study explored the potential incorporation of cognitive developmental concepts into student evaluation of teaching questionnaires by investigating relationships obtained between representative developmental items and other measures within a nomological net of cognitive structure. Initially, alternative methods for scoring the GEFT, PCT and Rep Test were explored. It appears that standard deviation and factor analytic scoring methods for the Rep Test provide important scoring alternatives to Bieri's "d".

Cognitive style and cognitive developmental measures were unrelated to sex and GPA with the exception of the PCT which was positively related to GPA. The verbal facility involved in the PCT tasks may explain this

relationship.

Cognitive development measured by constructed itmes and cognitive complexity measured by two Rep Test scoring procedures were positively related to differentiatedness of student ratings of their instructor as was field dependence (measured by lower scores on the GEFT). These results support previous studies' contentions that cognitive complexity is a response style pertaining similarly to Rep Test and student evaluation of instruction tasks, and that field dependent subjects are more sensitive to and able to differentiate between differing social stimuli. The results also suggest the potential usefulness and applicability of cognitive developmental items in student evaluation of instruction questionnaires.

Inter-test analyses indicated no significant interrelationships between measures except between all three cognitive developmental measures with each other. This is one approach to validating and clarifying the meaning of these items. An additional approach has been explicated in Church, A.T., Doyle, K., and Trabin, T. (Note 1). Many of the moderate correlations obtained in this study were of high statistical significance only because of the large number of subjects and actually accounted for about four per cent of the variance. This is not uncommon in personality research, especially within the area of cognitive structures. The constructs being assessed by the different measures used in this study, and especially the relationships between them, need more definitive explication on both theoretical and empirical grounds in the literature. The results of this study offer some support to the theoretical differences between stage theories of cognitive development and bipolar trait concepts of cognitive style. The results also suggest that cognitive developmental items can be constructed that are effective and meaningful measures which can be easily incorpor-

ated into student evaluation of instruction questionnaires. It is recommended that additional items be developed and further refined, and that their nomological net be further elaborated with experiments of their influence on student evaluation of teaching.

Acknowledgements

We wish to acknowledge the assistance of Clyde Parker with the conceptual development of this study, and Auke Tellegen and Phillip Wood with the data analysis.

Reference Notes

1. Church, A.T., Doyle, K., & Trabin, T. Student learning preferences and instructor evaluations. Submitted for publication, 1981.

2. Trabin, T., Doyle, K., & Wood, P. Cognitive complexity: Alternative scoring procedures for the modified Rep Test Grid. In preparation for submission for publication.

Reference List

Bieri, J., Atkins, A.L., Briar, S., Leaman, R.L., Miller, H., & Tripodi, T. Clinical and social judgment: The discrimination of behavioral information. New York: Wiley, 1966.

Blommers, P. Comments on Professor Messick's paper. In M.C. Witrock & Wiley, D.E. (Eds.), The evaluation of instruction. New York: Holt, Rinehart & Winston, 1970, 202-204.

Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H. & Kratwohl, D.R. Taxonomy of educational objectives: Handbook 1: Cognitive domain. New York: David McKay, 1956.

Crockett, W.H. Impressions of a speaker as a function of set to understand or to evaluate, of cognitive complexity, and of prior attitudes. Journal of Personality, 1975, 43, 168-178.

Cronbach, L.J. The two disciplines of scientific psychology. American Psychologist, 1957, 12, 671-684.

Cronbach, L.J. & Snow, R.E. Aptitudes and instructional methods. New York: Irvington Publishers, Inc., 1977.

Doyle, K. & Whitely, S. Student ratings as criteria for effective teaching. American Educational Research Journal, 1974, 7, 259-274.

Elliott, R. Interrelationships among measures of field dependence, ability and personality traits. Journal of Abnormal and Social Psychology, 1961, 63, 27-36.

Epting, F., & Wilkins, G. Comparison of cognitive structural measures for predicting person perception. Perceptual and Motor Skills, 1974, 38, 727-730.

Harvey, O.J., Hunt, D.E., and Schroder, H.M. Conceptual Systems and Personality Organization. New York: John Wiley and Sons, Inc., 1961.

Hunt, D.E., Butler, L.F., Noy, J.E., & Rosser, M.E. Assessing conceptual level by the paragraph completion method. Ontario: The Ontario Institute for Studies in Education, 1978.

Jenkins, J. New theoretical guidelines for attribute-treatment interaction research. Paper presented at the 1977 Annual APA convention.

Kelly, G.A. The psychology of personal constructs. New York: Norton, 1955.

Loevinger, J. The meaning and measurement of ego development. American Psychologist, 1966, __, 195-206.

Messick, S. The criterion problem in the evaluation of instruction: Assessing possible, not just intended, outcomes. In Witrock, M.C. & Wiley, D.E. (Eds.), The evaluation of instruction: Issues and problems. New York: Holt, Rinehart & Winston, 1970.

Packer, J. & Bain, J.D. Cognitive style and teacher-student compatability. Journal of Educational Psychology, 1978, 70, 864-871.

Perry, W.G. Forms of intellectual and ethical development in the college years: A scheme. New York: Holt, Rinehart and Winston, Inc., 1968.

Stewin. Integrative complexity: Structure and correlates. Alberta Journal of Educational Research, 1976, 22, 226-236.

Witherell, C. A structural-developmental analysis of teachers' conceptions of teaching and human development in relation to patterns of teaching behavior: Five case studies. Unpublished doctoral dissertation, University of Minnesota, 1978.

Witkin, H.A., Goodenough, D.R., and Karp, S.A. Stability of cognitive style from childhood to yo-ng adulthood. Journal of Personality and Social Psychology, 1967, 7, 291-300.

Witkin, H.A., Moore, C.A., Goodenough, D.R., and Cox, P.W. Field-dependent and field-independent cognitive styles and their educational implications. Review of Educational Research, 1977, 47, 1-64.

Wright, R.J. & Richardson, L. The effect of response style on cognitive complexity and course evaluation. Educational and Psychological Measurement, 1977, 37, 177-183.

Table 1

Cognitive Developmental Items

1. I learn best when the task is to (please rank these answers from 1-5 with 1=best, 2=next best, etc.):

___ show an idea can be applied to an actual situation

___ master a set of concrete facts or a body of information

___ closely examine abstract ideas and theories

___ understand basic principles about a subject

___ put together several differing ideas into a theory that makes sense to me

2. I learn best when the task is to (please rank these answers from 1-3 with 1=best, 2=next best, etc.):

___ take a position after considering many ways of interpreting a difficult problem

___ figure out the one right answer to a straightforward problem

___ choose the best answer from many possible ones which seem equally good

Table 2

Descriptive Statistics for the Measures

Measures	N	Mean	S.D.	Minimum	Maximum
Item #1	383	41.39	6.89	15.00	71.00
Item #2	382	11.64	1.75	6.00	20.00
PCT	370	16.77	4.02	8.00	30.00
GEFT	347	12.80	4.61	0.00	18.00
Rep "d"	315	301.44	36.33	166.00	367.00
Rep s.d.	315	1.03	.24	.39	1.88
Rep factor	315	3.85	.73	1.91	7.20
Rating s.d.	389	.76	.25	0.00	1.91
GPA*	249	2.53	.95	1.00	5.00
Sex	268	1.47	.50	1.00	2.00

* GPA of 3.6-4.0 was coded "1", with lower GPA groupings coded with numbers increasing to 5.

Table 3

Intercorrelation Matrix for Measures†

Measures	Item #1	Item #2	PCT	GEFT	Rep "d"	Rep s.d.	Rep Factor	Rating s.d.	GPA	Sex
Item #1	--									
Item #2	.35***	--								
PCT	.16*	.16*	--							
GEFT	-.05	-.07	-.06	--						
Rep "d"	.04	-.06	-.04	-.01	--					
Rep s.d.	.09	.04	-.07	-.03	.57***	--				
Rep factor	-.06	.02	.00	.06	.18**	.27***	--			
Rating s.d.	.15**	.14*	-.08	-.18**	.22**	.22**	.04	--		
GPA****	.06	.09	-.25	-.03	.04	.03	-.02	-.02	--	
Sex	.07	.04	.00	.15	.11	.08	.05	-.13	.13	--

† Pairwise deletion results in N's for each correlation of between 186 and 389.

* $P < .01$

** $p < .005$

*** $p < .001$

**** Because of the GPA coding described in Table 2, the above correlations show reversed signs from what would obtain if actual rather than coded GPAs had been used.