

ROLE CATEGORY QUESTIONNAIRE MEASURES OF COGNITIVE COMPLEXITY: RELIABILITY AND COMPARABILITY OF ALTERNATIVE FORMS

Daniel J. O'Keefe, Gregory J. Shepherd, and Thomas Streeter

COGNITIVE complexity has received considerable research attention as a potential determinant of communicative functioning. As commonly used, the concept refers to the relative number of constructs in a person's interpersonal construct system, and hence the variable might be more lucidly labeled "construct differentiation."¹ Whatever the label, the bulk of communication research in this area has employed Crockett's Role Category Questionnaire (RCQ) as the basis for the measure of complexity (differentiation).² The most common form of the RCQ is the "two-peer" version, in which respondents write impressions of two persons they know well, one whom they

like and one whom they dislike. The RCQ-based measure of construct differentiation is obtained by counting the number of constructs across the two impressions.³

This measure has been found to be significantly associated with a variety of indices of developed communicative functioning, with the research evidence ranging across age groups (children, adolescents, adults), communication situations (persuasive, feeling-centered, referential, regulative), and specific criterion measures (message strategies, message rationales, number of persuasive arguments, etc.). The instrument has also been found to be associated with a number of other indices of developed social cognition (for example, measures of social perspective-taking, construct abstractness, and the like); and, past early childhood, the measure appears to be unrelated to verbal ability, intelligence, verbal fluency, vocabulary, writing speed, and the like.⁴

Despite the frequency of the instru-

Daniel J. O'Keefe is an Assistant Professor, and Gregory J. Shepherd and Thomas Streeter are graduate students, in the Department of Speech Communication, University of Illinois at Urbana-Champaign. A version of this paper was presented at the International Communication Association convention, Minneapolis, Minn., 1981.

¹ See, for example, J. R. Adams-Webber, *Personal Construct Theory: Concepts and Applications* (New York: Wiley, 1979); Tony Tripodi and James Bieri, "Information Transmission in Clinical Judgments as a Function of Stimulus Dimensionality and Cognitive Complexity," *Journal of Personality*, 32 (1964), 119-37; James Bieri, "Cognitive Complexity-Simplicity and Predictive Behavior," *Journal of Abnormal and Social Psychology*, 51 (1955), 263-68.

² Walter H. Crockett, "Cognitive Complexity and Impression Formation," in *Progress in Experimental Personality Research*, ed. Brandan A. Maher, II (New York: Academic Press, 1965), 47-90.

³ The scoring procedures are described in Walter H. Crockett, Allan N. Press, Jesse G. Delia, and Charles T. Kenny, "The Structural Analysis of the Organization of Written Impressions," unpublished manuscript, University of Kansas, 1974.

⁴ The research bearing on these claims is reviewed and summarized in Daniel J. O'Keefe and Howard E. Sypher, "Cognitive Complexity Measures and the Relationship of Cognitive Complexity to Communication," *Human Communication Research*, 8 (1981), 72-92.

ment's use in communication research, however, two important methodological issues concerning the measure have received little empirical attention. The first concerns test-retest reliability. Differentiation is commonly conceived of as a relatively stable individual difference among adults; the number of constructs in a person's interpersonal construct system is not thought to exhibit wide variation over short periods of time.⁵ Thus, a finding that RCQ-based differentiation scores did not exhibit temporal stability (high test-retest reliability) would necessitate either conceptual revision (that is, treating differentiation as a temporally variable characteristic rather than as a relatively stable one) or abandonment of the RCQ as a basis for assessing differentiation. In either case, a reinterpretation of previous research findings would be called for. Presently, when individuals who differ in RCQ-based differentiation scores have been found to also differ in, for example, the kind of persuasive strategies they employ, that difference in persuasive strategy use has been attributed to underlying stable differences in the degree of differentiation of the interpersonal construct system.⁶ Such interpretations would obviously be

⁵ Ibid.; Adams-Webber.

⁶ For examples of such interpretations, see Jesse G. Delia and Ruth Anne Clark, "Cognitive Complexity, Social Perception, and the Development of Listener-Adapted Communication in Six-, Eight-, Ten-, and Twelve-Year-Old Boys," *Communication Monographs*, 44 (1977), 326-45; Jesse G. Delia, Susan L. Kline, and Brant R. Burleson, "The Development of Persuasive Communication Strategies in Kindergartners Through Twelfth-Graders," *Communication Monographs*, 46 (1979), 241-56; Barbara J. O'Keefe and Jesse G. Delia, "Construct Comprehensiveness and Cognitive Complexity as Predictors of the Number and Strategic Adaptation of Arguments and Appeals in a Persuasive Message," *Communication Monographs*, 46 (1979), 231-40; Jesse G. Delia, Ruth Anne Clark, and David E. Switzer, "The Content of Informal Conversations as a Function of Interactants' Interpersonal Cognitive Complexity," *Communication Monographs*, 46 (1979), 274-81.

undermined by a finding of low test-retest reliability for RCQ-based differentiation scores. To date, the only published test-retest reliability estimate is Crockett's four-month figure of .95.⁷ Although that figure is remarkably high, additional evidence would obviously be desirable, especially because Crockett's study used an early version of the RCQ which had subjects write eight (not two) impressions, whereas most recent research using the RCQ has employed the two-peer version. The two-peer version can be expected to yield lower reliability estimates for the differentiation measure (because of the shorter "test length"), but no evidence exists concerning whether the two-peer reliability estimates are unsatisfactorily low.

The second methodological issue concerns the time allotted for completion of the RCQ. For each impression, the standard RCQ form asks respondents "please do not spend more than five minutes describing this person," but in practice this time limit is not always strictly enforced. Thus, the question arises whether substantially different results would obtain under conditions of such strict time monitoring.

In part this question is connected to the frequently-voiced concern that factors such as verbal ability and vocabulary may be contaminants of RCQ-based differentiation scores, for such contamination could be exacerbated by a lack of strict time monitoring.⁸ A finding of low correlation between differentiation scores from

⁷ Crockett.

⁸ Such concerns have been voiced by, for example, Alan Miller and Paula Wilson, "Cognitive Differentiation and Integration: A Conceptual Analysis," *Genetic Psychology Monographs*, 99 (1979), 27; and William G. Powers, William J. Jordan, and Richard L. Street, "Language Indices in the Measurement of Cognitive Complexity: Is Complexity Loquacity?" *Human Communication Research*, 6 (1979), 69-73.

monitored and unmonitored RCQs might indicate the operation of such contaminating factors, suggesting caution in the use of unmonitored RCQs. Beyond this, such a finding would point to the need for a careful reconsideration of earlier research findings using the RCQ as a basis for measures of differentiation. At present, when investigators review previous research employing RCQ-based indices of differentiation, they do not distinguish results based on monitored RCQs from those of unmonitored RCQs.⁹ Such a distinction would be important in interpreting previous research if monitored and unmonitored RCQs do not yield comparable differentiation indices.

In light of these issues, the present research project (1) investigated the test-retest reliability of the construct differentiation measure based on the two-peer RCQ, using both "timed" (that is, strictly timed) and "untimed" (that is, without experimenter enforcement of the time limit) versions of the RCQ, and (2) compared the timed and untimed versions of the two-peer RCQ as bases for measures of differentiation. Subjects completed the RCQ twice, at an interval of four weeks. In one condition, both sessions were timed; in a second condition, neither administration was timed; in a third condition, one session was timed and the other was not, with the order (timed-untimed vs. untimed-timed) varied. The first two conditions permitted direct examination of the reliability of the measure in, respectively, timed and untimed versions. The third condition permitted examination of the equivalence of the two versions through the correlation of subjects' timed and untimed scores (regardless of order of completion).

⁹ See, for example, O'Keefe and Delia; Adams-Weber; O'Keefe and Sypher.

METHOD

Subjects

Subjects were 103 volunteers from undergraduate communication classes at a midwestern university. All sessions were held during regular class meetings.

Procedures

Subjects completed the two-peer version of the Role Category Questionnaire twice, at an interval of four weeks (range of interval from 26 to 30 days, with a mean interval of 28.75 days). For 30 subjects, both administrations were timed; for 22 subjects, both administrations were untimed; and for 51 subjects, one administration was timed, and the other was not. For 25 subjects the first session was timed, and the second untimed; for 26 subjects the first was untimed, and the second timed.

In all administrations, the standard cover-sheet instructions for the RCQ identified the investigators' concern as being "with the habits, mannerisms—in general, with the personal characteristics, rather than the physical traits—which characterize a number of different people." Subjects were then asked to identify (by initials, nicknames, etc.) "a person your own age whom you like" and "a person your own age whom you dislike," and were asked to "spend a few moments thinking about these people, mentally comparing and contrasting them." On the succeeding pages, subjects were asked to describe first the liked, then the disliked other. For each description they were asked to "describe this person as fully as you can," and to "pay particular attention to the person's habits, beliefs, ways of treating others, mannerisms, and similar attributes." For each description these instructions included the statement, "Please do not spend more than five (5) minutes describing this person."

In the untimed administrations, this

written time-limit instruction was not enforced by the experimenter. Subjects read the cover-sheet instructions, proceeded to the descriptions, and took as much time as they desired to complete the task.

In the timed administrations, the cover-sheet instructions included, at the bottom of the cover sheet, the statement "do not turn the page until instructed to do so." When all subjects in a class indicated they had read and completed the cover sheet, they were instructed to turn the page and begin the first description. After five minutes, they began the second description. After another five minutes, the questionnaires were collected.

For the second session in all conditions, subjects were told to approach the task as though they had not seen the questionnaire before. Thus, subjects were not restricted to, nor enjoined from, describing the same persons they had described in the first session. The possibility that subjects might describe different persons in the two sessions made for a more rigorous test of both the reliability and the equivalence of differentiation measures based on the timed and untimed versions of the RCQ.¹⁰

Following the procedures of Crockett et al.,¹¹ a trained coder counted the number of different constructs used to describe each peer; in this coding system, aspects of the other's personality and behavior are counted, while physical

¹⁰ It was not possible to determine exactly how many second-session descriptions were of the same persons as described at the first session, because subjects were asked to avoid providing full names for the persons described (but instead to use nicknames, initials, special symbols, etc.). Based on an examination of the descriptions and of the identifying information, however, it appeared that approximately 50 to 60 percent of the second-session descriptions were of different persons than those described at the first session.

¹¹ Crockett et al.

characteristics are not. A subject's differentiation score consisted of the sum across the two descriptions. A second trained coder independently scored 15 randomly-selected protocols, yielding an interrater reliability coefficient by Pearson correlation of .97.

RESULTS

Reliability Estimates

The four-week test-retest reliability estimates for the construct differentiation index were high for both the timed ($r = .84$, $df = 28$, $p < .001$) and the untimed ($r = .86$, $df = 20$, $p < .001$) versions of the two-peer RCQ.

Comparison of Timed and Untimed Versions

Differentiation scores based on timed and untimed RCQs were highly correlated, $r = .84$ ($df = 49$, $p < .001$). Not unexpectedly, the means and standard deviations for scores from the two versions were different. Based on the 51 subjects who completed both a timed and an untimed version, the mean for scores derived from the timed version was 22.59, with a standard deviation of 5.79; and the mean for scores derived from the untimed version was 25.39, with a standard deviation of 7.23.

These means and standard deviations, however, as based on both first- and second-session scores; as might have been expected, readministration of the RCQ tended to depress differentiation scores [for the total sample of 103, the mean across timed and untimed versions for the first session was 24.49, and the corresponding mean for the second session was 23.21; these means are significantly different, t ($df = 102$) = 2.56, $p < .05$]. Hence, the most illuminating comparison of means and standard deviations for timed and untimed versions is that based on first-session questionnaires

alone, independent of any readministration effects. For all first-session timed questionnaires ($n = 55$), the mean was 22.62 and the standard deviation was 7.11; for all first-session untimed questionnaires ($n = 48$), the mean was 26.63 and the standard deviation 9.28. These means are significantly different [t ($df = 101$) = 2.48, $p < .05$], as are the variances [F ($df = 47,54$) = 1.71, $p < .05$]. The greater variability of the untimed-RCQ scores is, as the higher mean suggests, largely a consequence of a larger number of high scores; for the 55 first-session timed RCQs, only one differentiation score exceeded 35, but for the 48 first-session untimed RCQs, seven scores were greater than .35.

DISCUSSION

These results further bolster one's confidence in Crockett's standard two-peer Role Category Questionnaire as a basis for indices of interpersonal construct differentiation. Satisfactorily high four-week test-retest reliability estimates were obtained for differentiation measures based either on the timed (.84) or the untimed (.86) versions of the RCQ.

Moreover, the timed and untimed versions yielded highly correlated measures of differentiation. The correlation between scores based on the two versions (.84) is especially striking given that the two versions were not completed in a single session, but at an interval of four weeks; the presence of even slight unreliability in either version would mitigate against a finding of substantial correlation between the two.¹² This

¹² Cohen and Cohen provide a formula for producing an unreliability-corrected coefficient from the obtained correlation between two imperfect measures (the numerator) and the square root of the product of the two reliability estimates (the denominator); see Jacob Cohen and Patricia Cohen, *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (Hillsdale, New Jersey: Lawrence Erlbaum, 1975), p. 63. In the present case, the

point can be expressed more precisely by considering the index of reliability, which is the square root of a reliability coefficient; "the index of reliability . . . is often used as an indication of the upper limit of correlation of any variable with another."¹³ The index of reliability for the timed-RCQ-based differentiation measure is .917—which suggests that even had the timed and untimed RCQs been completed in a single session, the correlation of timed and untimed differentiation scores would probably not have exceeded .917.¹⁴

It might also be noted that the high correlation of scores based on the timed and untimed versions is consistent with previous findings indicating only weak associations between RCQ-based differentiation scores and independent assessments of verbal ability, vocabulary, verbal intelligence, and the like.¹⁵ If verbal abilities significantly contributed to RCQ-based differentiation scores, one would expect a rather low association between scores based on untimed RCQs (where verbal ability would presumably

corrected coefficient for the relation of timed and untimed differentiation scores is .989. This corrected coefficient should be interpreted with caution, however, because the assumption of uncorrelated error is almost certainly false.

¹³ J. P. Guilford and Benjamin Fruchter, *Fundamental Statistics in Psychology and Education*, 5th ed. (New York: McGraw-Hill, 1973), p. 401.

¹⁴ If .917 is the upper limit for the correlation between timed differentiation scores and any other variable, then no obtained correlation can be expected to show more than 84 percent of the variance shared between timed differentiation scores and any other variable. The obtained correlation of .84 between timed and untimed differentiation scores yields 70.56 percent shared variance. Thus the obtained correlation may be said to have found 84 percent of the shared variance available to be found (.7056 divided by .84).

¹⁵ Brant R. Burleson, James L. Applegate, and C. M. Neuwirth, "Is Cognitive Complexity Loquacity? A Reply to Powers, Jordan, and Street," *Human Communication Research*, 7 (1981), 212-25; Howard E. Sypher and James L. Applegate, "Cognitive Differentiation and Verbal Intelligence: Clarifying Relationships," *Educational and Psychological Measurement*, in press; O'Keefe and Sypher.

have a large impact) and timed RCQs (where the effects of verbal ability might be minimized). Thus, the observed high correlation further buttresses the independence of Crockett's complexity measure from verbal abilities.

Of course, the untimed version (because it allowed for writing for a longer time) yielded more high differentiation scores, thereby giving a higher mean and a larger standard deviation. But despite these differences, the scores based on the two versions were very strongly related—and hence the present results rather strongly suggest the equivalence of the two versions of the RCQ as bases for the assessment of construct differentiation.

Finally, the results presented here appear to support the conceptualization of

construct differentiation as a relatively stable individual difference among adults. This, in turn, lends support to past interpretations of data which have linked many individual social-cognitive and communicative differences with corresponding differences in interpersonal construct differentiation.¹⁶ The reliability of both timed and untimed RCQ-based differentiation indices, as well as the apparent equivalence of the two versions, suggests that conclusions drawn in previous investigations which have employed the RCQ as a measure of differentiation can be viewed with increased confidence.

¹⁶ For examples of such interpretations, see Delia and Clark; Delia, Kline, and Burleson; O'Keefe and Delia; Delia, Clark, and Switzer.