Interpersonal Competence Instrument: Development and Preliminary Findings

Lawrence J. Stricker
Educational Testing Service

A prototype measure of interpersonal competence, designed to measure effectiveness in dealing with other people, was developed—the Interpersonal Competence Instrument (ICI). The ICI is based on the videotape presentation of scenes of subordinates talking to a superior in a business setting. The examinee takes the role of the superior, his or her tape-recorded replies are assessed for effectiveness and originality, and his or her written judgments are evaluated for accuracy. In a pilot study (1) the three scores for the ICI were reliable, in terms of both interrater agreement and internal consistency; (2) the effectiveness and originality scores for replies were only moderately correlated with the accuracy scores for judgments, but the two replies scores were highly intercorrelated; and (3) the correlations of the scores with background measures and other instruments, including measures of verbal ability and accuracy in social perception, generally supported the ICI's construct validity.

Interest in specialized abilities that come into play in social contexts dates back to Thorndike's (1920) seminal discussion of social intelligence in 1920. The broad spectrum of abilities constituting this domain—Guilford's (1967) structure of intellect theory envisages a total of 30 distinct kinds of abilities involving "behavioral content"—has been studied under such disparate rubrics as social intelligence, social perception,

empathy, and role taking. One of the most important abilities of this kind is interpersonal competence—effectiveness in dealing with other people. This construct has been the focus of theorists in psychology (Argyle, 1969, 1972; Argyris, 1962, 1964; White, 1959), in psychiatry (Adler, 1941; Sullivan, 1953), and in sociology (Foote & Cottrell, 1955; Weinstein, 1969).

A wide range of methods for measuring variables in this general area has been devised by workers in such diverse fields as ability testing (Hendricks, Guilford, & Hoepfner, 1969; O'Sullivan & Guilford, 1975), cognitive development (Flavell, Botkin, Fry, Wright, & Jarvis, 1968), social perception (Cline, 1964), experiential learning (Breen, Donlon, & Whitaker, 1977), behavior therapy (Eisler, 1976), and management assessment (Crooks, 1976). The methods include ability tests, self-report inventories, situational tests, role-playing procedures, self-ratings and ratings by others, and interviews. Two problems affect many of the measures. One drawback is that performance on the tests and on similar devices tends to be related to general ability, making it questionable that a trait distinguishable from intelligence is being measured (Walker & Foley, 1973). A second problem is that the measures are apt to be cumbersome or uneconomical, precluding their large-scale use.

The advent of videotape technology raises new possibilities for assessing interpersonal compe-

69

tence because videotape provides a means of portraying social situations in a comprehensive, standardized, and economical manner, and the examinee's responses upon viewing the scenes can then be evaluated. Videotape and motion pictures, of course, convey more information than print or any other communications medium, for they record changes in facial expressions, postural movements, and other potentially important behaviors. The research reviewed by Guilford (1967) as well as the work by Guilford and his colleagues (Hendricks, Guilford, & Hoepfner, 1969; O'Sullivan & Guilford, 1975) emphasizes the importance of visual stimuli in assessing interpersonal variables.

A videotape-based measure also shares the advantages of other situational tests that resemble real-life situations (Cronbach, 1970). Tests of this kind are (1) face valid, and hence more apt to engage the interest and motivation of the examinees; (2) "work" samples, similar to the activities of the test takers and thus more likely to be predictive of their behavior; and (3) complex in nature, making it difficult for the examinees to ascertain precisely what is being evaluated and, hence, reducing the probability that they will purposely distort their responses to give an unrealistically favorable impression.

The aim of this research was (1) to develop a prototype measure of interpersonal competence, the Interpersonal Competence Instrument (ICI), based on videotape presentations of a common kind of social situation; and (2) to obtain preliminary data on the ICI's functioning.

Planning and Development of Instrument

In planning the ICI, related work in the literature was reviewed, particularly the efforts of Guilford and his associates (Guilford, 1967; Guilford & Hoepfner, 1971; Hendricks et al., 1969; O'Sullivan & Guilford, 1975), Frederiksen and his co-workers (Frederiksen, Saunders, & Wand, 1957; Frederiksen & Ward, 1978; Frederiksen, Ward, & Samph, 1979), Argyris (1962, 1964), Argyle (1969, 1972), Bales (1950,

1970), and Kagan (1973); and authorities on interpersonal processses, nonverbal behavior, management assessment, and simulation measures were consulted.

On the basis of these deliberations, the basic features of the ICI were decided upon:

1. The social situation is that of a subordinate talking to a superior in a business setting. This situation is extremely common, permits a range of variations, and simplifies the preparation of scenarios as well as the assessment of examinees' responses.

2. An actor, playing the subordinate, talks directly to the camera (and, hence, to the viewer), as though speaking to the superior; the examinee, watching the videotape, takes the role of the superior. This relatively direct communication between the actor and the examinee maximizes resemblance to real life.

3. The actor-subordinate makes a single remark, appropriate to the beginning or middle of a conversation, after which the examinee gives his or her sole response. A longer segment of a conversation is impractical, for it is impossible for the actor-subordinate to make subsequent comments pertinent to the examinee's replies.

4. Several scenes are employed, systematically differing in the kinds of subordinates and tasks involved. This use of a variety of scenes makes it possible to evaluate the generality of the responses and enhances the reliability of scores based on them.

5. The examinee-superior has the same job throughout the entire testing situation; and all of the relevant features of the job, including background information about the subordinates and about the tasks presented in the scenes, are described in detail. Using the same job throughout permits the examinee to retain the same mental set in all of the scenes and facilitates his or her taking the required role. Providing details about the job increases realism and minimizes the tendency for examinees to make unnecessary

assumptions and thereby introduce irrelevant variation into their responses. This general approach is the same as the one used by Frederiksen et al. (1957) with the in-basket test.

6. The examinee-superior makes two different kinds of responses. For one set of scenes, he or she actually replies to the actor-subordinate by talking into a tape recorder. This arrangement is close to what happens in an actual conversation and ensures that the responses are spontaneous. For an equivalent set of scenes, the examinee makes judgments about the pertinent characteristics of the subordinate and about other features of the situation by writing down free responses to open-ended questions. The two kinds of responses are obtained from different scenes in order to minimize the possibility that one kind may affect the other.

7. Using expert raters, the replies are evaluated for effectiveness and originality, and the judgments are assessed for accuracy. The effectiveness of the replies, their originality, and the accuracy of the judgments correspond to three kinds of abilities that involve behavioral content in Guilford's (1967) theory: (1) convergent production (coping with the behavior of others in a correct way), (2) divergent production (coping with the behavior of others in a creative way), and (3) cognition (understanding the behavior of others). Global ratings by experts are a temporary expedient chosen as the most direct way of measuring these variables at this early juncture.

General instructions, including information about the superior's job, and 32 scenes with related background material, involving a range of subordinates and tasks, were prepared. The instructions, the scenes, and the background material were pretested, in paper-and-pencil form, both with individual office workers and with small groups of these office workers as well as with two classes of undergraduate business students. The material was successively revised on the basis of the subjects' reactions. This effort focused on making certain that the instructions were clear and that the scenes and the other information provided sufficient details for the examinees to give meaningful responses.

Ten scenes were videotaped in color. Each showed the actor, seated or standing in an office and directly facing the camera, in three-quarter view, with his or her head, torso, hands, and arms in full sight. Nine scenes were selected for use, one as a practice exercise. The instructions and background information for the scenes were narrated and added to the videotape. The videotape and accompanying questionnaire for the judgments section were pretested individually with office workers, and minor revisions were made in the questionnaire.

In brief, the final form of the ICI, entitled "Working with People," consists of general instructions that ask the examinee to take the role of the general manager of a medium-size company and to respond to each scene as though he or she were the manager. The essential elements of the manager's duties and the nature of the firm are described in nontechnical language. Additional instructions are provided for the two sections of the test, entitled "Making Replies" and "Making Judgments."

In the Replies section, the examinee orally replies to the subordinate, talking into a tape recorder. In the Judgments section, he or she makes two written judgments about the situation and the subordinate on an open-ended questionnaire. One judgment is "What is the main task facing you in the situation? In other words, what is the most important thing that you need to do in the present conversation?" The other judgment is "What are the thoughts, feelings, and intentions of the subordinate that are most relevant to the situation? In other words, what are the subordinate's characteristics of this kind that have an important bearing on the present conversation?"

Both sections consist of five exercises based on a scene and its background material: a practice exercise (the same one is used in each section)

and four actual exercises (different for the two sections). Each set of four actual exercises is roughly equivalent, being comprised of two scenes with male subordinates and two with female subordinates, one of each pair involving a task judged to be basically positive in nature (e.g., commending good work) and the other a task that seems more or less negative (e.g., correcting poor work). The characteristics of the exercises are listed in Table 1. As an illustration, the practice exercise appears in Figure 1.[1] The total time required to administer the ICI is approximately 35 minutes.

Each response to the actual exercises in the Replies section is rated on a 5-point scale for effectiveness (Not at all Effective—1 to Extremely Effective—5) and originality (Not at all Original—1 to Extremely Original—5). Effectiveness is defined as "the extent to which the reply contributes to dealing with the task in the conversation," and originality as "the extent to which the reply is unusual and imaginative." Each response to the actual exercises in the Judgments section (the two judgments for an exercise being considered as a unit) is rated for accuracy on a 5-point scale (Not at all Accurate—1 to Extremely Accurate—5). Accuracy is defined as "the extent to which the responses correctly describe all of the key characteristics of the subordinate and the task in the conversation."

Three total scores are derived from the ratings: Replies—Effectiveness, Replies—Originality, and Judgments—Accuracy; each score is the sum of the corresponding ratings for the exercises.

Pilot Study

In order to appraise the intercorrelations of the ICI's exercises and total scores, their reliability, and, at least to a limited extent, the validity of the scores as gauged from their correlations with several other relevant variables, a

pilot study was carried out. This study was done in conjunction with an independent investigation, by Rapaczynski (1979), of the relationship between social perception and field independence-dependence cognitive style. The general aims of the Rapaczynski investigation were sufficiently similar to the goals of the ICI pilot work that most of the measures in her research were pertinent to these efforts.

Subjects

The subjects were college-age women, paid volunteers recruited through advertisements in a local newspaper. The entire battery of instruments was administered to 58 women, but because of various technical problems in collecting and preparing the data for analysis, mainly subjects' failure to follow instructions, usable data were only available for 50 women for the ICI Replies section and 56 women for the ICI Judgments section. The 58 women's ages ranged from 18 to 27 years, with a mean of 20.5; their education varied from 10 to 20 years, with a mean of 14.0.

Procedures

The instruments were administered at two sessions. At the first, a group administration, the following instruments were given in the order shown:
1. Test of Implied Meanings (Sundberg, 1966), an auditory measure of accuracy in social perception via spoken language.
2. Group Embedded Figures Test (Witkin, Oltman, Raskin, & Karp, 1971), a paper-and-pencil test of field independence-dependence.
3. Extended Range Vocabulary Test (Ekstrom, French, & Harman, 1976), a test of verbal ability.
4. Mental Rotations Test (Vandenberg & Kuse, 1978), a paper-and-pencil test of spatial visualization.

---

[1]A transcript of the ICI instructions and exercises is available from the author.

Table 1
Characteristics of ICI Exercises

	Characteristic			
	Subordinate		Task	
Exercise	Sex	Job	Type	Content
Replies Section				
Practice	Female	Assistant Manager, Filing	Indefinite	Deciding about promotion
1	Male	Technical Staff	Positive	Commending good work
2	Female	Technical Staff	Negative	Correcting poor work
3	Male	Manager, Inventory	Negative	Dealing with office feud
4	Female	Indefinite	Positive	Commending good work
Judgments Section				
Practice	Female	Assistant Manager, Filing	Indefinite	Deciding about promotion
5	Male	Creative Director	Positive	Commending good work
6	Female	Manager, Credit	Negative	Handling complaint
7	Male	Manager, Sales Office	Negative	Motivating poor performance
8	Female	Staff Assistant	Positive	Commending good work

5. Affect Judgment Test (Feldstein, 1964) —Items 1 to 20 of Tape 2, an auditory measure of accuracy in social perception via spoken language.
6. Defense Mechanism Inventory (Gleser & Ihilevich, 1969), a self-report questionnaire for defense mechanisms.

At the second session, 2 to 34 days after the first, the following instruments were administered individually:
1. Biographical Information Form, a specially devised questionnaire covering age, handedness, native language, education, occupation, and if attended college, major, minor,

Figure 1
Transcript of Practice Exercise

Narrator:
"First, a practice exercise. This is the background for the practice exercise: Martha Robbins is the assistant manager of the filing department, which is headed by John Freeman. Martha has just walked into your office."

Actor:
"I'm glad you could squeeze me in. I just talked to John, and he told me he's leaving. He knows I'm interested in moving ahead. And we both thought that it would be a good idea if I told you myself. John thinks I'm the most natural person to take his place and, frankly, so do I. I also think there ought to be some changes made in the department, but maybe we should talk about that later."

psychology courses completed, Scholastic Aptitude Test (Donlon & Angoff, 1971) scores, and grade-point average.

2. Rod and Frame Test (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962), a laboratory measure of field independence-dependence.
3. Perceiving Affect Test (Rapaczynski, 1979), an auditory measure of accuracy in social perception via spoken language.
4. ICI. Each subject was given special instructions on using the videotape player so that the ICI was completely self-administered.[2] The subject was seated at a table. On it was (1) a Craig audio cassette recorder (Model 2603), equipped with an external microphone and stand, and a TDK AD cassette; and (2) a questionnaire for the ICI Judgments section, along with a ballpoint pen. To her right, on another table, was a Sony ¾" video cassette player (Model VP 2000), the location of its pause button marked with a red arrow. Directly in front of her, on a third table, was a Sony 17" color video receiver/monitor (Model CVM 1750). Its picture tube was approximately at the subject's eye level, 56 inches away.

The Biographical Information Form was administered first, followed by the other devices in counterbalanced order. The Mental Rotations Test, Defense Mechanism Inventory, and Biographical Information Form questions on handedness and language were not relevant to the purpose of this pilot study, and usable data on occupation, college minor, Scholastic Aptitude Test scores, and grade-point average were only available for a small number of subjects. In addition, scoring of the Perceiving Affect Test has not been completed. Consequently, the data for these variables will not be reported.

Rating ICI Responses

All of the responses to the actual exercises of the ICI were independently rated by two psychologists with training and experience in interpersonal processes who teach courses on this topic in graduate business schools. In rating the Replies section, the four exercises of each subject were rated in order, one after the other (e.g., Subject A's Exercise 1 was rated for effectiveness and originality, then her Exercise 2 was rated for these variables, and so on). Before the replies of the subjects in the study were rated, those for three other subjects (office workers) were rated as practice. The audio cassettes containing the replies for the subjects in the study, one individual recorded per side, were rated in random order.

In rating the Judgments section, each exercise was rated for all of the subjects before doing the exercise that followed it (i.e., Exercise 5 was rated for the sample, then Exercise 6, and so forth). As in the case of the replies, the ratings of the judgments for the subjects in the study were made after practice ratings of three other subjects' responses. The questionnaires containing the judgments for the subjects in the study were rated in random order.[3]

Results and Discussion

Interrater Reliability of ICI Exercises and Total Scores

The interrater reliabilities of the ICI exercises and total scores, estimated from the product-moment correlations based on ratings for each rater, are reported in Table 2. Reliabilities were substantial for all of the exercises and total scores, the reliabilities for Replies—Effectiveness being highest. The median correlations for the exercises were .84 for Replies—Effectiveness, .61 for Replies—Originality, and .68 for Judgments—Accuracy; the corresponding correlations for the total scores were .90, .80, and .80. All of the correlations were significant ($p < .01$, two tailed). These results indicate that

---

[2] A transcript of the special instructions used for self-administering the ICI is available from the author.

[3] A copy of the rating instructions together with examples of responses, for each exercise, with the highest and lowest ratings are available from the author.

Table 2
Interrater Reliability
of ICI Exercises

Exercise	Interrater Reliability
Replies--Effectiveness	
1	.90**
2	.80**
3	.75**
4	.87**
Replies--Originality	
1	.53**
2	.66**
3	.72**
4	.56**
Judgments--Accuracy	
5	.64**
6	.67**
7	.83**
8	.70**

Note. Interrater reliability was estimated by the product-moment correlation. Ns for the correlations vary from 47 to 50 for Replies--Effectiveness as well as Replies--Originality and from 55 to 57 for Judgments--Accuracy because of missing data.
**$p < .01$, two tailed.

the raters agreed closely in their ratings, a notable outcome, considering the broad definitions of the variables.

Halo Effect in ICI Exercises

The mean product-moment intercorrelations of the exercises for Replies—Effectiveness and Replies—Originality are shown in Table 3 for the same raters as well as for different raters. The first set of intercorrelations was based on the mean correlation (using the z transformation), for each pair of exercises, between the ratings by the same raters (e.g., Rater A's rating of Replies—Effectiveness for Exercise 1 versus his or her rating of this variable for Exercise 2, and

Rater B's rating of this variable for Exercise 1 versus his or her rating of it for Exercise 2). The second set was based on the mean correlation between the ratings by different raters (e.g., Rater A's ratings of Replies—Effectiveness for Exercise 1 versus Rater B's rating of this variable for Exercise 2, and vice versa). Both kinds of intercorrelations were roughly similar. The median correlations for the same raters were .50 for Replies—Effectiveness, .48 for Replies—Originality, and .42 for Replies—Effectiveness versus Replies—Originality; the corresponding correlations for different raters were .48, .44, and .42. Because the intercorrelations for different raters are free of halo effect, the similarity between these intercorrelations and those for the same raters, which are open to such an effect as a consequence of each subject's replies being appraised in sequence by the raters, indicates that this influence was negligible. The presence of halo effect would distort the interrelations of the exercises and, in turn, the intercorrelations as well as the internal-consistency reliability of the total scores derived from the exercises.

Internal Consistency Reliability of ICI Total Scores

The internal consistency reliability of the total scores (summing, for each exercise, the ratings for the two raters), estimated by coefficient alpha, was substantial: .81 for Replies—Effectiveness, .82 for Replies—Originality, and .74 for Judgments—Accuracy. This high level of reliability is striking in view of the small number of components on which the scores are based.

Intercorrelations of ICI Exercises and Total Scores

Exercises. The product-moment intercorrelations of the exercises (summing the ratings for the two raters) appear in Table 4. The intercorrelations of the exercises for Replies—Effectiveness, Replies—Originality, and Judgments

Table 3

Mean Intercorrelations of ICI Exercises

for Replies--Effectiveness and Replies--

Originality, Same and Different Raters

	Exercise							
	Replies--Effectiveness				Replies--Originality			
Exercise	1	2	3	4	1	2	3	4
Replies--Effectiveness								
1		.58	.44	.48	.58	.39	.31	.25
2	.56		.51	.56	.56	.51	.42	.37
3	.41	.48		.48	.40	.56	.72	.45
4	.44	.54	.47		.41	.43	.36	.67
Replies--Originality								
1	.48	.47	.42	.41		.45	.36	.49
2	.34	.40	.49	.43	.28		.56	.49
3	.29	.45	.63	.40	.36	.58		.46
4	.20	.42	.48	.61	.41	.52	.47	

Note. Mean correlations for the same raters are above the diagonal and those for different raters are below it. Means were computed by using the z transformation. Ns for the correlations range from 47 to 50 because of missing data.

—Accuracy were usually appreciable, and all but one was significant ($p < .05$, two tailed). The three sets of intercorrelations were approximately at the same level; the median correlations were .53, .58, and .45, respectively. The exercises in each set generally had similar correlations with each other, a major exception being that Exercise 8 for Judgments—Accuracy correlated minimally with the others in that section. No reason is apparent for this exercise's differential performance. All in all, these findings suggest that apart from Exercise 8 the exercises in each set operated in the same way, not varying with the characteristics of the subordinate and the task involved. The results also imply that the variables being measured had some degree of generality.

Total scores. The product-moment correlations between the total scores (summing the rating for the two raters) were appreciable: .76 between Replies—Effectiveness and Re-plies—Originality, .45 between Replies—Effectiveness and Judgments—Accuracy, and .45 between Replies—Originality and Judgments—Accuracy. All were significant ($p < .01$, two tailed). The substantial correlations of the two Replies scores with the Judgments score are noteworthy. These associations support the validity of both kinds of scores, for some relationship should exist between the scores if they are measuring separate aspects of interpersonal competence. The extremely high correlations between the two Replies scores, however, raises real doubts about whether they are assessing different characteristics.

Means of ICI Exercises and Total Scores

Exercises. The means and standard deviations of the exercises (summing the ratings for the two raters) are reported in Table 4. The means were uniformly low (the highest was 4.45

Table 4

Means, Standard Deviations, and Intercorrelations of ICI Exercises

Exercise	Mean	S.D.	Replies-- Effectiveness				Replies-- Originality				Judgments-- Accuracy			
			1	2	3	4	1	2	3	4	5	6	7	8
Replies-- Effectiveness														
1	3.66	2.11		.62**	.46**	.48**	.61**	.39**	.31*	.25*	.12	.43**	.33*	.03
2	3.04	1.43			.54**	.60**	.61**	.52**	.48**	.47**	.16	.38*	.33*	.00
3	3.88	1.76				.52**	.49**	.61**	.78**	.56**	.41**	.58**	.50**	.12
4	4.06	2.32					.48**	.49**	.42**	.74**	.23	.48**	.29*	.19
Replies-- Originality														
1	2.94	1.17						.45**	.43**	.58**	.13	.45**	.28*	-.14
2	3.52	1.52							.66**	.64**	.26*	.41**	.46**	.01
3	4.31	1.82								.57**	.29*	.50**	.57**	.12
4	3.00	1.40									.27*	.42**	.39**	.01
Judgments-- Accuracy														
5	4.45	1.84										.55**	.53**	.08
6	4.11	1.78											.69**	.37*
7	3.66	1.82												.23*
8	3.64	1.74												

Note. Ns for the correlations vary from 46 to 56 because of missing data.

*$p < .05$, two tailed.

**$p < .01$, two tailed.

Table 5
Correlations of ICI Scores with Background Variables and Tests

Variable	Reliability	Score		
		Replies-- Effectiveness	Replies-- Originality	Judgments-- Accuracy
Age	--	.30*	.36**	.25*
Education	--	.35**	.41**	.31*
Psychology Major	--	.54**	.57**	.08
Completed Psychology Course	--	.26*	.21	.43**
Extended Range Vocabulary Test	.89	.41**	.45**	.50**
Group Embedded Figures Test	.88	.27*	.20	.25*
Rod and Frame Test	.90	.00	.04	-.06
Test of Implied Meanings	.76	.31*	.22	.34**
Affect Judgment Test	.23	.00	.07	-.19

Note. Reliability was estimated by the split-half method. Ns for reliability coefficients vary from 57 to 58, and Ns for correlations vary from 46 to 47 for Replies--Effectiveness as well as Replies--Originality and from 54 to 55 for Judgments--Accuracy because of missing data.
*p < .05, two tailed.
**p < .01, two tailed.

and the lowest 2.94), considering that the theoretical range of the scores is 2 to 10. The means correspond to ratings of "Slightly Effective" for all of the exercises for Replies—Effectiveness; "Slightly Original" for all of the exercises for Replies—Originality, except Exercise 1, whose corresponding rating is "Not at all Original"; and "Slightly Accurate" for all of the exercises for Judgments—Accuracy. The means of the exercises for Replies—Effectiveness, like those for Judgments—Accuracy, were generally similar to each other, but the means of the exercises for Replies—Originality were noticeably different: those for the negative exercises (2 and 3) were considerably higher than the others.

These findings indicate that the exercises were of appropriate difficulty for this sample and imply that the exercises have sufficient "ceiling" and "floor" to be suitable for use with other groups of subjects that are either more or less able than the present one. The results also suggest that the level of performance only varied markedly with the exercises' characteristics in the case of Replies—Originality, responses of greater originality being made to exercises that

were negative in nature. It may well be that the kinds of scenes portrayed in the positive exercises, unlike those in the negative exercises, inevitably limit appropriate responses to those that are relatively stereotyped and conventional.

Total scores. The means for the total scores were 14.85 for Replies—Effectiveness, 13.96 for Replies—Originality, and 15.93 for Judgments—Accuracy; the corresponding standard deviations were 6.16, 4.83, and 5.37. These means, like those for the exercises, were moderate, given a theoretical range of scores from 8 to 40. The means correspond to ratings of "Slightly Effective" for Replies—Effectiveness, "Slightly Original" for Replies—Originality, and "Slightly Accurate" for Judgments—Accuracy. In line with the findings for the exercises, these results imply that the ICI's overall level of difficulty was satisfactory for this sample and is apt to be adequate for groups with different levels of ability.

Correlations of ICI with Other Measures

The product-moment correlations of the ICI total scores (summing the ratings for the two

raters) with the other variables in the study are given in Table 5 together with the internal consistency reliability coefficients for the latter, where computable.

The ICI scores generally correlated both significantly ($p < .05$, two tailed) and appreciably with age (.25 to .36), education (.31 to .41), and training in psychology (Psychology Major, .54 to .57; Completed Psychology Course, .26 to .43). Each of the ICI scores had approximately similar correlations with these variables, with two major exceptions, both involving indices of training in psychology: Replies—Originality correlated substantially with Psychology Major (.57), but not with Completed Psychology Course (.21); and Judgments—Accuracy did not correlate with the former (.08) but correlated substantially with the latter (.43). The reason for these divergences is unclear. The ICI scores, in line with their correlations with education, consistently correlated appreciably with verbal ability (Extended Range Vocabulary Test, .41 to .50, $p < .01$).

The ICI scores' significant ($p < .05$) correlations with the other variables were scattered. Two of the ICI scores (Replies—Effectiveness and Judgments—Accuracy) correlated moderately with one index of field independence-dependence (Group Embedded Figures Test, .25, .27), but none of the scores correlated with another measure of this cognitive style (Rod and Frame Test, .04 to −.06). Similarly, these two ICI scores correlated modestly with one measure of accuracy in social perception (Test of Implied Meanings, .31, .34), but none correlated with a second (Affect Judgment Test, .07 to −.19). However, the lack of association with the latter may be due to that test's minimal reliability (.23).

Overall, these results support the validity of the ICI as a measure of interpersonal competence, for this variable should be associated with age, education, training in psychology, and accuracy in social perception. The close relationships with verbal ability, however, are more problematical. Verbal ability, like education, ought to be implicated in interpersonal competence, but the magnitude of the obtained correlations raises an obvious question about the extent to which the ICI taps something other than intelligence. As noted earlier, the pervasive influence of general ability is a common problem with tests in this area.

The findings with the field independence-dependence measures are puzzling, because the correlations are inconsistent with each other, and the modest but positive correlations with the Group Embedded Figures Test are contrary to the negative associations expected by theory and research in this area, reflecting superiority by field dependent subjects (Witkin & Goodenough, 1977).

Conclusions

The ICI appears promising, judging from the preliminary findings. The level of the reported correlations must be interpreted with caution because of the relatively small sample involved. Nonetheless, it appears that (1) the total scores are reliable, in terms of both interrater agreement and internal consistency; (2) the exercises generally function similarly, with the notable exception of one exercise in the Judgments section; (3) the two total scores for Replies are only moderately correlated with the total score for Judgments, but the Replies scores may be too highly intercorrelated to justify using both; and (4) the correlations of the total scores with background variables and other tests generally support the ICI's construct validity but raise the possibility of confounding with general ability.

These tentative results need to be verified and extended in further research (1) using large samples of different kinds of subjects, especially business executives; and (2) employing other indices of interpersonal skills, such as peer ratings, instruments that tap potentially confounding variables, particularly general ability, and devices that assess the processes underlying this performance, notably verbal fluency. In the event that these results are also favorable, atten-

tion can then turn to adapting the ICI for group administration and simplified scoring.

References

Adler, A. *Understanding human nature.* Cleveland OH: World, 1941.

Argyle, M. *Social interaction.* London: Methuen, 1969.

Argyle, M. *The psychology of interpersonal behaviour* (2nd ed.). Harmondsworth, England: Penguin, 1972.

Argyris, C. *Interpersonal competence and organizational effectiveness.* Homewood IL: Dorsey, 1962.

Argyris, C. *Integrating the individual and the organization.* New York: Wiley, 1964.

Bales, R. F. *Interaction process analysis.* Chicago: University of Chicago Press, 1950.

Bales, R. F. *Personality and interpersonal behavior.* New York: Holt, Rinehart, & Winston, 1970.

Breen, P., Donlon, T. F., & Whitaker, U. *Teaching and assessing interpersonal competence—A CAEL handbook.* Princeton NJ: Cooperative Assessment of Experiential Learning, 1977.

Cline, V. B. Interpersonal behavior. In B. A. Maher (Ed.), *Progress in experimental personality research* (Vol. 1). New York: Academic Press, 1964.

Cronbach, L. J. *Essentials of psychological testing* (3rd ed.). New York: Harper & Row, 1970.

Crooks, L. A. The selection and development of assessment center techniques. In J. L. Moses & W. C. Byham (Eds.), *Applying the assessment center method.* New York: Pergamon, 1976.

Donlon, T. F., & Angoff, W. H. The Scholastic Aptitude Test. In W. H. Angoff (Ed.), *The College Board Admissions Testing Program: A technical report on research and development activities relating to the Scholastic Aptitude Test and Achievement Tests.* New York: College Entrance Examination Board, 1971.

Eisler, R. M. Behavioral assessment of social skills. In M. Hersen & A. S. Bellack (Eds.), *Behavioral assessment.* New York: Pergamon, 1976.

Ekstrom, R. B., French, J. W., & Harman, H. H. *Kit of factor-referenced cognitive tests.* Princeton NJ: Educational Testing Service, 1976.

Feldstein, S. *Identification of simulated affective vocal expression.* Unpublished manuscript, 1964.

Flavell, J. H., Botkin, P. J., Fry, C. L., Wright, J. W., & Jarvis, P. E. *The development of role-taking and communication skills in children.* New York: Wiley, 1968.

Foote, N. N., & Cottrell, L. S., Jr. *Identity and interpersonal competence.* Chicago: University of Chicago Press, 1955.

Frederiksen, N., Saunders, D. R., & Wand, B. The in-basket test. *Psychological Monographs,* 1957, *71* (9, Whole No. 438).

Frederiksen, N., & Ward, W. C. Measures for the study of creativity in scientific problem-solving. *Applied Psychological Measurement,* 1978, *2,* 1–24.

Frederiksen, N., Ward, W. C., & Samph, T. *Development of methods for selection and evaluation in medical education.* Princeton NJ: Educational Testing Service, 1979.

Gleser, G. C., & Ihilevich, D. An objective instrument for measuring defense mechanisms. *Journal of Consulting and Clinical Psychology,* 1969, *33,* 51–60.

Guilford, J. P. *The nature of human intelligence.* New York: McGraw-Hill, 1967.

Guilford, J. P., & Hoepfner, R. *The analysis of intelligence.* New York: McGraw-Hill, 1971.

Hendricks, M., Guilford, J. P., & Hoepfner, R. *Measuring creative social intelligence* (Psychological Laboratory Report No. 42). Los Angeles: University of Southern California, 1969.

Kagan, N. Can technology help us toward reliability in influencing human interaction? *Educational Technology,* 1973, *13,* 44–51.

O'Sullivan, M., & Guilford, J. P. Six factors of behavioral cognition: Understanding other people. *Journal of Educational Measurement,* 1975, *12,* 255–271.

Rapaczynski, W. Personal communication, May 1979.

Sullivan, H. S. *An interpersonal theory of psychiatry.* New York: Norton, 1953.

Sundberg, N. D. A method for studying sensitivity to implied meanings. *Gawein (Tijdschrift voor Psychologie),* 1966, *15,* 1–8.

Thorndike, E. L. Intelligence and its uses. *Harper's Monthly Magazine,* 1920, *140,* 227–235.

Vandenberg, S. G., & Kuse, A. R. Mental Rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills,* 1978, *47,* 599–604.

Walker, R. E., & Foley, J. M. Social intelligence: Its history and measurement. *Psychological Reports,* 1973, *33,* 839–864. (Monograph)

Weinstein, E. A. The development of interpersonal competence. In D. A. Goslin (Ed.), *Handbook of socialization theory and research.* Chicago: Rand McNally, 1969.

White, R. W. Motivation reconsidered: The concept of competence. *Psychological Review,* 1959, *66,* 297–333.

Witkin, H. A., Dyk, R. B., Faterson, H. F., Goodenough, D. R., & Karp, S. A. *Psychological differentiation.* New York: Wiley, 1962.

Witkin, H. A., & Goodenough, D. R. Field dependence and interpersonal behavior. *Psychological Bulletin*, 1977, *84*, 661–689.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. *A manual for the Embedded Figures Tests.* Palo Alto CA: Consulting Psychologists Press, 1971.

Author's Address

Send requests for reprints or further information to Lawrence J. Stricker, Educational Testing Service, Princeton NJ 08541.