CRITERION-RELATED VALIDITY OF 
THE ACTIVE LISTENING COMPONENT 
OF THE BEHAVIOURAL TEST OF 
INTERPERSONAL SKILLS

Joanne K. Olson, Carroll L. Iwasiw and Brian A. Gerrard

Active listening is a major component of therapeutic communication with clients. Although active listening has received much attention in many disciplines, valid and reliable tools to measure a health professional’s active listening skills are lacking. The absence of behavioural measures of empathy limits research in this important dimension of clinical practice.

The purpose of this study was to extend the validity testing of the active listening component of the Behavioral Test of Interpersonal Skills (BTIS), a test of communication skills (Gerrard & Buzzell, 1980). Specifically, it was criterion-related validity that was assessed. Validation of this aspect of the BTIS was critical because active listening has been identified as one of the most essential aspects of therapeutic communication.

Review of the Literature

Active listening and empathy

Active listening is defined as "the skill of understanding what your patient is saying and feeling and communicating to your patient in your own words what you think he is saying and feeling" (Gerrard, Boniface & Love, 1980, p. 133). The literature frequently refers to the same type of behaviour as empathy. Active listening can be considered to be the measurable dimension of empathy. Through analysis of a helper's verbal response to a patient it is possible to determine the accuracy of the helper's understanding of the thoughts and feelings expressed by the patient.

Joanne K. Olson, R.N., M.S. is Associate Professor in the Faculty of Nursing, at the University of Alberta, in Edmonton. Carroll L. Iwasiw, R.N., M.Sc.N. is Associate Professor, in the Faculty of Nursing at the University of Western Ontario, in London. Brian A. Gerrard, B.A., M.A., Ph.D. is Associate Professor in the School of Education at the University of San Francisco, in California.
Rogers (1958; 1961) has identified empathy as one of the major elements in establishing a helping relationship. Of all aspects of the therapeutic communication process in nursing, empathy has been cited as one of the most essential and complex (Forsyth, 1980; Gagan, 1983; Kalisch, 1973; La Monica, 1981). In attempting to be supportive, nurses sometimes make statements such as "don't worry" or "don't be upset". These "Don't Feel" statements are antithetical to empathy (Gerrard & Buzzell, 1980, p. 9) because they negate the patient's feelings and deny the right to experience those feelings. Although empathy is essential in a helping relationship, nurses have not always demonstrated high levels of empathy (Hughes & Carver, 1990; Iwasiw & Olson, 1987; Olson & Iwasiw, 1987).

**Relationship of health professionals' interpersonal skills to patient outcomes**

Health professionals require good interpersonal skills to establish, maintain and terminate effective helping relationships with patients and to establish and maintain collaborative relationships with each other. Poor communication skills on the other hand could anger and threaten patients so that their purpose in seeking health care is defeated (Gerrard et al., 1980).

Various studies, reported mainly in the medical and counselling literature, have examined the behavioural relationship between health professionals’ interpersonal skills and patient outcomes. Nursing studies of this nature have been limited in number; however, several have specifically provided evidence that the nurses’ use of empathy positively influences client outcomes. Williams (1979) found that the level of nurse empathy was related to changes in self-concept in elderly patients. In a study by La Monica, Wolf, Madea and Oberst (1987), the immediate effects of nurses’ empathy, before and after empathy training, were measured on the variables anxiety, depression, hostility and satisfaction with care. There was less anxiety and hostility reported in patients cared for by nurses who had received empathy training. In a 1990 study, Vanderlee reported that nurses' levels of empathy were positively correlated with clients' feelings of being understood, as measured by the empathy subscale of the Barrett-Lennard Relationship Inventory.

**Purpose and Hypotheses**

The purpose of this study was to test the criterion-related validity of the active listening component of the Behavioural Test of Interpersonal Skills. Actual clinical interactions were the standard against which the BTIS was tested. Previous criterion-related validity testing has been based on supervisor and peer ratings of subjects' communication skills - and not on assessment of actual interactions between health professionals and patients. The interpersonal categories of the Behavioural Test of Interpersonal Skills which this study tested were:
1. Feeling: Any relevant general (e.g. upset) or specific (e.g. angry) reference to feeling.

2. Content: The reason for the speaker's feeling.

3. Don't Feel: Any attempts to suppress or discourage expression of speaker's feelings (Gerrard & Buzzell, 1980, p. 43).

Based on the assumption that nursing personnel would structure interactions such that patients would feel free to disclose Feeling and Content, we hypothesized that there would be positive correlations of at least .50 between:

1. subjects’ BTIS percentage scores on the behavioural category Feeling and the percentage of opportunities to which they correctly identify patients’ Feelings during clinical interactions.

2. subjects’ BTIS percentage scores on the behavioural category Content and the percentage of opportunities to which they correctly identify the Content of patients’ statements during clinical interactions.

3. subjects’ BTIS percentage scores on the behavioural category Don’t Feel and the percentage of opportunities to which they make Don’t Feel statements to patients during clinical interactions.

Methods

Design

This study was a descriptive correlational investigation. Nursing personnel were audiotaped while responding to the simulated patients on the BTIS and while interacting with real patients. Subjects’ statements of Feeling, Content and Don’t Feel, in response to the BTIS patient situations and during patient interactions, were scored. Correlations were computed for each category between BTIS percentage scores and percentage scores in actual subject-patient interactions.

Instrument

The Behavioural Test of Interpersonal Skills (BTIS) is a test that can be used to assess the interpersonal or interviewing skills of any health professional student or practitioner. The tool is useful for comparing the interpersonal skills of different groups, for giving feedback on the effectiveness of interpersonal skills training and for research.

The test consists of two main parts: a videotape that is used to elicit verbal responses (see Figure 1 for an example of one BTIS situation and its content analysis); and, a content-analysis scoring sheet that is used to rate the verbal responses.
Female Patient:

"It's a dull nagging pain. I don't know what else I can tell you. It just goes on and on night and day. I don't think it's ever going to go away."

<table>
<thead>
<tr>
<th>Content Analysis Category</th>
<th>Scoring Guidelines</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feeling</td>
<td>Underlying: hopeless, helpless, scared.</td>
</tr>
<tr>
<td></td>
<td>Surface: in pain, worried, upset.</td>
</tr>
<tr>
<td>Content</td>
<td>&quot;Because ... the pain never stops, you think you won’t get better.&quot;</td>
</tr>
</tbody>
</table>

**Figure 1**
Sample BTIS patient situation and content analysis.

There are 26 problem situations on the BTIS videotape to which a subject can respond with active listening. Thirteen are patient situations and 13 are health professional situations. Only the patient situations were used. There are three categories of patient situations: Aggression (verbal attacks and unreasonable requests); Distress (depression/sadness, pain, anxiety and anger); and Positive Emotion (happiness and affection).

*Reported use of the BTIS.* The BTIS has been used to study the interpersonal skills of practicing professionals (Hills & Knowles, 1983; Iwasiw & Olson, 1985, 1987). Student communication skills have also been assessed with the BTIS (Anderson & Gerrard, 1984; Gerrard, 1982; Olson & Iwasiw, 1987).

*Reliability of the BTIS.* The non-reactivity of the Active Listening categories (Feeling, Content and Don’t Feel) has been demonstrated. There were no significant differences in subjects’ initial scores on the BTIS and their scores at 6 and 16 week intervals (Gerrard & Buzzell, 1980).

*Validity of the BTIS.* Content validity of the BTIS was established through an extensive literature review and input of health professionals in the development of the initial problem situations (Gerrard & Buzzell, 1980). There is moderate support for the construct validity of the active listening component of the BTIS (Anderson & Gerrard, 1984; Gerrard & Buzzell, 1980; Olson & Iwasiw, 1987).

Criterion-related validity of the BTIS categories of Feeling, Content and Don’t Feel was demonstrated in a study where peers and supervisors used
semantic differential items to rate a sample of 26 nurses. These ratings were based on supervisors’ familiarity with the nurses and not on direct observation of their communication skills. Significant positive correlations (p<.05) were found between the semantic differential item "Has Empathy for Others", and the categories of both Feeling and Content. A significant negative correlation (p<.05) was found between the category Don’t Feel and the item "Has Empathy for Others" (Gerrard & Buzzell, 1980, pp. 22-26).

**Study sample and sampling procedures**

The population was composed of registered nurses, registered nursing assistants, nursing aides and nursing orderlies (all employed on selected units in two acute care and two chronic care institutions) and registered nurses in two community health agencies. Nursing personnel whose positions required regular interaction with patients were invited to participate. A variety of nursing personnel was included in the expectation of yielding a wide distribution of scores on the study variables.

Eligible staff were identified through the agency personnel offices and a letter was sent inviting them to attend an informational session about the project. Signed consent was obtained from volunteer subjects and appointments were made for data collection. The sample consisted of 41 volunteers. Twenty-five subjects were registered nurses and 16 were non-professional nursing personnel.

**Procedures to obtain patient participation**

On the day of patient audiotaping the head nurse identified the alert and oriented patients who were assigned to the subject. A research assistant explained the study to these patients, requested their participation and obtained their written consent to be tape-recorded while receiving nursing care. Written information was also provided to these patients. Two or three patients for each subject agreed to participate and none withdrew from the study.

**Data collection procedures**

There were two parts to the data collection. One part consisted of nurse subjects being audiotaped for 30-60 minutes during their usual interactions with patients. Subjects were also audiotaped while responding to the BTIS. The order of the patient and BTIS taping sequence was varied in an attempt to eliminate the possible effects of one taping situation on the other taping situation.

The research assistant instructed each subject in the use of the tape recorder and informed him or her about the patients who had agreed to participate.
The research assistant then left the unit and returned later to retrieve the recording. Each subject was in control of the taping session and decided precisely when, for how long and with which consenting patients taping would occur. Voice-activated, pocket-sized recorders were used for convenience and to reduce the artificial nature of the taping session. Sixty-minute tapes eliminated the need to change or turn over the tape.

A room on each unit was used for the BTIS taping sessions. Subjects were alone and were audiotaped while responding aloud to the BTIS. This tape was immediately given to the research assistant.

**Audiotape analysis**

**BTIS.** Responses to the BTIS were scored by two of the investigators. Active Listening was scored in accordance with the guidelines for the categories of Feeling, Content and Don’t Feel, as described in the BTIS User’s Manual (Gerrard & Buzzell, 1980). See Figure 1. Intra-rater and inter-rater scoring reliability had previously been established for each behavioural category (Cohen’s Kappa Statistic 0.85-1.00). For each behaviour (Feeling, Content, Don’t Feel) a score of "one" was obtained when the behaviour was present in the subject’s response. The absence of the behaviour resulted in a score of "zero".

**Subject-patient.** The subject-patient audiotapes were analyzed according to a procedure developed by the investigators. The investigators randomly selected three, five-minute segments of each subject-patient audiotape for analysis. The interactions were first analyzed for the feelings and content expressed and these constituted the number of opportunities to respond with active listening. The interactions were then analyzed to determine the subject’s Feeling, Content and Don’t Feel responses to the opportunities. Each subject received a final percentage score for these three communication behaviours. This score represented the percentage of opportunities to which the subject gave Feeling, Content and Don’t Feel responses. The possible range of scores for each category was 0-100% (See Figure 2 for an example of a subject-patient interaction analysis). Intra- and inter-rater scoring reliability for subject-patient interactions had been established for feelings and content expressed by the patient and therefore, the opportunities for active listening responses and for subjects’ responses of Feeling, Content and Don’t Feel (Kappa statistics 0.85-0.90).

Subject tapes had been coded by the research assistant in a manner to ensure that the investigators could not link any subject’s BTIS and patient interview tapes. This procedure was employed to reduce bias in scoring.
Nurse-Patient Interaction
Nurse: I understand that your arthritis has been really bothering you.
Patient: I’ve tried to do everything I’m supposed to. I take my pain pills and I rest and I never overdo it, but still... the pain is bad.

Nurse: You feel discouraged that the pain is bad, even though you follow the doctor’s orders.

Patient: I try not to let it get me down, even when I have these flare-ups. Sometimes though, it’s hard to keep smiling. I do try to be pleasant.

Nurse: Don’t be discouraged. You’ve made a lot of progress since you came into hospital.

Patient: Yeah, I guess so.

Total 3 3 1 1 1
\[ \frac{1}{3} = \frac{1}{3} = \frac{1}{3} \]

Subject’s Score 33\(\frac{1}{3}\)% 33\(\frac{1}{3}\)% 33\(\frac{1}{3}\)%

Figure 2.
Sample nurse-patient interaction and content analysis.

Results and Discussion

Hypothesis testing

BTIS percentage scores and interview percentage scores were analyzed using Pearson product-moment correlations for each of the three categories (Feeling, Content, Don’t Feel).

Hypothesis 1, that there would be a positive correlation between Feeling responses to the BTIS and to real patients, was not supported \( (r = .12; K = .07; p > .05) \). There was no significant correlation between BTIS percentage
scores on the behavioural category Feeling and the percentage of Feeling responses in clinical interactions.

Hypothesis 2 was not supported \((r = -.10; K = -.7; p > .05)\). There was no significant correlation between BTIS percentage scores on the behavioural category Content and the percentage of Content responses in clinical interactions.

Hypothesis 3 was not supported \((r = .24, p > .05)\). There was no significant correlation between BTIS percentage scores on the behavioural category Don’t Feel and the percentage of Don’t Feel responses in clinical interactions.

Post-hoc content validation of the BTIS. To ensure that similar situations were being presented to the subjects on the BTIS and during actual clinical interactions, each patient statement was labelled according to the BTIS categories of patient situations. The clinical interviews gave the subjects opportunities to respond to situations of Distress (pain, depression/sadness, anxiety and anger) and Positive Emotions. There were no situations of Aggression (verbal attacks and unreasonable requests) in the interviews. See Table 1 for a comparison of the percentages of the types of patient situations presented to subjects on the BTIS and during actual client interviews. The BTIS presented an approximately equal percentage of all types of patient situations, whereas situations of pain predominated with real patients.

Table 1

<table>
<thead>
<tr>
<th>Percentage of Types of Patient Situations Presented to Subjects</th>
<th>BTIS</th>
<th>Interview</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distress</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain</td>
<td>15.38</td>
<td>38.57</td>
</tr>
<tr>
<td>Depression</td>
<td>15.38</td>
<td>12.69</td>
</tr>
<tr>
<td>Anxiety</td>
<td>15.38</td>
<td>21.31</td>
</tr>
<tr>
<td>Anger</td>
<td>15.38</td>
<td>13.19</td>
</tr>
<tr>
<td>Positive Emotion</td>
<td>7.69</td>
<td>14.21</td>
</tr>
<tr>
<td>Aggression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal attack</td>
<td>15.38</td>
<td>0</td>
</tr>
<tr>
<td>Unreasonable request</td>
<td>15.38</td>
<td>0</td>
</tr>
</tbody>
</table>
Additional discussion

Study results may be interpreted in two ways. The active listening component of the BTIS may indeed lack concurrent criterion-related validity. It is also possible that the BTIS does have criterion-related validity but that method problems interfered with the demonstration of the validity.

One such problem was in the scoring procedures. Four subjects received a score of 0% when there were no opportunities to respond with Feeling. For 22 subjects, 0% meant a failure to identify the Feeling stated by the patient, no matter how many opportunities were present. A score of 100% meant, in some instances, simply that the subject had responded accurately only once to one opportunity. Therefore, a final score of 0% or 100% may not have represented a large quantitative difference in Feeling responses.

Clinical interaction is the ultimate external criterion for assessing tests of health professional communication. Nonetheless, the Hawthorne effect may have operated so that subjects may not have responded in their usual manner during data collection. It is also possible that only subjects who felt confident about their communication skills and comfortable with audiovisual equipment volunteered. Another issue is whether interaction with only one or two patients truly represented the subject’s usual patterns of interaction with many patients.

This study was premised on the assumption that nursing personnel would structure interactions so that patients would feel free to disclose Feeling and Content. This assumption was inaccurate. In the 15-minute analyzed segments, there were 20 of the 41 subjects who were given only two or fewer opportunities to respond with Feeling. The patient’s lack of disclosure seemed to result from the nurses’ monopolization of the interaction (often with a focus on self), or directing of the interaction to superficial or social topics (e.g., the weather). Although social communication is appropriate at times, the predominance of this type of interaction fails patients who need encouragement to describe their situation and perspective.

Conclusions and Implications

Concurrent criterion-related validity of the active listening component of the Behavioral Test of Interpersonal Skills was not demonstrated in this study. Clinical active listening skills cannot be predicted from individual scores on the BTIS. Therefore, critical decisions about students’ or practitioners’ communication skills should not be based on BTIS scores. However, because the BTIS situations have content validity, they continue to be useful for teaching purposes.
This study has added to the literature in three ways. A method was developed to identify and score the opportunities for active listening given by the patient to the health care provider. This method of content analysis has potential for use in many types of communication studies. Content analysis of patient statements has provided support for the categories of patient situations on the BTIS and thus, further evidence for the content validity of the instrument. The major strength of this investigation has been the development of a method to validate communication tools during actual clinical practice. Few instruments have been subjected to such rigorous testing.
REFERENCES


This project was funded by Ontario Ministry of Health Research Grant #02538.
RÉSUMÉ

L'écoute attentive Validité du volet écoute active
du test comportemental du sens des relations humaines

L'objet de cette étude était d'appliquer le test de validité du volet écoute active du test comportemental du sens des relations humaines (BTIS). Ce test comportemental est une vidéocassette de scénarios simulés entre un patient et un co-travailleur. Par écoute active, on entend la compréhension de ce qu'une autre personne dit et ressent et le renvoi de cette compréhension à cette personne. Les sujet soumis à cette étude corrélationnelle descriptive étaient 41 infirmiers/infirmières travaillant dans des établissement de soins aigus, de soins prolongés et de santé communautaire. Les sujets ont été filmés tandis qu'ils réagissaient aux patients simulés dans le cadre du test et alors qu'ils interagissaient avec des patients réels. On a analysé les bandes audio au niveau de l'écoute active. Les corrélations établies entre les résultats obtenus au BTIS et les interactions infirmière-patient n'ont pas atteint le seuil d'importance statistique, sans doute en raison de l'échelle restreinte des résultats. On n'a pas non plus réussi à démontrer la validité du volet écoute active, même s'il faut tenir compte dans cette constatation des difficultés inhérentes aux méthodes d'étude. Cette étude est venue s'ajouter à la littérature sur les méthodes qui permettent de mesurer les interactions cliniques et d'évaluer la validité simultanée des instruments.