DIFFERENCES IN AUDIOTAPED VERSUS VIDEOTAPED PHYSICIAN-PATIENT INTERACTIONS

Dawn L. Riddle, Terrance L. Albrecht, Michael D. Coovert, Louis A. Penner, John C. Ruckdeschel, Christina G. Blanchard, Gwendolyn Quinn, and Daniel Urbizu

ABSTRACT: Most medical interaction studies have been conducted on audiotaped recordings of physician-patient encounters. Empirical studies have not previously demonstrated whether coders' scores differ on audio-only versus videotaped data. Data from a convenience sample of forty-seven physician-patient interactions were analyzed using the same coding systems to judge audio-only versus video-based data formats. All coding conditions demonstrated acceptable reliability, using intra-class correlation coefficients. However, MANOVA analyses show that ratings of audiotaped physician-patient interactions are not equivalent to ratings of video-taped encounters. Exploratory factor analyses show differences in the underlying structures of the data derived from the audio-only versus the video information. The differences in the video-based factor solutions account for more total variance and are more consistent with theoretical expectations.

KEY WORDS: analysis; audiotaped; interactions; validity; videotaped.

Dawn L. Riddle, Medical Interaction Research Group, Moffitt Cancer Center. Terrance L. Albrecht, Medical Interaction Research Group, Moffitt Cancer Center and University of South Florida College of Medicine. Michael D. Coovert and Louis A. Penner, Medical Interaction Research Group, Moffitt Cancer Center and University of South Florida. John C. Ruckdeschel, Medical Interaction Research Group, Moffitt Cancer Center and University of South Florida College of Medicine. Christina G. Blanchard, Medical Interaction Research Group, Moffitt Cancer Center. Gwendolyn Quinn, Medical Interaction Research Group, Moffitt Cancer Center and University of South Florida College of Medicine. Daniel Urbizu, Medical Interaction Research Group and Project Link, Moffitt Cancer Center.

Address correspondence to Dr. Dawn L. Riddle, H. Lee Moffitt Cancer Center and Research Institute at the University of South Florida, MRC-CANCONT, MDC-44, 12902 Magnolia Drive, Tampa, FL 33612; e-mail: riddled@moffitt.usf.edu.

A previous version of this paper was presented at the Annual Meeting of the Society for Behavioral Medicine, Washington, D.C., April 3–6, 2002.

This study was supported by grants from the National Cancer Institute (NIH RO1 CA75003-01 and CA 76292S1) and the H. Lee Moffitt Cancer Center and Research Institute Foundation (#G80502). The authors wish to thank Thomas Gordon, Timothy Willis, Thomas King, Kim Hoffman, Jennifer Powell, Todd Manson and Hazel-Anne Johnson for their coding assistance with this project.

Introduction

Real-time recording of physician-patient interactional behavior has been an important methodology for investigating how physicians and patients actually communicate. Investigators have used audiotaping procedures for capturing on-going conversations and also have undertaken more extensive videotaping efforts (that include audiotaping) to record and later analyze live interactions in the medical setting.

Analyses of recorded interactions between health providers and patients have moved beyond the simple categorical tracking of messages (e.g., checklists of message types, Blanchard, Labrecque, Ruckdeschel, & Blanchard, 1988; Roter, 1991) to include observers' judgments of the characteristics and overall nature of the physician-patient relationship as it unfolds (Gillotti, Thompson, & McNeilis, 2002). The quality of the therapeutic relationship, reflected in constructs that include perceived rapport, supportiveness, and positive affect, is a key factor in the patient's trust of the physician, the types of treatment decisions he or she is likely to make, and the overall coping and adjustment of the patient and his/her family to the disease (Albrecht, Blanchard, Ruckdeschel, Coovert, & Strongbow, 1999; Mechanic & Meyer, 2000; Rosser & Kasperski, 2001). The notion of the quality of the interaction is particularly relevant in the setting of a life threatening disease like cancer (Albrecht et al., 1999).

One might expect that audio-visual information (obtained through videotaping) would be necessary for fully and more accurately assessing the content of the patient-physician relational constructs identified above. For instance, previous work in healthcare settings suggests nonverbal communication behaviors such as eye gaze, smiling, and touch, play a critical role in the development of rapport between patients and providers (Caris-Verhallen, Kerkstra, & Bensing, 1999; DiMatteo, Taranta, Friedman, & Prince, 1980; Friedman, 1979a; Larsen & Smith, 1984). Further, the communication literature suggests that relational information is most often conveyed via paralinguistic (tone of voice) together with kinesthetic (nonverbal facial expressions or gestures) cues (Burgoon, 1994). Specific displays of certain nonverbal cues and gestures actually facilitate the development of coherent dialogue, and relational rapport (see Bavelas, 1994; Bavelas, Chovil, Coates, & Roe, 1995). Bavelas and Chovil (2000) demonstrate in recent experimental studies that face-to-face communication is comprised of both "audible and visual acts . . . completely interwoven" in interactional behavior (p. 164) and thus inseparable when making assessments of interactions as they develop. Indeed, the importance of visual input in understanding relationship-building has been shown across a wide variety of communication contexts such as medical interviews (Ruusuvuori, 2001), video versus audio conferencing (Daly-Jones, 1998), business interactions (Griffin, 1998), and video tutoring and initial teacher training (Nichol & Watson; 2000). Ford, Hall, Ratcliffe, and Fallowfield (2000) suggest videotaping is the "only observational strategy" (p. 557) that captures the variety of behaviors inherent in face-to-face interaction, most notably physician-patient exchanges.

Despite of the call for video-based data, the process of audiotaping interactions continues to be a widely used procedure (e.g., Greene, Adelman, Friedmann, & Charon, 1994; Roter, Geller, Bernhardt, Larson, & Doksum, 1999; Roter, Larson, Fischer, Arnold, & Tulsky, 2000), primarily because it is less costly and potentially less intrusive than is the process of videotaping the same events. And, audiotaped data have been shown to have acceptable reliability (Ford et al., 1996; Ong et al., 1998; Roter 1991).

Though video recording may be conceptually and intuitively more appealing than simple audiotaping alone, few direct comparisons exist for empirically determining the precise added value warranting a research decision to move from collecting audio-based data to recording and coding more complex video-based physician-patient communication data.

Thus, the purpose of this study is to compare coding results from audio versus video-based analyses of the same patient-physician interactions in the context of a life-threatening disease. It is hypothesized that the same coding systems applied to audio and video data will yield differing results (Ong et al., 1998) in terms of the consistency and the construct validity (as seen in the underlying factor structures of the data) of the ratings. Assessments of videotaped data in this study are expected to yield ratings that are more consistent with theory-based expectations about aspects of the encounters and their outcomes than will assessments based on audiotaped data alone.

Method

Participants

The physician-patient interaction data were taken from a convenience sample collected at the H. Lee Moffitt Cancer Center and Research Institute from 1996–1997 (see Albrecht et al., 1999). The 47 patients (25% male, 75% female; average age = 58) recruited for the study were attending one of several multidisciplinary clinics (thoracic, malignant hematology, breast, neuro-oncology, pain management, senior adult, and gastrointestinal oncology). Only one patient refused to participate for personal reasons. Twelve medical oncologists also participated in the study (two women, 10 men, average age = 55 years; all those approached agreed to take part in the research).

The interactions were videotaped during the time the physician (and, in four cases, a research nurse) presented the patient with the option to enroll in a clinical trial as a treatment choice. The procedure for the taping was described to each patient, physician, and research nurse (if present) and informed consent was obtained from each study participant. Two small video cameras were arranged unobtrusively in the exam room to record the physician and the patient during the encounter. At the end of the interaction, the tapes from the cameras were inserted into videocassette recorders (VCRs) and processed through an audiovisual mixer unit enabling synchronized, split-screen viewing of the entire exchange between the physician/health care provider and the patient on a single video monitor. The split view was then duplicated onto a standard VHS cassette, suitable for insertion into standard VCR units for playback and coding.

Measures

Moffitt Accrual Analysis System (MAAS; Albrecht et al., 1999). This observational coding system is designed to assess video-recorded communication behaviors occurring between oncologists and patients, during oncology consultations in which patients are presented with a clinical trial as a treatment option (and thereby considering alternatives to selecting standard therapy). The MAAS contains a set of global subjective ratings that focus on interaction-level variables (e.g., trust, hierarchical rapport) and a behavioral checklist. Only the global ratings were used for the current study, as they are more relevant for analyzing relationship and rapport-building behaviors. The global items (rated on a 1–7 scale) include the following: *Hierarchical rapport* (extent to which the physician preserves his/her status as the medical expert with arrogance or with cordiality shown toward the patient; *Connectedness/closeness* (extent there appears to be a warm relationship between the physician and patient during the interaction); *Trust level* (degree the patient appears to have confidence in the physician's integrity, ability, and judgment and the extent to which the physician seems to recognize and respond); *MD code* (degree to which physician uses technical language in conversing with patient versus using nontechnical, layperson-oriented language/terminology); *PT code* (degree to which patient uses technical language in conversing with physician versus using nontechnical, layperson-oriented language/terminology); *MD* *responsiveness to PT concerns/questions* (extent to which physician invites and responds with reassurance/resources to patient comments, questions, and concerns regarding the trial option); *MD directedness* (extent to which physician moves the discussion toward guiding the patient to sign the consent form); *Conformity to legal consent form* (degree to which physician's discourse conforms to the language, structure, format of the formal consent form); *MD dominance of floor time* (degree to which physician talks vs. engaging in conversational turn-taking); *PT dominance of floor time* (degree to which patient talks versus engaging in conversational turn-taking); *Amount of information given to PT* (extent to which physician gives patient appropriate or inappropriate amount of information, either too much or too little, for the patient's needs); *MD manner of delivery* (extent to which physician appears prepared and structured in his/her presentation of treatment options vs. appearing unplanned and disorganized); *MD information orientation/opinion-based* (extent to which evaluations or conclusions provided by physician to patient are framed as his/her personal opinion); *MD information orientation/data-based* (extent to which evaluations or conclusions provided by physician to patient are framed as accepted scientific findings).

Roter Interaction Analysis System (RIAS; Roter, 1991). This observational behavior coding system was originally designed for analyzing audiotaped interactions between physicians and patients during primary care visits. It has been used more recently in videotaped studies (Ong et al., 1998) and also has been extended to analyzing physician-patient communication in cancer care (e.g., Ford, Fallowfield, & Lewis, 1996; Ong et al., 1998). As with the MAAS, the subjective ratings component of the scoring system was used in the present study, given the relevance for obtaining relational judgments of the physician-patient communication. Ratings are made separately for the physician and the patient in each encounter on the following items (using a 1–6 scale): *Anger/Irritation, Dominance/Assertiveness; Interest/Attentiveness; Friendliness/Warmth; Responsiveness/Engagement; Sympathetic/Empathetic; Enabling/Involving; Energy; Overall Disruption/Distraction*. One item, *Physician Technical Talk*, is coded for the physician only.

Procedure

Eight coders, blind to the purpose of the study, were trained to use the MAAS and RIAS coding systems. To examine our hypotheses, four coding conditions were established: MAAS audio, RIAS audio, MAAS video, RIAS video. Four raters were randomly assigned to code the audio conditions (the videotapes were played in a 13″ TV/VCR with the TV screens covered). Of these, two raters were randomly assigned to code approximately half the tapes using the MAAS and the other half using the RIAS. The remaining two raters coded the tapes that had previously been rated using the MAAS, with the RIAS; conversely, they rated the tapes that had been rated with the RIAS, with the MAAS. This approach was also applied to the video conditions. Four raters coded the MAAS and the RIAS using the videotaped data. Each tape, then, was rated eight times by different raters (two raters coded MAAS audio, two coded RIAS audio, two coded MAAS video, and two coded RIAS video). Each coder rated approximately 44 tapes (missing data due to technical malfunction and limitations of rater availability) in the same format, either in the audio or video format.

Analysis

Reliability

Selecting the appropriate statistic to use in assessing reliability is determined by the type of data being examined. Although Pearson's *r* and Cohen's kappa have typically been used to assess rater agreement of interaction data (e.g., Ong et al., 1998; Roter, 1991), in the case of assessing agreement between raters coding continuous variables, a number of authors argue the intraclass correlation coefficient (ICC) is the more appropriate statistic (Streiner, 1995; Bartko & Carpenter, 1976; Bartko, 1976; Shrout & Fleiss, 1979; Bédard, Martin, Krueger & Brazil, 2000; Saw & Ng, 2001).

ICCs have been used previously to evaluate interrater agreement in healthcare contexts including: rater observations of sedation recovery made from videotaped data, (Macnab, Levine, Glick, Phillips, Susak, & Elliott; 1994), clinical ratings made by experienced psychiatrists and trainees (Crippa, Sanches, Hallak, Loureiro, & Zuardi, 2001), concordance of self assessment and caregiver assessment of patient symptoms (Fransson, Tavelin, & Widmark, 2001). Most relevant for the present investigation, Ford et al. (2000) used ICC to assess the interrater agreement of their behavioral observation rating scale, the Medical Interaction Process System.

In this study, ICCs were computed to assess interrater reliability for the MAAS (audio and video data) and the RIAS (audio and video data) coding schemes, using Shrout and Fleiss's (1979) ICC Model 1 (a random effects ANOVA model). ICCs were calculated as the ratio of variance between participants over the total between-and-within participant variance. Hence, if rater agreement is high, then within- participant variance is low, resulting in a high ICC.

Comparison of Audiotaped Versus Videotaped Formats

Several analyses were used to test for differences in coder ratings using the same coding scheme, but with different observational format (audiotape versus videotape). The following analyses were applied to both the MAAS and the RIAS. First, MANOVA was used to determine if differences in coding occurred due to format. MANOVA was used to avoid the possibility of an experiment-wide error rate caused by computing many different ANOVAs. Based on the MANOVA results, univariate analyses were conducted to assess the effects of format on specific variables.

Additionally, exploratory factor analysis (following guidelines by MacCallum, Widaman, Zhang, & Hong, 1999) was employed to analyze the underlying structure of the data obtained from rating each coding scheme using different formats. Following established procedures for determining the number of factors (Rummell, 1970; Fabrigar, Wegener, MacCallum, & Strahan, 1999) preliminary solutions were evaluated using eigenvalues, scree plots, and the amount of total variance explained (using the STATISTICA program with assumptions similar to those of SAS, SPSS, and BMDP). The factor loadings were generated by maximum likelihood and are reported in Tables 5 through 8.

Results

Reliability

Reliability estimates for the MAAS audio-based ratings range from .60 to .99, with a mean ICC of .89; the video-based ratings using the MAAS range from .67 to .99, averaging .87 (see Table 1).

For audio ratings using the RIAS, reliabilities range from .29 to .99 with a mean of .78; video-based reliabilities range from .40 to .99 with a mean reliability of .86 (Table 2). Both formats yield good reliability (using the guidelines set by Streiner & Norman, 1995, an ICC of .75 or higher is considered acceptable).

To compare reliabilities across audio and video conditions for the MAAS and the RIAS, ICCs were transformed to z-scores. Mean z-scores were computed for each coding condition and compared using Fisher's *z* transformation. Based on Fisher's z transformations, the difference between

TABLE 1

MAAS—Average Reliability Results (Intraclass Correlations) for Audio Vs. Video Data (n = 47)

	Intraclass correlations	
Item	Audio	Video
Hierarchical Rapport***	0.60	0.99
Connectedness/Closeness**	0.86	0.67
Trust**	0.90	0.68
MD Code	0.90	0.90
PT Code	0.87	0.84
Physician Responsiveness***	0.99	0.88
Physician Directedness (toward consent)	0.83	0.91
Conformity to Legal Consent Form*	0.94	0.84
MD Dominance	0.88	0.82
PT Dominance***	0.96	0.99
Amount of Information Given to PT***	0.97	0.84
MD Manner of Delivery**	0.96	0.86
MD Info Orientation: Opinion**	0.87	0.96
MD Info Orientation: Data***	0.87	0.97

Note: Asterisks indicate a significant difference between audio and video ICCs based on Fisher's z transformations.
*$p < .05$; **$p < .01$; ***$p < .001$.

audio and video ICCs were not significant for either the MAAS ICCs (difference z = .07, ns) or the RIAS ICCs (difference z = .48, ns).

Comparison of Audiotaped Versus Videotaped Formats

As hypothesized, MANOVA revealed a significant effect of format for the MAAS coded data (audio versus video; $F\,(14, 79) = 1.997$; $p < .028$). Univariate analysis showed differences between audio and video ratings for connectedness/closeness, physician dominance, and appropriate amount of information given by the physician. Means and standard deviations of audio and video assessments of MAAS variables along with significant *F* values are provided in Table 3.

The hypothesis is further supported by the MANOVA results for the RIAS coded data. The structure of the RIAS as a printed coding guide leads

TABLE 2

RIAS—Average Reliability Results (Intraclass Correlations) for Audio Vs. Video Data

Item	Intraclass correlation	
	Audio	Video
Physician		
Anger/Irritation	0.61	0.73
Dominance/Assertiveness***	0.78	0.99
Interest/Attentiveness*	0.84	0.91
Friendliness/Warmth*	0.95	0.86
Responsiveness/Engagement***	0.95	0.99
Sympathetic/Empathetic*	0.99	0.97
Enabling/Involving	0.80	0.84
Energy***	0.74	0.93
Overall Disruption/Distraction***	0.44	0.94
Technical talk***	0.92	0.70
Patient		
Anger/Irritation**	0.54	0.84
Dominance/Assertiveness	0.99	0.99
Interest/Attentiveness***	0.93	0.66
Friendliness/Warmth	0.93	0.92
Responsiveness/Engagement***	0.99	0.93
Sympathetic/Empathetic	0.29	0.40
Enabling/Involving	0.94	0.95
Energy***	0.76	0.95
Overall Disruption/Distraction***	0.37	0.86

Note: Asterisks indicate a significant difference between audio and video ICCs based on Fisher's r to z transformations.

*$p < .05$; **$p < .01$; ***$p < .001$.

the coder to rate patient and physician dimensions separately; hence, separate MANOVAs were conducted on physician and patient variables. Significant main effects of format were found for both analyses: a significant effect of format on physician variables ($F\,(10, 83) = 4.41$, $p < .01$) and on patient variables ($F\,(9, 84) = 8.72$, $p < .01$). Specific differences occur for physician variables including: enabling, energy, disruption, and for patient variables including: anger, dominance, friendliness, sympathetic, enabling,

TABLE 3

Means and Standard Deviations for MAAS Item Ratings (Audio and Video)

	Audio (n = 47)		Video (n = 47)		
Item	*x*	*sd.*	*x*	*sd.*	*F*
Hierarchical Rapport	5.35	0.89	5.15	1.19	
Connectedness/Closeness	4.53	0.97	3.99	1.35	6.18**
Trust	4.95	0.86	4.92	1.16	
MD Code	4.30	0.93	4.07	1.32	
PT Code	5.83	0.90	6.09	1.02	
MD Responsiveness	5.46	0.77	5.19	1.08	
MD Directedness (toward consent)	4.98	0.96	5.18	1.31	
Conformity to Legal Consent Form	4.97	1.13	4.87	1.47	
MD Dominance	5.40	0.69	5.67	0.93	3.98*
PT Dominance	3.24	1.25	2.91	1.26	
Amount of Information Given to PT	5.74	0.79	5.37	1.18	4.18*
MD Manner of Delivery	5.28	0.99	5.24	1.18	
MD Info Orientation: Opinion	3.65	1.32	3.70	1.18	
MD Info Orientation: Data	5.02	1.19	5.02	1.06	

*$p < .05$; **$p < .01$.

energy, and disruption (see Table 4 for means, standard deviations and significant *F* values).

Examination of the means for the MAAS and RIAS variables discussed above shows that observers in the audio conditions rate variables higher than do observers in the video conditions. Audio-based means were higher for two of the three MAAS variables; audio versus video differences were found for ten of the RIAS variables. Further, those variables for which significant differences between audio and video were not found still demonstrated a similar trend of elevated audio ratings.

Given that the MANOVA reveals differences in mean ratings between audio and video assessment for both the MAAS and the RIAS coding systems, the next task was to determine whether differences existed in the underlying structures of the two data formats. Exploratory factor analysis was used to investigate whether additional differences exist relative to the

TABLE 4

Means and Standard Deviations for RIAS Item Ratings (Audio and Video)

	Audio (n = 47)		Video (n = 47)		
Item	*x*	*sd.*	*x*	*sd.*	*F*
Physician Variables					4.41**
Anger/Irritation	1.64	0.67	1.39	0.58	
Dominance/Assertiveness	4.03	0.90	3.85	0.68	
Interest/Attentiveness	4.03	0.80	4.00	0.76	
Friendliness/Warmth	3.64	0.83	3.52	0.96	
Responsiveness/Engagement	3.95	0.85	3.87	0.87	
Sympathetic/Empathetic	3.65	0.76	3.36	0.84	
Enabling/Involving	3.95	0.80	3.56	0.92	4.74*
Energy	3.81	0.96	3.39	1.07	3.99*
Overall Disruption/Distraction	2.66	1.04	1.74	0.85	22.1**
Physician Technical talk	2.92	0.67	3.12	0.85	
Patient Variables					8.72**
Anger/Irritation	1.93	0.71	1.62	0.64	4.83*
Dominance/Assertiveness	3.18	1.01	2.66	1.14	5.50*
Interest/Attentiveness	4.21	0.79	4.30	0.84	
Friendliness/Warmth	3.66	0.80	2.91	0.95	16.9*
Responsiveness/Engagement	3.88	0.89	3.57	1.03	
Sympathetic/Empathetic	2.84	0.80	2.44	0.84	5.50*
Enabling/Involving	3.66	0.95	3.10	0.90	8.76*
Energy	3.45	1.01	2.88	1.05	7.88*
Overall Disruption/Distraction	2.51	0.95	1.47	0.54	43.0*

*$p < .05$; **$p < .01$.

underlying structures of audio-based data versus the video-based data. Maximum likelihood estimates are provided for the MAAS solutions below (see Tables 5 and 6; for ease of discussion, only factor loadings of .35 or greater are presented).

The theory of factor analysis states that observed scores correlate to the extent they are influenced by an underlying common factor (construct). Factor analysis of the MAAS audiotaped data results in a three-factor solu-

TABLE 5

Factor Loadings for the MAAS Audio-Based Ratings

Item	Factors: *Relational affiliation*	*Physician task orientation*	*Physician information*
Hierarchical Rapport	.76		
Connectedness/Closeness	.67		
Trust	.76		
MD Code	.40		
PT Code*			
MD Responsiveness	.66		
MD Directedness (toward consent)		.62	
Conformity to Legal Consent Form		.92	
MD Dominance			
PT Dominance			
Amount of Information Given to PT			.43
MD Manner of Delivery			.45
MD Info Orientation: Opinion			−.63
MD Info Orientation: Data			.96

Note: Relational Affiliation explained variance (eigenvalue) = 2.44, 42% common variance; *Physician Task Orientation* explained variance (eigenvalue) = 1.54, 26% common variance; *Physician Information* explained variance (eigenvalue) = 1.86, 32% common variance. Three factor solution = 42% total variance.

*Factor loadings for this variable were <.35.

tion. As shown in Table 5, Hierarchical rapport, Connectedness/closeness, Trust, MD code, and MD Responsiveness load on Factor 1. This factor can be interpreted as reflecting the *Relational Affiliation* apparent between the physician and the patient. The second factor involves the extent to which the physician is directed toward consent in the interaction and conforms to the legal consent document in his/her verbal review. This factor reflects *Physician Task Orientation*. Finally, the third factor includes the appropriateness of the amount of information given to patients, the manner of delivery, and the physician's expression of opinion and/or expression of data in providing treatment alternatives; it can be labeled *Physician Information.*

The *Relational Affiliation* factor explains most of the common variance (42%; eigenvalue = 2.44), followed by the *Physician Information* factor (32%; eigenvalue = 1.86). The *Physician Task Orientation* factor explains the least amount of common variance (26%; eigenvalue = 1.54). When combined, the three factors account for 42% of the total variance in the audio ratings.

The analysis of the MAAS video data also yields a three-factor solution but as expected, the structure of the factors differs from that of the audio data. The first factor includes variables similar to those seen in the audio-based analysis (rapport, connectedness/closeness, trust, physician code, and the responsiveness of the physician) and can be also be labeled *Relational Affiliation*. The video-based *Relational Affiliation* factor however involves two additional items: physician and patient dominance.

Physician directedness toward consent, conformity to legal document, the MD and patient dominance items, appropriateness of the amount of information, and manner of delivery all load on the second factor reflecting a *Physician Task Orientation* factor very differently from the audio *Physician Task Orientation* factor. The third factor contains the MD opinion and MD data items, reflecting *Physician Reasoning Style.*

Together, the three factors explain 55% of the total variance in the video ratings. Examining the common variance in this analysis, the *Physician Task Orientation* factor explains 39% of the common variance (eigenvalue = 3.06); the *Relational Affiliation* factor explains 38% of the common variance (eigenvalue = 2.99), and the *Physician Reasoning Style* factor explains the least amount of common variance (22%; eigenvalue = 1.74).

As with the MAAS results, the factor analyses for the RIAS show differing structural solutions for the audio and video data. Even though the factor analyses yield predictably predominant factors of "patient" and "physician" for both audio and video analyses (based on the format of the RIAS coding guide noted earlier) differences in the solutions are still apparent.[1]

The coded audiotaped data demonstrate a three-factor solution (see Table 7): *Patient* (36% of common variance; eigenvalue = 3.80), *Physician* (40% of common variance; eigenvalue = 4.23), and *Physician-Patient Negative Affect* (24% of common variance; eigenvalue = 2.50). The *Patient* factor predictably includes six patient variables (mostly reflecting positive affect) and the *Physician* factor includes seven physician variables plus the patient friendliness variable (again, mostly positive affect).

TABLE 6

Factor Loadings for the MAAS Video-Based Ratings

	Factors		
Item	*Relational affiliation*	*Physician task orientation*	*Physician reasoning style*
Hierarchical Rapport	.73		
Connectedness/Closeness	.86		
Trust	.69		
MD Code	.44		
PT Code*			
MD Responsiveness	.74		
MD Directedness (toward consent)		.70	
Conformity to Legal Consent Form		.70	
MD Dominance	−.48	.46	
PT Dominance	.39	−.51	
Amount of Information Given to PT		.79	
MD Manner of Delivery		.82	
MD Info Orientation: Opinion			.98
MD Info Orientation: Data			−.80

Note: Relational Affiliation explained variance (eigenvalue) = 2.99, 38% common variance; *Physician Task Orientation* explained variance (eigenvalue) = 3.06, 22% common variance; *Physician Reasoning Style* explained variance (eigenvalue) = 1.74, 39% common variance. Three factor solution = 55% total variance.

*Factor loadings for this variable were <.35.

The third factor, *Physician-Patient Negative Affect* includes a hybrid of physician and patient variables. These three factors combined account for 55% of the total variance in RIAS audio ratings.

A four-factor solution results for the videotaped data (Table 8): *Patient* (41% of common variance; eigenvalue = 5.08), *Physician* (36% of common variance; eigenvalue = 4.38), *Patient Negative Affect* (11% of common variance; eigenvalue = 1.31), and *Physician Influence on Patient*

TABLE 7

Factor Loadings for the RIAS Audio-Based Ratings

Items	*Patient*	*Physician*	*Physician-patient negative affect*
Physician			
Anger/Irritation			.63
Dominance/Assertiveness		.88	
Interest/Attentiveness		.79	
Friendliness/Warmth		.58	−.35
Responsiveness/Engagement		.64	
Sympathetic/Empathetic		.55	0.84
Enabling/Involving		.57	
Energy		.80	
Overall Disruption/Distraction			.58
Physician Technical talk			−.44
Patient			
Anger/Irritation			.59
Dominance/Assertiveness	.93		
Interest/Attentiveness	.75		
Friendliness/Warmth	.39	.59	
Responsiveness/Engagement	.79		
Sympathetic/Empathetic			.55
Enabling/Involving	.70		
Energy	.84		
Overall Disruption/Distraction			.70

Note: Patient explained variance (eigenvalue) = 3.80, 36% common variance; *Physician* explained variance (eigenvalue) = 4.23, 40% common variance; *Physician-Patient Negative Affect* explained variance (eigenvalue) = 2.50, 24% common variance. The three factors account for 55% of the total variance.

Anger (12% of common variance; eigenvalue = 1.51). The *Patient* factor includes eight patient variables plus the MD talk variable (mostly positive affect); the *Physician* factor includes seven physician variables (showing positive affect). The third factor, *Patient Negative Affect*, includes negative loadings for patient interest, friendliness, and sympathy. The fourth factor, *MD Influences on Patient Anger*, includes physician dominance, talk, and

TABLE 8

Factor Loadings for the RIAS Video-Based Ratings

	Factors			
Items	*Patient*	*Physician*	*Patient negative affect*	*Physician influence on patient anger*
Physician				
Anger/Irritation		−.58		
Dominance/Assertiveness				−.39
Interest/Attentiveness		.80		
Friendliness/Warmth		.90		
Responsiveness/Engagement		.89		
Sympathetic/Empathetic		.78		
Enabling/Involving		.84		
Energy		.61		
Overall Disruption/Distraction				
Physician Technical talk	−.40			−.71
Patient				
Anger/Irritation				.51
Dominance/Assertiveness	.86			
Interest/Attentiveness	.64		−.44	
Friendliness/Warmth	.83		−.40	
Responsiveness/Engagement	.89			
Sympathetic/Empathetic	.43		−.78	
Enabling/Involving	.88			
Energy	.95			
Overall Disruption/Distraction	.53			

Note: Patient explained variance (eigenvalue) = 5.08, 41% common variance; *Physician* explained variance (eigenvalue) = 4.38, 36% common variance; *Patient Negative Affect* explained variance (eigenvalue) = 1.30, 11% common variance; *Physician Influence on Patient Anger* explained variance (eigenvalue) = 1.51, 12% common variance. The four factors account for 64% of the total variance.

patient anger. This four-factor solution accounts for 64% of the total variance in RIAS video ratings.

Discussion

The findings of this study show that ratings of audiotaped physician-patient interactions are not equivalent to ratings of videotaped encounters. Specifically, judges tend to code relational communication behavior based on audiotaped interaction data differently from the way they code videotaped interaction data. Though they are using the same coding system and analyzing the same physician-patient exchanges, coders' judgments result in differing interpretations of the nature of the communication behavior.

Differences in judges' judgments arise in terms of the general factor structures of of their scores. The results show non-isomorphic underlying factor structures for the audio and video datasets. For the MAAS, audio ratings yield the *Relational Affiliation, Physician Information*, and *Physician Task Orientation* factors, together accounting for 42% of the common variance. In contrast, three different factors emerge for the videotaped data accounting for 55% of the common variance: *Physician Task Orientation*, *Relational Affiliation*, and *Physician Reasoning Style*. The audio/video-based ratings for the MAAS may be viewed as having more construct validity given that the solution derived from the video data accounts for more variance in ratings than the solution derived from the audio data (55% versus 42%).

In integrating the theoretical model with the results, it is clear that the pattern of findings for the video-based MAAS ratings reflect a greater theoretical coherence than do the results for the audio-based MAAS ratings. Theories behind the development of the MAAS as an analytical tool for communication focus on the *content* of the interaction and the quality of the *relationship* given that communication processes are comprised of both content and relational dimensions of exchange (Watzlawick, Beavin, & Jackson, 1967). And, context influences the relative nature of both the content and relational dimensions (Gillotti et al., 2002). Thus, in the context of informed consent in oncology treatment consultations, it is expected the physician will emphasize informational content describing the alternative treatment decisions and demonstrate positive relational affect and support to offset the patient's feelings of uncertainty, fear, and misunderstanding (Albrecht & Adelman, 1984; Albrecht & Adelman, 1987; Albrecht & Goldsmith, in press; Mortensen, 1997).

The MAAS is better able with video data to capture the two key dimensions that theoretically should be driving the interactions (the *Physician Task Orientation* and the *Relational Affiliation* factors). In contrast, the MAAS is limited in its ability to capture both these constructs using audio data. A further example involves the factor loadings for the physician and patient dominance variables. They are not captured in the audio data (loadings are <.35) yet load on the *Relational Affiliation* and *Physician Task Orientation* factors in the video-based data in a coherent manner.

As noted earlier, the structure of the RIAS guides the coder to rate patient and physician dimensions separately; hence, the factor analysis results show clear factors of "patient" and "physician" for both types of analyses. These two factors more likely represent an instructional bias than they do the underlying structure of the data. That said, the structural solutions for the audio and video formats are still somewhat dissimilar.

As seen in Tables 3 and 4, the audio and video ratings of both MAAS and RIAS variables generally show acceptable reliabilities. Though there is not a standard procedure for evaluating differences between ICCs, Fisher's z transformation shows no difference in the ability of observers to reliably measure variables from audio and video information formats. In short, as expected, both techniques were reliable.

However, for both coding systems, the video scoring led to a more robust understanding of the interaction. What this implies is that though reliabilities may be acceptable, using coding systems for audiotaped data versus videotaped data may lead researchers to draw very different conclusions regarding the physician-patient interaction process. True, videotaping may be a more expensive and seemingly more intrusive research procedure that artificially biases participants' behavior. Though cost is a factor in videotaping, prices for the technology have significantly decreased in recent years. Further, research has shown that participants' knowledge of the presence of cameras does not significantly alter their behaviors or the interaction (see Albrecht et al., 1999; Ickes, 1994; Ickes & Tooke, 1988).

This study underscores empirically the idea that differences do exist in rater assessments of physician-patient interactions based on the availability of visual information. It is suggested from these results that the research question needs to drive the format of the data to be analyzed (audio versus video). Though additional research is needed to guide researchers in understanding when specifically video versus audio should be used, we speculate that if the research task is to examine some type of verbal behavior or conversational structure (e.g., question-asking, forms of address, specific information-giving) and the tone of that behavior, audiotaping is likely to be sufficient. Such a research task is elegant in its goal, and thus does not require a complex strategy (i.e., videotaping) to capture more data than are needed at face value. If, alternatively, the research task is focused on capturing a more complicated purpose (such as the process of relationship building, the negotiation of meaning, influence, and so on) where nonverbal cues are intricately interwoven with verbal behavior, videotaping is more likely to take hold of the more complex magnitude of data. For example, if our research interest focuses on patients' relative abilities to negotiate uncertainty during medical encounters, we will likely gain richer data from video than audio. The healthcare environment, and even more, the context of life threatening disease, increases individual sensitivity to nonverbal cues (Friedman, 1979a, 1979b) as patients actively seek verbal and nonverbal information, to reduce uncertainty. Analyses based on audio-only data may not be sufficiently sensitive to coders' interpretations of behaviors, especially when judgments need to be rendered to address incongruence between verbal and nonverbal cues.

In conclusion, this study by itself, does not settle the question of whether audiotaping or videotaping is the preferable format. Rather, these findings provide evidence that researchers should choose audio or video based on the research question and the desired focus. This analysis requires extension to larger data sets and other settings.

Note

1. Although we computed separate MANOVAs for the RIAS physician and patient variables, all variables were included in the factor analysis. This was done in order to present comparable factor analyses of the MAAS and RIAS coded data. Separate factor analyses of physician and patient variables provide similar results.

References

Albrecht, T.L., & Adelman, M.B. (1984). Social support and life stress: New directions for communication research. *Human Communication Research*, *11*, 3–32.

Albrecht, T.L., Adelman, M.B., and Associates (1987). *Communicating social support.* Newbury Park, CA: Sage.

Albrecht, T.L., Blanchard, C.G., Ruckdeschel, J.C., Coovert, M., & Strongbow, R. (1999). Strategic physician communication and oncology clinical trials. *Journal of Clinical Oncology*, *17*, 3324–3332.

Albrecht, T.L., & Goldsmith, D.J. (in press). Social support, social networks, and health. Forthcoming in T. Thompson et al. (Eds.), *Handbook of health communication.* Hillsdale, NJ: Erlbaum.

Bartko, J.J. (1979). Various intraclass correlation reliability coefficients. *Psychological Bulletin*, *83*, 762–765.

Bartko, J.J., & Carpenter, W.T. (1976). Methods and theory of reliability. *Journal of Nervous Mental Disorders*, *163*, 307–317.

Bavelas, J. B. (1994). Gestures as part of speech: Methodological implications. *Research on Language and Social Interaction*, *27*, 201–221.

Bavelas, J. B., & Chovil, N. (2000). Visual acts of meaning: An integrated message model of language in face-to-face dialogue. *Journal of Language and Social Psychology*, *19*, 163–194.

Bavelas, J. B., Chovil, N., Coates, L., & Roe, L. (1995). Gestures specialized for dialogue. *Personality and Social Psychology Bulletin*, *21*, 394–405.

Bédard, M., Martin, N., Krueger, P., & Brazil, K. (2000). Assessing reproducibility of data obtained with instruments based on continuous measurements. *Experimental Aging Research*, *26*, 353–365.

Blanchard, C.G., Labrecque, M. S., Ruckdeschel, J. C., & Blanchard, E.G. (1988). Information and decision-making preferences of hospitalized adult cancer patients. *Social Science and Medicine*, *27*, 1139–45.

Burgoon, J. (1994). Nonverbal signals. In M.N. Knapp, & G. R. Miller (Eds.), *Handbook of interpersonal communication, 2nd Ed*. (pp. 229–285). Thousand Oaks, CA: Sage.

Caris-Verhallen, W.M.C.M., Kerkstra, A., & Bensing, J. M. (1999). Non-verbal behaviour in nurse-elderly patient communication. *Journal of Advanced Nursing, 29*, 808–818.

Crippa, J. A., Sanches, R. F., Hallak, J. E., Loureiro, S. R., & Zuardi, A. W. (2001). A structured interview guide increases Brief Psychiatric Rating Scale reliability in raters with low clinical experience. *Acta Psychiatrica Scandanavia*, *103*, 465–470.

Daly-Jones, O. (1998). Some advantages of video conferencing over high-quality audio conferencing: Fluency and awareness of attentional focus. *International Journal of Human-Computer Studies, 49*, 21–58.

DiMatteo, M. R., Taranta, A., Friedman, H. S., & Prince, L. M. (1980). Predicting patient satisfaction from physicians' nonverbal communication skills. *Medical Care, 18*, 376–387.

Fabrigar, L. R., Wegender, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4,* 272–299.

Ford, S., Fallowfield, L., & Lewis, S. (1996). Doctor-patient interactions in oncology. *Social Science and Medicine, 42*, 1511–1519.

Ford, S., Hall, A., Ratcliffe, D., & Fallowfield, L. (2000). The Medical Interaction Process System (MIPS): An instrument for analyzing interviews of oncologists and patients with cancer. *Social Science and Medicine*, *50*, 553–66.

Fransson, P., Tavelin, B., & Widmark, A. (2001). Reliability and responsiveness of a prostate cancer questionnaire for radiotherapy-induced side effects. *Supportive Cancer Care, 9*, 187–98.

Friedman, H. S. (1979a). Nonverbal communication between patients and medical practitioners. *Journal of Social Issues, 35*, 82–99.

Friedman, H. S. (1979b). The interactive effects of facial expressions of emotion and verbal messages on perceptions of affective meaning. *Journal of Experimental Social Psychology, 15*, 453–469.

Gillotti, C., Thompson, T., & McNeilis, K. (2002). Communicative competence in the delivery of bad news. *Social Science and Medicine, 54*, 1011–1023.

Greene M.G., Adelman R.D., Friedmann E., & Charon R. (1994). Older patient satisfaction with communication during an initial medical encounter. *Social Science and Medicine, 38,* 1279–88.

Griffin, J. (1998). *How to say it at work: putting yourself across with power words, phrases, body language, and communication secrets.* Paramus, NJ: Prentice Hall.

Ickes, W. (1994). Methods of studying close relationships. In A. Weber, & J. Harvey, Eds., *Perspectives on close relationships* (pp. 18–44). Boston, MA: Allyn & Bacon.

Ickes, W., & Tooke, W. (1988). The observational method: Studying the interaction of minds and bodies. In S. Duck, & S. Hay (Eds.), *Handbook of personal relationships: Theory, research, and interventions* (pp. 79–97). Chichester, England: John Wiley.

Larsen, K. M., & Smith, C. K. (1984). Sequential nonverbal behavior in the patient-physician interview. *Journal of Family Practice*, *18*, 257–261.

MacCallum, R. C., Widaman, K. F., Zhang, S. B., & Hong, S. H. (1999). Sample size in factor analysis. *Psychological Methods, 4*, 84–99.

Macnab, A. J., Levine, M., Glick, N., Phillips, N., Susak, L., & Elliott, M. (1994). The Vancouver sedative recovery scale for children-validation and reliability of scoring based on videotaped instruction. *Canadian Journal Anesthesiology*, *41*, 913–918.

Mechanic, D., & Meyer, S. (2000). Concepts of trust among patients with serious illness. *Social Science and Medicine*, *51,* 657–668.

Mortensen, C. D. (1997). *Miscommunication.* Thousand Oaks, CA: Sage.

Nichol, J. D., & Watson K. E. (2000). Videotutoring, non-verbal communication and initial teacher training. *British Journal of Educational Technology*, *31,* 135–144.

Ong, L. M. L., Visser, M. R. M., Kruyver, I.P.M., Bensing, J.M., van den Brink-Muinen, A., Stouthard, J.M.L., Lammes, F.B., & de Haes, J.C.J.M. (1998). The Roter Interaction Analysis System (RIAS) in oncological consultations: Psychometric properties. *Psycho-Oncology*, *7*, 387–401.

Rosser, W. W., & Kasperski, J. (2001). The benefits of a trusting physician-patient relationship. *The Journal of Family Practice*, *50*, 329–330.

Roter, D.L. (1991). *The Roter Method of Interaction Process Analysis: RIAS Manual*. Baltimore, MD: Johns Hopkins University.

Roter D. L., Geller, G., Bernhardt, B. A., Larson, S.M., & Doksum T. (1999). Effects of obstetrician gender on communication and patient satisfaction. *Obstetrics and Gynecology, 93*, 635–641.

Roter, D.L., Larson, S., Fischer, G. S., Arnold, R. M., & Tulsky, J. A. (2000). Experts practice what they preach—A descriptive study of best and normative practices in end-of-life discussions. *Archives of Internal Medicine, 160,* 3477–3485.

Rummell, R.J. (1970). *Applied factor analysis*. Evanston, IL: Northwestern University Press.

Ruusuvuori, J. (2001). Looking means listening: Coordinating displays of engagement in doctor patient interaction. *Social Science and Medicine*, *52*, 1093–1108.

Saw, S.M., & Ng, T.P. (2001). The design and assessment of questionnaires in clinical research. *Singapore Medical Journal, 42,*131–5.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86,* 420–428.

Streiner, D.L. (1995). Learning how to differ: Agreement and reliability statistics in psychiatry. *Canadian Journal of Psychiatry, 40,* 60–66.

Streiner, D.L, & Norman, G.R. (1995). *Health measurement scales (2nd ed).* New York: Oxford University Press.

Watzlawick, P., Beavin, J., & Jackson, D. (1967). *Pragmatics of human communication*. New York: Norton.