Detecting Emotional Expression in Face-to-Face and Online Breast Cancer Support Groups

Anna Liess, Wendy Simon, and Maya Yutsis
Stanford University School of Medicine

Jason E. Owen
Loma Linda University

Karen Altree Piemme
Stanford University School of Medicine

Mitch Golant
The Wellness Community—National

Janine Giese-Davis
Stanford University School of Medicine

Accurately detecting emotional expression in women with primary breast cancer participating in support groups may be important for therapists and researchers. In 2 small studies ($N = 20$ and $N = 16$), the authors examined whether video coding, human text coding, and automated text analysis provided consistent estimates of the level of emotional expression. In Study 1, the authors compared coding from videotapes and text transcripts of face-to-face groups. In Study 2, the authors examined transcripts of online synchronous groups. The authors found that human text coding significantly overestimated Positive Affect and underestimated Defensive/Hostile Affect compared with video coding. They found correlations were low for Positive Affect but moderate for negative affect between Linguistic Inquiry Word Count (LIWC) and video coding. The implications of utilizing text-only detection of emotion are discussed.

Keywords: emotional expression, text analysis, breast cancer, group therapy, video coding

Women with breast cancer experience reductions in distress and pain, a decrease in emotional suppression, and increased restraint

of hostility after participation in face-to-face (F2F) groups that facilitate emotional expression (Classen et al., 2001; Giese-Davis et al., 2002). They also experience decreased distress, greater heart-rate habituation to writing, and declines in physical symptoms while expressing emotions in written text (Low, Stanton, & Danoff-Burg, 2006). Online synchronous groups (OSGs) for breast cancer combine group support and written text and may similarly decrease distress and pain (Lieberman et al., 2003; Winzelberg et al., 2003) but may lead to increased emotional suppression (Lieberman et al., 2003).

Detecting emotional expression in these groups may be important for therapists and researchers testing whether expression mediates outcomes. Researchers often analyze expression in text transcripts of therapy rather than use video analysis (Low et al., 2006). Likewise, in OSGs, therapists have only text on which to rely for emotional cues. Detecting emotion in text for either would be compromised if some emotion categories rely heavily on nonverbal channels.

Early research on emotion communication channels (i.e., voice, face, text content) indicated that people watching videotapes of social interactions were significantly more accurate (those reading a transcript not exceeding chance levels) when asked interpretive questions (Archer & Akert, 1977). Detecting deceptive negative affect relies on vocal, facial, and text content cues (O'Sullivan, Ekman, Friesen, & Sherer, 1985). However, honest positive affect is difficult to detect in the absence of vocal (Krauss, Appel, Morency, Wenzel, & Winton, 1981), facial, or body cues (O'Sullivan et al., 1985). Since a host of computer-mediated

interventions for cancer patients (e-mail, real-time text correspondence, online forums, and electronic support groups; Davison, Pennebaker, & Dickerson, 2000) are now in common practice, we thought it timely to investigate detection of emotional expression comparing human coding of videotape and text transcripts of F2F groups as well as human and text analysis (Pennebaker & Francis, 1999) of F2F groups and OSGs. By using trained human coders and strict reliability standards, we believe we offer the best-case scenario for detection of expression in both videotape and text.

Detecting emotional expression in therapy is a complex process using (a) observational coding from videotape (Giese-Davis, DiMiceli, Sephton, & Spiegel, 2006; Giese-Davis, Piemme, Dillon, & Twirbutt, 2005), (b) coding of transcripts (Grabhorn, 1998), (c) automated text analysis (Low et al., 2006), or (d) content analysis (LaBarge, Von Dras, & Wingbermuehle, 1998). Regardless of method, affect constructs are often labeled with the same words (e.g., positive, negative, or defensive affect) even when assumptions differ dramatically. Trained raters may be in the best position to accurately identify expression in therapy contexts, but instead automated text analysis is increasingly used. The use of these programs assumes that words carry psychological information independent of context (Krippendorf, 2004). Because expression is highly contextual, irony, sarcasm, and anxiety may be missed by text analysis (Davison et al., 2000). Few studies have compared methods. In our two studies, we compare levels of emotion detected from human coding of video versus text (Study 1) in F2F groups and whether text analysis correlates with human coding of similar constructs (Study 1 and 2) in F2F groups and OSGs.

Hypothesis 1: Differences in levels of affect.

Due to greater reliance on non-verbal cues, human text coding will underestimate some video-coding categories: validation, high and low affection (included in Positive Affect), high fear, sadness, direct anger (included in Primary Negative Affect), tension, micro-moment contempt, belligerence, disgust, and stonewalling (included in Defensive/Hostile Affect). However, because some categories may be overestimated by text coding, our tests are two-tailed (e.g., *interest*, which is coded from video when there is a genuine positive voice tone, not just when a woman asks a question). Our study was powered to test summary variables.

Hypothesis 2: Multimethod–multitrait consistency.

Though levels of video and text may differ, strong convergent correlations will indicate that coders utilized similar cues. The strength of correlations among categories will replicate within each method (Kenny & Campbell, 1989). These analyses are necessary to examine whether constructs are consistent across methods.

Exploration: Human and automated text analysis.

Human coding will correlate significantly with text analysis of similar constructs. These correlations will be larger in OSGs than F2F groups because communication must rely on clarity in text rather than on non-verbal cues.

Method

Therapy Model

Both studies were community/research collaborations (March 8, 2000–January 16, 2002) between The Wellness Community (TWC), Stanford University, and the University of California San Francisco (UCSF). In Study 1, we selected women randomized to TWC in a study comparing supportive–expressive (SET) groups with TWC groups. In Study 2, women participated in TWC OSGs (Lieberman et al., 2003).

TWC offered free F2F therapy in Study 1 as part of their ongoing support programs serving over 5,000 cancer patients each week. Groups of 12 participated in weekly, 2-hr groups led by one therapist for 16 weeks. In TWC's "patient-active" model, therapists encourage patients to (a) become empowered; (b) partner with physicians; (c) access resources; (d) make active choices in their recovery; and (e) reduce unwanted aloneness, loss of control, and loss of hope. In Study 2, TWC offered free OSGs where groups of 8 participated in two non-randomized 90-min groups (4 total) led by one therapist for 16 weeks. Women could access a private, 24-hr, TWC-based newsgroup by cohort. The same therapy model was used for both studies.

Participants

Recruitment for both studies was through collaborating community organizations and general advertising, and for Study 2 through online postings on breast-cancer-related sites. Women were eligible if they were over 18 years old, diagnosed with physician-confirmed primary breast cancer (Stages I–III, without metastasis or recurrence), English-literate, less than 18-months posttreatment, and had not attended more than 8 support-group sessions. Women lived in the San Francisco East Bay in Study 1 and in California and throughout the United States in Study 2. Study procedures were approved by institutional review boards at Stanford and UCSF. Participants signed written informed consent and physician contact consents. In Study 1, women could earn up to $40 for completing questionnaires, but no payment was given in Study 2.

In Study 1, of 108 women contacted, 92 consented, 66 completed baseline data, 45 were randomized at the Walnut Creek site (22 to TWC, 23 to SET) and 18 were randomized to a second community site in San Francisco (10 to The Cancer Support Community, 8 to SET). Some attrition occurred due to time delays associated with block randomization. Of 22 women randomized to TWC, 20 attended sessions in two groups. A videographer taped each session, focusing on the woman speaking. In Study 2, women registered at www.twc-chat.org which provided information, consent details, and an invitation to participate. Of 67 women recruited, 35 did not participate due to scheduling difficulties, 32 consented and completed online measures at baseline, and 26 women completed the 16-week measures. We randomly selected 2 of 4 groups ($N = 16$). The OSG closely mimics F2F group interactions. Demographic and medical variables for both studies are in Table 1. Final sample for Study 1 is $N = 20$ and for Study 2 is $N = 16$.

Human and Automated Coding

For each study, we coded Sessions 2, 6, 10, and 15 for two groups. For F2F groups, we coded each participant's expression by

Table 1

Demographic and Medical Characteristics for Primary Breast Cancer Patients Participating in F2F Breast Cancer Support Groups (N = 20) and in TWC OSGs (N = 16)

Characteristic	TWC F2F San Francisco, East Bay (%)	TWC OSG (%)
Age		
30–39 years	05	12
40–49 years	25	54
50–59 years	30	28
60 years	40	06
Marital status		
Married	60	67
Divorced/separated	30	17
Widowed	00	10
Single	10	06
Education		
High school or less	05	10
College/trade school	60	70
Graduate school	35	20
Stage at diagnosis		
I	35	22
II	55	56
III	10	12
Other	00	10

Note. F2F = face-to-face; TWC = The Wellness Community; OSG = online synchronous groups.

using Specific Affect (SPAFF) for Breast Cancer (Giese-Davis et al., 2005) for Videotape and for Text (Giese-Davis et al., 2005) following professional transcription. For OSGs, we coded transcripts by using SPAFF for Text. We also conducted Linguistic Inquiry Word Count (LIWC) text analysis (Pennebaker & Francis, 1999) for each participant-by-session segment. We used mean scores across four sessions per woman for analyses.

We used SPAFF for Breast Cancer and SPAFF for Text (Giese-Davis et al., 2005) for human coding and tested hypotheses with Positive Affect (affection, affection with touch, interest, validation, genuine humor, and excitement), Primary Negative Affect (direct anger, low and high sadness, verbalized and high fear), and Defensive/Hostile Affect (tension, tense humor, whining, disgust, micro-moment contempt, verbalized contempt, domineering, and belligerence; Giese-Davis et al., 2005). For video, we coded 1 woman at a time at least twice (68 person-by-tape segments: Mean

kappa = 0.70, SD = 0.09). For a kappa of .60 or higher, a coin toss determined which coder's data we used (50 segments). If kappa was below .55, it was recoded (9-by-3 and 3-by-5 coders). Six segments (kappa between .55 and .60) were consensus coded to maintain thresholds. We gave a transcriptionist the timing of speaking turns so that video and text segments were comparable in hours:minutes:seconds:frames. The median correlation between two coders of each F2F transcript was 0.54 (SD = 0.11) for percent time and 0.66 (SD = 0.14) for word count. Either coder's work provided the same magnitude of results. We used consensus coded OSG transcripts because median correlations were 0.58 (SD = 0.27) for percent time but 0.38 (SD = 0.34) for word count. Each separate emotional expression in a stream of data over time has a duration in seconds that we used to calculate percent-time data.

We used LIWC text analysis, which automatically matches each word to 1 or more of 82 language dimensions (Pennebaker & Francis, 1999). Summary scores are the number of words matching a dimension divided by total number of words. Current analysis focused on two dimensions: (a) Expression of Emotion included Positive Emotion (happy), Positive Feeling (joy, love), Optimism (pride, certainty), Assents (yes, OK), Question Marks, Negative Emotion (range of negative words), Sadness (grief, cry), Anger (pissed, hate), Anxiety (nervous, tense, afraid), and Negations (no, not); (b) Cognitive Mechanisms included Inhibition (always, never), Tentativeness (perhaps, might), and Discrepancy (should, would).

Data Analysis

We utilized percentages to assure comparability across methods (F2F groups ran for 2 hr while OSGs ran for 1.5 hr). We utilized percent time to compare video coding (which typically uses mean duration of time; Giese-Davis et al., 2006), with SPAFF text percent time. Time in transcripts was calculated as total time for each utterance divided by number of words. SPAFF text percent word count was used for correlations with LIWC (Table 2).

In Study 1, to compare methods of human coding, we used the non-parametric Friedman test (due to non-normal distributions) to compare summary variables for three related samples: SPAFF video percent time, SPAFF text percent time, and SPAFF text percent word count (Table 2 and 3; Figure 1). In the analysis of emotion, two measurement traditions exist: one based on duration

Table 2

Type of Coding Measurement by Study Question

	Face-to face (F2F) groups				Online synchronous groups (OSGs)	
	Human			Automated	Human	Automated
Measurement	SPAFF Video (time)	SPAFF Text (time)	SPAFF Text (word count)	LIWC (word count)	SPAFF Text (word count)	LIWC (word count)
Sample size	20	20	20	20	16	16
Hypothesis 1—differences in levels of affect	X	X	X			
Hypothesis 2—multimethod–multitrait consistency	X	X				
Exploration—human and automated text analysis	X		X	X	X	X

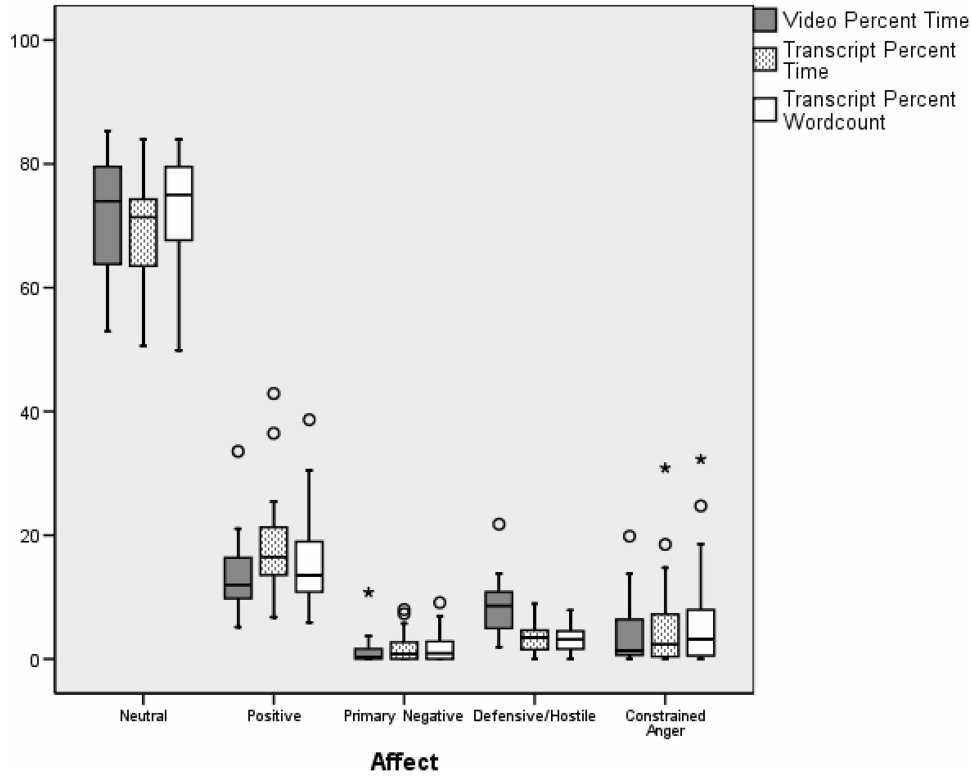*Note.* An X indicates we used these data to test this hypothesis.

Figure 1. Graphed are box-and-whisker plots for each summary measure of each affect coded for Study 1 from videotape (Specific Affect for Breast Cancer) and transcripts of the videotapes (Specific Affect for Text): video percent time, transcript percent time, and transcript percent word count. Bottom line on whisker = the smallest observation; bottom line on box = lower quartile; middle line on box = median; top line on box = upper quartile; top line on whisker = largest observation; circles = mild outlier; stars = extreme outlier.

of affect, one on word count. Because no prior comparisons give an indication of level differences between methods, we compared all three. If significant, we examined three pairwise comparisons with Wilcoxon signed ranks tests. To examine convergent correlations within and between methods, we utilized Spearman correlations (Table 3). We also explored whether the associations among SPAFF categories were consistent whether coded by video or text with a multitrait–multimethod matrix and Kenny and Campbell's method (Campbell & Fiske, 1959; Kenny & Campbell, 1989). If variables within both methods are similar, the same

Table 3
Median, 25th, and 75th Percentiles for Specific Affect Summary Codes for Videotape and Text in Women With Primary Breast Cancer (N = 20)

| | Face-to-face groups | | | | | | | | | |
| | Video percent time | | | Text percent time | | | Text percent word count | | | Friedman test |
Measure	25th	50th	75th	25th	50th	75th	25th	50th	75th	
Neutral	63.4	**74.0**	79.7	62.6	**71.4**	74.5	67.0	**75.1**	79.6	5.7
Positive Affect	9.6	**12.0**	16.4	13.2	**16.5**	21.6	10.8	**13.6**	19.3	25.9***
Primary Negative Affect	0.0	**0.4**	1.7	0.0	**0.8**	3.0	0.0	**0.9**	3.3	1.0
Defensive/Hostile Affect	4.8	**8.6**	10.8	1.5	**3.5**	4.7	1.7	**3.4**	4.6	17.1***
Constrained Anger	0.6	**1.4**	6.5	0.3	**2.3**	8.1	0.4	**3.2**	8.8	2.8

Note. We calculated the percent of time each individual spent expressing each type of affect, and we present the group distributions here. Because we utilized non-parametric analyses, we present the 25th, 50th (median in bold), and 75th percentiles for the distributions of three types of variables: video-coded percent time, transcript-coded percent time, and transcript-coded percent word count. For Positive Affect, video was significantly lower than either text category, and text categories were significantly different from each other. For Defensive/Hostile Affect, video was significantly higher than either text category, but text categories did not differ from each other.
*** $p < .001$.

patterns should emerge in all hetero-trait triangles, both within the mono-method triangle (Table 4, italicized numbers) and hetero-method triangle (Table 4, bold numbers). For instance, the size of correlation between variables 1 and 2 in method A ought to be similar to the correlation between variables 1 and 2 in method B. Lastly, we explored correlations among LIWC and SPAFF categories thought to represent similar constructs (Table 5) in F2F groups and OSGs.

Results

Similarity Between Coding Methods in F2F Groups (Table 3, Figure 1)

Significantly more Positive Affect was coded in transcripts than in videotapes. Levels of Primary Negative Affect were generally low and were not different between transcripts and videotape. Significantly less Defensive/Hostile Affect was coded in transcripts than in videotapes. We had no hypotheses about Neutral and Constrained Anger, but we present these for the possible interest of readers.

For Positive Affect, video percent time was significantly lower than both text percent time ($z = -3.65$, $p < .001$) and percent word count ($z = -2.14$, $p = .03$). Text percent time was significantly higher than percent word count ($z = -3.11$, $p = .002$). For Defensive/Hostile Affect, video percent time was significantly higher than either text percent time ($z = -3.43$, $p = .001$) or text percent word count ($z = -3.81$, $p < .001$). Text percent time and word count did not differ.

Multimethod–Multitrait Consistency

Results indicate that emotion constructs are consistent across methods except for Defensive/Hostile Affect. We found strong convergent validity (Table 4, bold italicized numbers) between methods for video and text percent time: Neutral, Positive Affect, Primary Negative Affect, and Constrained Anger, but not for Defensive/Hostile Affect. We found similar patterns of correlation levels among emotion categories other than those with Defensive/Hostile Affect. For instance, in both the mono-method triangles

Table 5

Spearman Correlations Between SPAFF Video Coding and SPAFF Text Word Count With LIWC Affect Categories

LIWC	SPAFF for Video (percent time) Video F2F ($N = 20$)	SPAFF for Text (percent word count) Text F2F ($N = 20$)	Text OSG ($N = 16$)
Positive Affect			
Positive Emotion	0.19	0.05	0.44
Positive Feeling	0.11	−0.29	0.36
Optimism	−0.34	0.11	−0.15
Assent	0.13	0.05	−0.32
Question Marks	0.19	0.28	0.31
Primary Negative Affect			
Negative Emotion	0.59**	0.60**	0.02
Sadness	0.48*	0.41	0.27
Anger	0.01	0.24	−0.03
Anxiety	0.13	0.06	0.15
Constrained Anger			
Negative Emotion	0.58**	0.74***	−0.14
Sadness	0.50*	0.64**	0.15
Anger	0.09	0.42	−0.01
Anxiety	0.08	0.39	0.14
Defensive/Hostile Affect			
Anxiety	0.30	−0.19	0.34
Negate	0.26	0.12	0.52
Inhibit	0.10	0.12	0.12
Tentative	0.21	−0.32	0.30
Discrepancy	0.34	0.46*	0.28

Note. Categories along the top of this correlation table are from Specific Affect (SPAFF) Coding (human coding) either from videotape or from text transcripts (using word count) of the same sessions. Categories along the side of this correlation table are from text analysis using the Linguistic Inquiry Word Count (LIWC) with the same text transcripts. F2F = face-to-face; OSG = online synchronous group.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 4

Spearman Correlations Among Emotion Categories for Specific Affect for Breast Cancer Video (Percent Time) and Specific Affect for Text (Percent Time)

Emotion category	1	2	3	4	5	6	7	8	9	10
1. Video Neutral Affect	—									
2. Video Positive Affect	−.21	—								
3. Video Primary Negative Affect	−.56*	−.24	—							
4. Video Defensive/Hostile Affect	−.51*	−.36	.27	—						
5. Video Constrained Anger	−.72**	−.19	.83**	.48*	—					
6. Text Neutral	**.60**	**−.07**	**−.32**	**.08**	**−.24**	—				
7. Text Positive Affect	**−.08**	**.61**	**−.11**	**−.00**	**−.64**	−.25	—			
8. Text Primary Negative Affect	**−.33**	**−.30**	**.70**	**.12**	**.57**	−.29	−.43	—		
9. Text Defensive/Hostile Affect	**−.17**	**−.23**	**.06**	**.16**	**.45***	.10	−.25	.32	—	
10. Text Constrained Anger	**−.36**	**−.29**	**.57**	**.15**	**.67**	−.26	−.51*	.46*	.03	—

Note. $N = 20$. Kenny and Campbell's multitrait–multimethod matrix to examine the consistency of constructs across methods. Hetero-trait triangles: mono-method (italicized numbers); hetero-method triangles (bold numbers); convergent validity (bold italicized numbers). Video = video percent time; text = text percent time.
*$p < .05$. **$p < .01$.

(Table 4, italicized numbers) for video (emotion categories 1–5) and text coding (emotion categories 6–10), and the hetero-method video- and text-coding block (Table 4, bold numbers), Primary Negative Affect is moderately negatively correlated with Positive Affect, rs = −.24, −.39, and −.27, respectively.

SPAFF and LIWC Correlations (Table 5)

Positive affect variables from LIWC does not correlate significantly with SPAFF Positive Affect coded from video (r = −.34 to .19) or text (r = −.29 to .28) in F2F groups. In OSGs, where participants use emoticons to increase clarity, correlations between SPAFF and LIWC are higher but not significant for Positive Emotion, Positive Feeling, and Questions Marks. There are four negative correlations. Negative affect variables from LIWC correlates moderately with SPAFF F2F video and text Primary Negative Affect (r = .01 to .60, three significant) and Constrained Anger (r = .08 to .74, four significant), indicating that they may measure similar constructs. For negative affect variables in OSGs, SPAFF and LIWC correlations are not significant, and three are negative. LIWC variables thought to be similar correlate moderately with SPAFF F2F video and text Defensive/Hostile Affect (r = −.32 to .46). For Defensive/Hostile Affect in OSGs, SPAFF and LIWC correlations are moderate but not significant.

Discussion

We compared levels of affect detected by human coders in F2F TWC groups and found that text coding overestimated Positive Affect and underestimated Defensive/Hostile Affect compared with video coding. Differences are likely because text cannot convey intonation, facial expression, and body posturing. We also examined correlations between human coding and automated text analysis in TWC F2F groups and OSGs, finding significant positive correlations for Primary Negative Affect, Constrained Anger, and Defensive/Hostile Affect, but none for Positive Affect.

Our research indicates that genuine positive affect is difficult to judge accurately from text in OSG and F2F groups and supports an earlier finding that accurate assessment of positive affect is more related to voice tone than content (O'Sullivan et al., 1985). We were surprised that significantly more Positive Affect was detected by coders in text than videotape and that text analysis correlated so poorly with human coding. An examination of video and text segments of mismatches indicated that a statement may seem positive on paper but may be interpreted as Defensive/Hostile Affect or Primary Negative Affect in the presence of defensive body posturing, a raised voice, or tears in the eyes. One can only speculate about how the probable increase in perception of positive affect and lack of cues for defensive/hostile affect might affect OSGs.

Non-verbal cues may be crucial for detecting Defensive/Hostile Affect because human coders identified significantly less from text than videotape, and convergent validity was low. If, like our coders, OSG therapists cannot detect hostility accurately, a lasting impact on participants' emotion regulation may be curtailed.

Rates of Primary Negative Affect detected in video and text were similar, and correlations between human coding and text analysis were higher, implying that Primary Negative Affect is conveyed to a greater extent through content. We found equally low levels in both F2F groups and OSGs.

This study was limited by small sample sizes, and the lack of economic and cultural diversity may have restricted the range of affect. Future research could randomize women to F2F groups versus OSGs so that statistical comparison is possible. However, these preliminary studies indicate that video and text methods of detecting emotion yield substantially different results.

We highlight ways in which the interpretation of written text can be changed by non-verbal cues in the following example. A common training technique for SPAFF is to ask the trainee simply to open a book to a random sentence. The trainee reads that sentence with the facial muscle movement, voice tone, and body movement of each one of the 20+ expressions coded by SPAFF. The same words can thus convincingly be used to convey emotional expressions as varied as validation, contempt, domineering, sadness, affection, and joy. Given that written text can be interpreted so broadly, examination of the coherence between methods of coding emotional expression, and caution when interpreting words in text as evidence of a particular emotion, seem crucial.

Based on these results, we recommend that researchers use caution when assessing computer-based and electronic psychotherapy services that depend solely on text. We also recommend that OSG therapists receive training to increase attention to cues for emotion in the absence of voice tone, facial movement, and body movement. OSG therapists may need to frequently double check their perception of affect to counteract their own tendency to see greater positive affect than is warranted (as did our human text coders) in order for these groups to be effective.

References

Archer, D., & Akert, R. M. (1977). Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of Personality and Social Psychology, 35,* 443–449.

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait–multimethod matrix. *Psychological Bulletin, 56*(2), 81–105.

Classen, C., Butler, L. D., Koopman, C., Miller, E., Dimiceli, S., Giese-Davis, J., et al. (2001). Supportive–expressive group therapy and distress in patients with metastatic breast cancer: A randomized clinical intervention trial. *Archives of General Psychiatry, 58*(5), 494–501.

Davison, K. P., Pennebaker, J. W., & Dickerson, S. S. (2000). Who talks? The social psychology of illness support groups. *American Psychologist, 55*(2), 205–217.

Giese-Davis, J., DiMiceli, S., Sephton, S., & Spiegel, D. (2006). Emotional expression and diurnal cortisol slope in women with metastatic breast cancer in supportive–expressive group therapy. *Biological Psychology, 73*(2), 190–198.

Giese-Davis, J., Koopman, C., Butler, L. D., Classen, C., Cordova, M., Fobair, P., et al. (2002). Change in emotion regulation-strategy for women with metastatic breast cancer following supportive–expressive group therapy. *Journal of Consulting and Clinical Psychology, 70*(4), 916–925.

Giese-Davis, J., Piemme, K. A., Dillon, C., & Twirbutt, S. (2005). Macrovariables in affective expression in women with breast cancer participating in support groups. In J. Harrigan, K. R. Scherer, & R. Rosenthal (Eds.), *Nonverbal behavior in the affective sciences: A handbook of research methods* (pp. 399–445). Oxford, England: Oxford University Press.

Grabhorn, R. (1998). Affective experience in a case of group therapy with psychosomatic inpatients. *Psychoanalytic Inquiry, 18*(3), 490–511.

Kenny, D. A., & Campbell, D. T. (1989). On the measurement of stability in over-time data. *Journal of Personality, 57*(2), 445–481.

Krauss, R. M., Appel, W., Morency, N., Wenzel, C., & Winton, W. (1981). Verbal, vocal, and visible factors in judgments of another's affect. *Journal of Personality and Social Psychology, 40*(2), 312–320.

Krippendorf, K. (2004). *Content analysis.* Thousand Oaks, CA: Sage.

LaBarge, E., Von Dras, D., & Wingbermuehle, C. (1998). An analysis of themes and feelings from a support group for people with Alzheimer's disease. *Psychotherapy: Theory, Research, Practice, Training, 35*(4), 537–544.

Lieberman, M. A., Golant, M., Giese-Davis, J., Winzelberg, A., Benjamin, H., Humphreys, K., et al. (2003). Electronic support groups for breast carcinoma: A clinical trial of effectiveness. *Cancer, 97,* 920–925.

Low, C. A., Stanton, A. L., & Danoff-Burg, S. (2006). Expressive disclo sure and benefit-finding among breast cancer patients: Mechanisms for positive health effects. *Health Psychology, 25,* 181–189.

O'Sullivan, M. M., Ekman, P., Friesen, W. V., & Sherer, K. (1985). What you say and how you say it: The contribution of speech content and voice quality to judgments of others. *Journal of Personality and Social Psychology, 48*(1), 54–62.

Pennebaker, J. W., & Francis, M. E. (1999). *Linguistic inquiry and word count.* Mahwah, NJ: Erlbaum.

Winzelberg, A. J., Classen, C., Alpers, G. W., Roberts, H., Koopman, C., Adams, R. E., et al. (2003). Evaluation of an internet support group for women with primary breast cancer. *Cancer, 97*(5), 1164–1173.