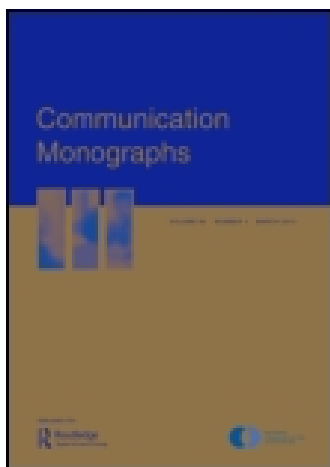


This article was downloaded by: [130.39.209.60]

On: 03 November 2014, At: 16:44

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Communication Monographs

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/rcmm20>

### Examining the Construct Validity of Enacted Support: A Multitrait-Multimethod Analysis of Three Perspectives for Judging Immediacy and Listening Behaviors

Graham D. Bodie, Susanne M. Jones, Andrea J. Vickery, Laura Hatcher & Kaitlin Cannava

Published online: 22 Sep 2014.

To cite this article: Graham D. Bodie, Susanne M. Jones, Andrea J. Vickery, Laura Hatcher & Kaitlin Cannava (2014) Examining the Construct Validity of Enacted Support: A Multitrait-Multimethod Analysis of Three Perspectives for Judging Immediacy and Listening Behaviors, Communication Monographs, 81:4, 495-523, DOI: [10.1080/03637751.2014.957223](http://dx.doi.org/10.1080/03637751.2014.957223)

To link to this article: <http://dx.doi.org/10.1080/03637751.2014.957223>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

# Examining the Construct Validity of Enacted Support: A Multitrait–Multimethod Analysis of Three Perspectives for Judging Immediacy and Listening Behaviors

Graham D. Bodie, Susanne M. Jones, Andrea J. Vickery,  
Laura Hatcher & Kaitlin Cannava

*Scholars of supportive communication are primarily concerned with how variations in the quality of enacted support affect individual and relational health and well-being. But who gets to determine what counts as enacted support? There is a large degree of operational heterogeneity for what gets called enacted support, but little attention has been afforded to the issue of whether these assessments are substitutable. In two studies we use self-reports, conversational partner-reports, and third-party ratings of two quintessential behavioral support indicators, namely, listening and immediacy. Using a multitrait–multimethod (MTMM) design, Study 1 found (1) little association between the enacted support assessments and (2) a high degree of common method variance. A second study found moderate-to-high degrees of effective reliability (i.e., consistency of judgments within a set of judgments, or mean judgments) for enacted support evaluations from the perspective of unacquainted and untrained third-party judges. In general, our data provide cautionary evidence that when scholars examine evaluations of enacted support, perspective matters and might ultimately contribute differently to well-being and health.*

---

Graham Bodie is Associate Professor in the Department of Communication Studies, Louisiana State University and Agricultural & Mechanical College and Visiting Associate Professor in the School of Media and Communication, Korea University, Seoul, South Korea. Susanne M. Jones is Associate Professor and Donald V. Hawkins Professor of Communication in the Department of Communication Studies, University of Minnesota, Twin Cities. Andrea Vickery, Laura Hatcher, and Kaitlin Cannava are doctoral candidates at LSU A&M. Correspondence to: Graham Bodie, Department of Communication Studies, LSU A&M, Baton Rouge, LA 70803, USA. Email: [gbodie@lsu.edu](mailto:gbodie@lsu.edu).

Data presented in Study 1 were presented as part of a separate manuscript at the 2012 convention of the International Communication Association, Phoenix, AZ. A previous version of this manuscript was awarded with a top paper distinction at the 2013 convention of the National Communication Association, Washington, DC.

*Keywords:* Active Listening; Comfort; Stress; Judgment Studies; Perceived Support; Social Support

Studies of social support have generated one resounding finding: The availability of positive social support is beneficial to mental and physical well-being (Sarason & Sarason, 2009; Uchino, Carlisle, Birmingham, & Vaughn, 2011). Communication research in particular has demonstrated that enacted support—what is said and done in the service of helping a distressed other—contributes to our ability to cope with stressful events (for review, see MacGeorge, Feng, & Burleson, 2011). Still, a critical problem remains: A close inspection of the enacted support literature reveals a lack of consensus regarding the methodological specificity for evaluating enacted support. For example, several studies have explicitly stated a focus on enacted support, yet operationalized enacted support with recalled “partner responsiveness” (Gable, Gosnell, Maisel, & Strachman, 2012, p. 967) or simply “perceived received support” (Chen & Feeley, 2012, pp. 610, 614). We present two studies that examine three frequently used judgment perspectives that operationalize evaluations of enacted support: Self-reports, conversational partner-reports, and third-party observer assessments. We examine these perspectives with two behavioral enactments of support, immediacy and listening. Our objective is to show that different perspectives used to assess enacted support behaviors, such as immediacy and listening, are not isomorphic and render different kinds of information. The studies we present here might ultimately bear on questions such as what “kind” of enacted support is beneficial to health and human functioning. We begin by conceptualizing enacted support and then advance a rationale for our studies.

### Enacted Support as Communication

In line with Goldsmith (2004), we define enacted support as those verbal and nonverbal manifestations of social support that serve to help another person cope with stressful experiences and emotions. Making conceptual space for *communicatively* enacted support means we must acknowledge that “social support ... is ultimately conveyed through messages directed by one individual to another” and is not merely a perceptual variable (MacGeorge et al., 2011, p. 323). There are, however, many examples in the literature of researchers claiming to examine *enacted support* but that ultimately conflate communicated support with support that was received or provided (Chen & Feeley, 2012; Gable et al., 2012; Lakey, Orehek, Hain, & VanVleet, 2009).

We concur with Chen and Feeley (2012) that observing enacted support in “real-world settings” is difficult (p. 610). In addition, we acknowledge that enacted support is complex and consists of various forms (e.g., emotional, informational, and tangible) and functions (e.g., dispensing information and advice, fostering a sense of belonging or inclusion, and enhancing the recipient’s self-esteem) that can be executed more or

less skillfully (Burleson, 2003). Nevertheless, we argue that assessing what is labeled enacted support with different perspectives has consequences. Although extensive research suggests that patterns of variability in the enactment of support reliably “influence various cognitive, affective, behavioral, and physiological processes associated with physical and psychological health” (MacGeorge et al., 2011, p. 325), others have questioned the utility of the enacted support construct (Lahey & Orehek, 2011) because it often fails to predict important outcomes.

Two behavioral clusters that have been frequently used to capture enacted support are immediacy and listening behaviors (Goldsmith, 2004, p. 20), which seem nomologically connected (Cronbach & Meehl, 1955). For instance, each behavioral cluster is thought to be an observable mechanism through which people form bonds and convey affective concern; each is also operationally defined with similar behaviors, such as eye contact and head nods (Andersen & Andersen, 2005; Bodie, St. Cyr, Pence, Rold, & Honeycutt, 2012; Floyd, 2014; Jones, 2011).

The connection between nonverbal immediacy (NI) and enacted support is well supported empirically (Jones & Guerrero, 2001; Trees, 2000). Several theories also make specific claims about the connection between immediacy cues, social support networks, and positive outcomes (Anders & Tucker, 2000; Bernieri & Gillis, 2001; Fredrickson, 2013; Gable et al., 2012; Tickle-Degnen & Rosenthal, 1992). Thus, the inclusion of NI as a manifestation of enacted support seems relatively unproblematic.

Although the theoretical link between listening and enacted support is not as well established, scholarly prescriptions often freely dispense advice that “good support” must be based on “active listening” behaviors (e.g., Burleson, 2003, 2007; MacGeorge, Feng, & Thompson, 2008). Such advice is not unwarranted because lay conceptualizations of “supportive people” and “supportive listeners” are “strikingly similar” (Bodie, Vickery, & Gearhart, 2013, p. 47). In addition, third-party observers rate support providers trained to deliver more sophisticated forms of verbal and nonverbal support as better listeners than those enacting substandard support (Bodie & Jones, 2012). Thus, several measures used in the literature include items that explicitly assess listening, such as the Inventory of Socially Supportive Behaviors (e.g., “Listened to you talk about your private feelings”), a measure of received support that has been used as a proxy for enacted support (Barrera & Baca, 1990). Likewise, interaction-based studies include listening as a primary coding category (Feeney, Cassidy, Lemay, & Ramos-Marcuse, 2009; Feeney & Collins, 2001) and emphasize listening when training confederates to engage in more or less supportive behavior (Afifi, Afifi, Merrill, Denes, & Davis, 2013). Perhaps the clearest theoretical connection between listening and enacted support lies in the most commonly used verbal operationalization of enacted support, person centeredness. Person centeredness includes “a repertoire of behavioral strategies and tactics” (Burleson, 2003, p. 580) that are considered behavioral manifestations of “good” listening. These include the use of open-ended questions, paraphrasing, backchanneling, and reflecting feelings, all of which are direct, observable manifestations of “active listening” (Bodie, Vickery, Cannava, & Jones, in press; Elliott, Bohart, Watson, & Greenberg, 2011; Jones, 2011; Weger, Castle, & Emmett, 2010).

### Different Decoder Perspectives of Enacted Support

Lay persons generally (i.e., without regard to relational cues) agree on the behaviors that constitute “good” listening (e.g., Bodie et al., 2012). In the context of research, however, decoders are typically asked to judge listening and immediacy cues as indicators for enacted support from different perspectives (e.g., while observing a conversation versus after being involved in a conversation). These decoders may have different (or no) stakes in the support process. Consequently, they may judge the enactment of support differently (or similarly), which may have consequences for the outcomes we attribute to supportive communication. Indeed, several researchers have pointed to the importance of examining how various decoder perspectives influence message evaluations and effects (e.g., Feeley, 2002; Street, 1985). Particularly relevant for enacted support is the meta-analysis by High and Dillard (2012), which found that the relationship between manipulations and evaluations of verbal person centeredness (VPC) is moderated by research design. The average effect of VPC on evaluation measures in hypothetical scenario studies was .67, whereas the average effect in confederate studies was .37 (see Afifi et al., 2013 for a similar finding). In other words, third-party judges seem to make greater distinctions among different types of enacted support than individuals within the conversations in which these behaviors are enacted.

Because it is relevant to our claim, High and Dillard’s (2012) confederate versus hypothetical scenario distinction needs to be made more perspicuously. Work using hypothetical vignettes has either asked participants to imagine they are in a stressful situation and are the targets of support attempts or to read a hypothetical conversation and evaluate the support manipulated in these conversations. The former method places participants in the role of participant/recipient of enacted support, whereas the latter places them in the role of observer (Cappella & Street, 1989). These two perspectives are not isomorphic, and thus hypothetical scenario studies are not homogeneous with respect to enacted support judgments. As for interaction-based research, the most common method is to manipulate the behavior of a confederate recruited to provide support based on a researcher defined set of supportive behaviors. Trained raters or untrained third-party judges are then used to provide assessments of supportiveness either as a manipulation check (Afifi et al., 2013; Jones & Guerrero, 2001) or to render estimates of *accuracy*, defined as “the extent to which a person’s perceptions of support received ... correspond with the observable supportive behaviors the partner enacts,” and *bias*, defined as “when an individual perceives more or less supportiveness than is actually occurring” (Priem, Solomon, & Steuber, 2009, p. 532). Here again, we argue that the two judgments generated from trained or untrained observers are not isomorphic and thus interaction-based studies also are not homogeneous.

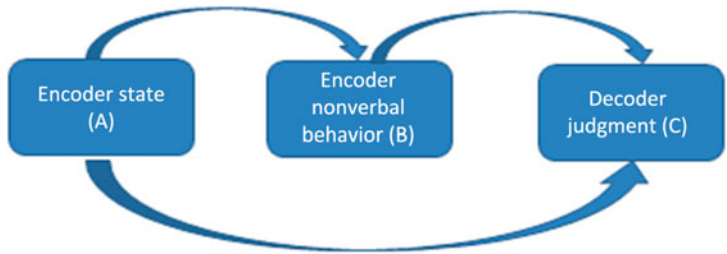
Taken together, it seems to us that scholars interested in enacted support have ignored the issue of who gets to decide what counts as quality-enacted support. Indeed, past work has not focused on the degree of convergence among various perspectives that can be used to judge enacted support but has assumed a preference

that certain perspectives represent coherent or stable facets of enacted support. An essential consideration for scholars of supportive communication is this: What are best practices for measuring evaluations of enacted support? We note that although some outcomes of supportive communication are rather unproblematic in this regard (e.g., emotional improvement is best measured from the perspective of the distressed individual), what is perceived to have been enacted can be judged from a variety of perspectives, assessed at behavioral micro- or macrolevels, from the vantage point of one specific interaction or a set of interactions, or in reference to an individual in general. None of these perspectives or levels is necessarily more valid or accurate than the other; rather, each is more or less useful depending on the specific focus of a particular study.

### Focus of the Current Studies

Ambiguity over whether decoder evaluations (e.g., third-party observers, interlocutors, and trained versus untrained coders) are in fact isomorphic points to crucial concerns about the construct validity of enacted support. Carlson and Herdman (2011) stated that “a critical issue for researchers is judging when the convergent validity of proxies is strong enough to permit one measure to substitute for another in a research design” (p. 20). In the case of work on enacted support, who gets to evaluate that support and when can another individual (or similar set of individuals) act as a substitute for that perspective? We argue that the level of methodological specificity in assessing enacted support is lacking and posit that different decoder (i.e., data) sources for enacted support evaluations render different values for behavioral immediacy and listening indicators. In the two studies, we compare enacted support evaluations of a representative sample of the total population of possible assessment choices for stranger dyads using two observational measures of enacted support: Immediacy and listening behaviors. Each perspective represents an assessment that reflects our conceptual definition of enacted support, namely, what people do when called on to provide comfort to others. We present evidence that different decoding sources represent different perspectives of enacted support.

To help organize our thinking, we draw theoretical insight from Rosenthal's (1987) model of judgment studies. A judgment study involves “one or more encoders characterized by one or more attributes (A) observed by one or more decoders who make one or more judgments (C) about the encoders on the basis of selectively presented behavior (B)” (Rosenthal, 1987, p. 4). Using the model depicted in Figure 1 (see Rosenthal, 2005, p. 200), the typical design of studies that examine enacted support is to manipulate one or more supportive behaviors (B) produced by one or more actual or imagined help providers (A) in the service of assessing the evaluation of that behavior (C). Research exploring the A–B (encoder attributes–observed behavior) link has detected individual variability in supportive message sophistication related to empathic motivation and social cognitive abilities (e.g., Burleson, 1983). Research exploring the B–C (observed behavior–decoder judgment) link has shown that message evaluations are a linear function of various message qualities such as



**Figure 1** A simple model of judgment studies (from Rosenthal, 2005, p. 200).

VPC (High & Dillard, 2012). Finally, research exploring the A–C (encoder state–observed behavior) link has documented that certain characteristics of people influence support judgments (e.g., similarity and conscientiousness; Lakey, Ross, Butler, & Bentley, 1996, Study 2); other research has suggested that person and message characteristics interact with these message evaluations (i.e., the A–B–C chain; e.g., Bodie & Bursleson, 2008; Uno, Uchino, & Smith, 2002). Of primary concern for our studies is the evaluation component of Rosenthal’s model and the degree to which different decoder evaluations of enacted support are substitutable (Study 1) and internally consistent (Study 2).

In Study 1, we generate a multitrait–multimethod (MTMM) matrix including correlations between various ways to operationalize the evaluation of immediacy and listening behaviors (Rosenthal’s “C”). Patterns of correlations are inspected for evidence of convergent and divergent validity. For Study 2, we predict that if different perspectives are represented with different approaches to operationalization, then ratings within perspective should exhibit a high degree of reliability. Interestingly, whereas past work with confederate support providers (e.g., Afifi et al., 2013) has appropriately stressed the internal consistency of trained raters, the similarity among other third-party observers has not been empirically scrutinized, even though the third-party observer perspective is the most commonly used perspective to judge enacted support quality. Study 2 explores the *effective reliability* (i.e., the consistency of judgments within a set of judgments, or mean judgments) for multiple reports from untrained raters taking a third-party evaluation perspective.

### Study 1: A MTMM Matrix of Active Listening and NI

Goldsmith (2004) noted that “there is likely to be individual variability (and even some idiosyncrasy) in how an action is interpreted and evaluated. However, meaning is not completely subjective, nor is it simply correspondence between two subjective individuals” (p. 34). In Study 1, we compare and contrast three perspectives from which one can judge enacted support. To test the convergence of measurement techniques for immediacy and listening behaviors, we used Campbell and Fiske’s (1959) MTMM approach and utilized three data sources, namely, self, partner, and observational reports.

Using MTMM, this study explores the extent to which the measures commonly used to operationalize active listening are construct-valid measures of enacted support (see Carlson & Herdman, 2011). We assess four types of correlations, each providing unique evidence for (or against) construct validity. The first type of correlation generated by an MTMM analysis is the *monotrait–monomethod* correlation. This correlation is synonymous with the reliability coefficient of a scale. The second type of correlation is the *monotrait–heteromethod* correlation, which represents the association between different measurement methodologies used to measure the same construct. These correlations provide direct evidence of convergent validity. Often referred to as validity coefficients, these values should, at a minimum, be statistically different from zero and sufficiently larger than heterotrait correlations to demonstrate evidence for divergent or for discriminant validity. The MTMM produces two heterotrait correlations: *Heterotrait–heteromethod* correlations are associations between different measurement methodologies used to measure different constructs, and *heterotrait–monomethod* correlations are associations between different constructs measured by a common methodology. When heterotrait–monomethod correlations are larger than the heterotrait–heteromethod correlations, measurement bias is a concern.

### Method

*Participants and procedures.* The data-set for this study is a subset of a larger pool of 171 dyadic interactions.<sup>1</sup> All 206 participants included in this study were undergraduate students enrolled in Communication Studies courses at Louisiana State University (LSU) A&M. The sample consisted of two subgroups: One person (the “discloser,”  $n = 103$ ) who disclosed an upsetting event to another (the “listener,”  $n = 103$ ), who was asked to attend to the problem discloser. Disclosers and listeners were on average 20 years of age ( $SD = 4.88$ ) and primarily female (61%). Although most (73.5%) self-reported as Caucasian, others self-reported as African-American/Black (16.2%) and Asian (3.9%); other categories constituted 1% or less of the data-set.

Students signed up for the study through an online bulletin board, were sent a link to a secure URL, and then completed a set of self-report scales including those for general enactments of listening and immediacy behaviors. Once surveys were completed, two participants were invited to a multiroom laboratory. After verifying that they were unacquainted, participants (1) provided informed consent and (2) were randomly assigned the role of a discloser or of a listener. Support is frequently enacted in established relationships, but given our objective to examine methodological differences in how enacted support is assessed, we elected to use stranger dyads because (1) the use of stranger data is an empirical reality in much of the social support literature and (2) stranger dyadic data serve as one valid decoder perspective that is frequently represented in research on enacted support (e.g., Burleson & Samter, 1985; High & Dillard, 2012; Jones & Wirtz, 2006). Researchers have justified the use of zero-history relationships because Institutional Review Board (IRB) have

not permitted the use of established relationships, or because these data remove commonly investigated “relational characteristics (like commitment or satisfaction) that may bias perceptions” (Priem et al., 2009, p. 532). We are certainly not arguing that stranger dyadic data be preferred in all cases; rather, our choice ultimately permits us to compare our results to those of past research.

Participants were separated so that disclosers could complete information about the emotionally distressful event they would be talking about in the subsequent 5-minute conversations; listeners completed measures not germane to this article. We asked disclosers to identify daily hassles as potential conversational topics for three reasons: (1) Daily hassles constitute the majority of our stressful live events, and (2) they are important predictors of distress and other mental and physical ailments (Almeida, 2005). Finally, (3) the vast amount of research in Communication Studies focuses on enacted support that is provided in response to minor daily hassles (see Burleson & Goldsmith, 1998). Events were everyday stressors relevant to college students, including academic stressors (e.g., failing a test), relationship problems (e.g., roommate spats), family problems (e.g., parents fighting), and health-related issues (e.g., a recent surgery). Once the event was identified, participants were reunited and given instructions about the conversation.

The dyad was given one minute to get acquainted, after which a video and audio recording was made of the 5-minute supportive conversation for later analysis. The camera was positioned in a corner of the room to capture a profile of both participants who sat across from one another in chairs. After five minutes, a research assistant knocked on the wall to indicate the end of the conversation, then entered the room, and separated the participants to gather postconversation data. Per IRB regulations, each participant was debriefed and offered information about the LSU A&M Mental Health Service Center.

### *Analysis Plan and Measures*

All measures were tested with confirmatory factor analytic (CFA) techniques using AMOS 19.0. Model fit was determined by examining the comparative fit index (CFI) and the standardized root mean square residual (SRMR) value (Kline, 2005).<sup>2</sup> Unless otherwise noted, standardized residual covariances for all models were below 2.58 in absolute value, and error terms were left uncorrelated. Standardized loadings were deemed acceptable if  $\lambda > .50$ .

*Self-report measures.* For self-reported enacted listening, we used a version of the Active-Empathic Listening Scale (AELS-SR), which contains 11 statements (e.g., “I assure others that I will remember what they say”) and uses 7-point scaling (1 = *Never or almost never true*; 7 = *Always or almost always true*; 4 = *Occasionally true*) (Bodie, 2011, Study 1). This constitutes a measure of general tendencies to enact supportive listening behaviors. Each item tends to load on one of three first-order factors (i.e., sensing, processing, and responding), which in turn converge on a second-order factor (Bodie, 2011, Study 1; Bodie, Gearhart, Denham, & Vickery,

2013; Drollinger, Comer, & Warrington, 2006; Gearhart & Bodie, 2011). For the current data, out of the 1,133 total responses, there were 4 missing values (0.4%) which were replaced with the mean of surrounding items (Tabachnick & Fidell, 2007). Each item revealed approximately normal kurtosis ( $M = -0.03$ ,  $SE = 0.47$ ) and slight skew ( $M = -0.59$ ,  $SE = 0.24$ ); indeed, the mean for each item was above the midpoint ( $M = 5.32$ ,  $SD = 1.23$ ), and the mode for 10 out of 11 items was 6. The overall scale mode was 5. The second-order model fit the data well (see Table 1). We therefore treated active-empathic listening (AEL) as one construct.

The Nonverbal Immediacy Scale (NIS) developed by Richmond, McCroskey, and Johnson (2003) contains 26-items (e.g., “I have relaxed body posture when I talk to people”) that are rated on 5-point scales (1 = *Never*; 5 = *Very Often*). Similar to the AELS, this constitutes a measure of general tendencies to enact immediacy behaviors. The NIS is proposed to be unidimensional but possesses only a small validity portfolio featuring empirical evidence that is generated from student–teacher interactions (e.g., McCroskey, Richmond, & Bennett, 2006). Four out of 2,575 total responses were missing and replaced with the mean of surrounding items. No item showed problematic skew ( $M = 0.24$ ,  $SE = 0.24$ ) or kurtosis ( $M = 0.39$ ,  $SE = 0.47$ ), and the mean for each item was close to the midpoint ( $M = 2.86$ ,  $SD = 0.85$ ).

Our data failed to conform to a one-factor model (see Table 1). A close inspection of the correlation matrix suggested that the 26 items might be better represented by 7 first-order factors. Although this model was an improvement, it still did not produce acceptable fit indices. A close inspection of the factor structure showed that indicators loaded sufficiently high enough onto respective factors (i.e.,  $\lambda_s > .50$ ).

**Table 1** Model fit indices for all scales used, Studies 1 and 2.

Scale	$\chi^2$	<i>df</i>	SRMR	CFI	RMSEA	Low <sup>a</sup>	High <sup>a</sup>
<i>Study 1</i>							
AELS-SR	50.32*	41	.05	.97	.05	.00	.09
NIS-SR, unidimensional	827.95	275	.13	.47	.14	.13	.15
NIS-SR, first-order correlated	430.02	231	.09	.80	.09	.08	.11
NIS-SR, second-order	364.73	202	.09	.82	.08	.07	.10
NIS-SR, parceled	32.05	14	.07	.88	.11	.06	.16
NIS-SR, parceled (modified)	21.91*	13	.06	.94	.08	.00	.14
AELS-PR	98.62	41	.07	.89	.12	.09	.15
NIS-PR	24.85	9	.04	.93	.11	.04	.17
<i>Study 2</i>							
AELS	103.51	41	.03	.97	.06	.04	.07
NIS	248.92	35	.04	.92	.12	.11	.14

Note: All model Chi-square statistics were significant at  $p < .05$  unless marked with an asterisk;  $*p > .05$ .

<sup>a</sup>Low and high values are the lower and upper bound estimates of the 90% confidence interval for RMSEA, respectively.

Thus, all NIS items seemed to be valid indicators. In addition, modification indices did not produce substantive suggestions for fit improvement, and no standardized residual covariances were above 2.58 in absolute value. For these reasons, we chose to parcel NIS items based on seven sets of nonverbal behaviors (hands/gestures, touch, voice, eye behavior, body position, closeness, and facial expressions). Each parcel contained between three and five items (on the advantages of parceling see Little, Cunningham, Shahar, & Widaman, 2002).

The unidimensional seven-parcel model was suggestive of better fit. Although the CFI is slightly lower than desired, an inspection of the standardized residual covariance matrix did not return any problematic items; instead, modification indices suggested correlating the error terms associated with touch and closeness. A close inspection of items in these two categories warranted such a correlation: Consider the touch item “I move away from others when they touch me while we are talking [reversed coded]” and the closeness item “I sit close or stand close to people while talking with them.” This modification did improve relative structural fit (see Table 1). Taken together, our statistical solution seems viable for present purposes, though these data warrant caution; future psychometric research is necessary before the NIS can be used with impunity.

*Partner-report measures.* After the 5-minute conversation, disclosers were asked to respond to several items that provided data on perceptions of the conversation. There were no missing data for these measures.

The partner report version of the AELS (AELS-PR) was modified so that the 11 statements began with “My conversational partner.” One item had one missing value, which was replaced with the value of surrounding items. No item showed problematic skew ( $M = -0.82$ ,  $SE = 0.24$ ) or kurtosis ( $M = 0.50$ ,  $SE = 0.47$ ), and the mean for each item was close to 5; the scale had the following characteristics:  $M = 4.98$ ,  $SD = 1.11$ ,  $Mdn = 5.18$ ,  $Mode = 4$ . The confirmatory model was adequate (see Table 1); one standardized loading was slightly below .50 ( $\lambda = 0.47$ ).

Six items that reflected the conceptual nature of immediacy constituted the partner-report of NI (NIS-PR) ( $M = 5.08$ ,  $SD = 0.98$ ;  $Mdn = 4.17$ ,  $Mode = 5.17$ ). Items were scaled along 7-points (1 = *Never or almost never true*; 7 = *Always or almost always true*). Past work has indicated that immediacy perceptions are a result of a set of nonverbal cues that are processed as a *gestalt* (P.A. Andersen & Andersen, 2005). Thus we chose to tap fundamental features of immediacy rather than all 26 items: body language (e.g., “My conversational partner used inviting body language”), nonverbal expression (e.g., “My conversational partner expressed understanding nonverbally”), and body tenseness (e.g., “My conversational partner felt very tense talking to me”). None of these items showed problematic skew or kurtosis. Table 1 displays model fit information.

*Behavioral coding.* The videotaped interactions were assessed after a training that consisted of (1) having a theoretical discussion of the relevant construct (i.e., AEL and NVI); (2) discussing and visually demonstrating the level of each nonverbal cue;

(3) rating videotaped interactions; and (4) discussing and adjusting differences in ratings. When training the raters, the end points of the scales were conceptualized to fit the specific cue. For example, for the immediacy coding, a high level of eye contact was conceptualized as exhibiting eye contact 80% or more of the time; a low level of eye contact was conceptualized as exhibiting eye contact 20% of the time or less. Raters assessed the cues twice, once after the first half of the conversation and then again after the second half of the conversation. All raters were allowed to stop, rewind, and fast-forward the videos as much as they needed. Inter-item correlations across Times 1 and 2 were quite high ( $r_s > .65$ ) and homogeneous across all cues, suggesting that there were only minor variations in the first and second half of the conversations. Cues for both Times 1 and 2 were consequently collapsed within the assessment rubrics.

Listening behaviors were coded using five items from the Active Listening Observation Scale (ALOS; Fassaert, van Dulmen, Schellevis, & Bensing, 2007); each is assessed from 0 (*Never*) to 4 (*Always*). Five research assistants blind to the study's purpose were trained by the first and third authors to assess the following: (1) Uses exploring questions; (2) expresses understanding verbally; (3) reflects feelings; and (4) paraphrases information. Coders also made a global rating of active listening that reflected their assessment of the four composite behaviors. Prior to training, the first and third authors rated five videotaped conversations and modified the rubric in an iterative fashion. Raters were trained only after establishing agreement between the experts. Inter-coder reliability (Krippendorff's alpha) was then assessed among the five raters using 12 dyads (see Table 2).

NI cues were rated with a modified version of Andersen, Andersen, and Jensen's (1979) NIS-Coded. Two research assistants blind to the study's purpose were trained by the second author and evaluated nine immediacy cues (e.g., "orient her/his body toward the other person," "smile when it's appropriate?") and one global immediacy evaluation. The immediacy cue stem read "To what extent is/does the person(s)...," and raters were directed to rate the immediacy cues of the listener along 7-points (1 = *Not at all*; 7 = *Very much*). Table 2 reports reliability.

## Results

The monotrait–monomethod diagonal of Table 2 contains standardized Cronbach's alpha values for self and partner scales;<sup>3</sup> values for the coded listening and immediacy behaviors are Krippendorff's alpha and the two-way random intra-class correlation coefficient, respectively. The remainder of the matrix contains Pearson's product-moment correlation coefficients, organized by trait and method. The values above the diagonal are disattenuated using Spearman's (1904) original formula, and the values below the diagonal represent values with measurement error.

**Table 2** Multitrait–multimethod correlation matrix.

	Active-empathic listening			Nonverbal immediacy		
	Self-report	Partner-report	Coded behavior	Self-report	Partner-report	Coded behavior
AEL self-report	<b>.87</b>	−.14	−.07	.50	.02	.13
AEL partner-report	−.12	<b>.90</b>	.35	−.01	.75	.09
AEL behavior	−.06	.30	<b>.80</b>	.00	.28	.42
NIS self-report	.43	−.01	.00	<b>.84</b>	.11	.19
NIS partner-report	.02	.63	.22	.09	<b>.78</b>	.16
NI behavior	.11	.08	.34	.16	.13	<b>.81</b>

Note: AEL = active-empathic listening; NIS = Nonverbal Immediacy Scale; NI = nonverbal immediacy. Correlations above the diagonal are corrected for measurement error, whereas those below the diagonal are not. Color coding key:

Monotrait–monomethod correlation	Monotrait–heteromethod correlation	Heterotrait–monomethod correlation	Heterotrait–heteromethod correlation
----------------------------------	------------------------------------	------------------------------------	--------------------------------------

*Monotrait–monomethod.* Based on conventional standards, all measures achieved an adequate level of internal consistency; the only value below .80 was the partner report of immediacy. In general, statistics are consistent with prior reports (Bodie, 2011; Bodie, Gearhart, et al., 2013; Drollinger et al., 2006; Gearhart & Bodie, 2011; Jones & Wirtz, 2007; McCroskey et al., 2006; Richmond et al., 2003). Nevertheless, all measures featured internal inconsistencies. Therefore, we report all other correlations corrected for measurement error. Consistent conclusions result from both sets of correlations, making differential patterns of measurement error a nonviable explanation for results.

*Monotrait–heteromethod.* The monotrait–heteromethod block of correlations provides evidence against convergent validity; scores on valid alternative measures of the same construct should be a result of the same construct. AELS values ranged from −.12 to .30 (−.14 to .35, corrected). NIS values ranged from .09 to .16 (.11 to .19, corrected). The highest correlations within traits were between (1) the AELS-PR and AELS-Coded and (2) the NIS-SR and NIS-Coded.

*Heterotrait–heteromethod.* The heterotrait–heteromethod correlations in Table 2 provide evidence for divergent validity. In particular, of the six heterotrait–heteromethod correlations, five were within sampling error of zero. The only correlation that was statistically different from zero was AELS-Coded/NIS-PR, which was small in magnitude ( $r = .22$ ).

*Heterotrait–monomethod.* The final set of correlations provides evidence against discriminant validity, since the association between different constructs measured by common measurement methodologies were consistently larger (uncorrected  $r_{ave} =$

.47; corrected  $r_{\text{ave}} = .56$ ) than the monotrait–heteromethod correlations (uncorrected  $r_{\text{ave}} = .08$ ; corrected  $r_{\text{ave}} = .10$ ). These results suggest a high degree of common method variance. The AELS-PR and NIS-PR generated the largest method effects, and coded behaviors produced the smallest.

### *Discussion of Study 1*

Study 1 examined the degree to which two factors important to conceptualizations of enacted support—listening and immediacy—are evaluated similarly (or differently) as a function of three data sources (self, partner, and observers). Using a MTMM approach, Study 1 explored the extent to which active listening and immediacy are construct valid measures of the verbal and nonverbal behavioral manifestations of enacted support.

In general, the AELS and NIS measures each seem to be internally consistent and produce reliable proxies for the constructs under operationalization. That is, the indicators tapping immediacy and those tapping active listening seem to reliably measure and consequently represent their respective constructs. The second highest coefficients, however, were along the heterotrait–monomethod diagonal. We detected correlations between (1) self-reported AEL and NIS ( $r = .43/.50$ ); (2) partner-reported AEL and NIS ( $r = .63/.75$ ); and, (3) AEL and NIS behaviors ( $r = .34/.42$ ; see Table 2). These results suggest that studies investigating validity issues with single-method measures of enacted support may be reporting correlations inflated by common method variance (CMV); indeed, many past correlations may be spurious (Podsakoff & Organ, 1986) and may have to be disattenuated to account for this bias.

CMV is defined as systematic error variance shared among variables that is introduced as a function of the same method and/or source (Richardson, Simmering, & Sturman, 2009). For example, if a researcher measures both predictor and outcome variables with a common measure or method (e.g., observer judgments and self-reports), then the estimated coefficient may be much higher than if different measures or methods of assessment are used. No matter its degree, the consequences of CMV are considerable. In egregious cases, estimates may be within interpretable bounds, yet be entirely a function of shared methods across constructs that inflate Type I error, a problem that Bank, Dishion, Skinner, and Patterson (1990) referred to as the “glop” problem. The question then becomes how one can minimize (rather than completely eradicate) this systematic error variance. The interested reader is referred to Richardson et al. (2009) to determine appropriate solutions to CMV. At the very least, the degree to which CMV changes the conclusions regarding past work should be addressed (e.g., the degree to which relations among enacted support and health outcomes change as a function of decoder perspective used to operationalize enacted support).

The remaining coefficients seem to suggest that different measurement techniques produce different perspectives on enacted support. Self-reports were largely unrelated to evaluations made by conversational partners or trained observers (see monotrait–heteromethod correlations). The lack of correspondence between self-reports and

coded behaviors is consistent with research on instructional communication, suggesting that student reports of instructor immediacy only minimally correlate with actual teacher behavior (i.e., their actual enactment of supportive behaviors; Smythe & Hess, 2005), as well as with work on verbal aggressiveness and assertiveness (Kotowski, Levine, Baker, & Bolt, 2009; Levine, Kotowski, Beatty, & Van Kelegom, 2012). Likewise, partner-reports of the NIS were not statistically associated with the coded behaviors, whereas partner-reported and coded AEL were only moderately correlated ( $r = .30/.35$ ). Thus, conversational partners may rely on different cues or may process similar cues in different ways compared to trained coders.

One possible explanation for the lack of correspondence between self and partner/coded evaluations is that these sets of evaluations are contextually heterogeneous. That is, whereas self-reports were assessed with respect to what a supportive listener does in general (or thinks he or she does), partner- and coded-evaluations were assessed after a specific conversational instance of enacted support. It is instructive, however, that the monotrait–heteromethod block for the NIS produced correlations of similar magnitude regardless of this contextual difference. The AELS fared only slightly better in this regard. In addition, the method bias already mentioned seems to suggest that CMV is the most potent influence on our results. Thus, contextual specificity may not fully explain the pattern of results in Study 1. Instead, we suspect that the different operational definitions of enacted support represent different perspectives from which one judges these behaviors, providing empirical backing for Goldsmith's (2004) charge to take into account "multiple perspectives of evaluation" when studying supportive communication (p. 161).

We present a second study that explored the possibility that our results might be profitably explained by the following hypothesis: If different operational definitions of enacted support constitute different perspectives, then evaluations within perspective should exhibit a high degree of what Rosenthal calls *effective reliability* (ER)—the *consistency of judgments* within a set of judgments, or *mean judgments*. We focus particularly on third-party evaluators, namely, untrained raters, because this evaluation type is frequently used in the assessment of supportive message quality. By assessing ER within this perspective we are ultimately investigating the substitutability criterion, or the degree to which another individual (or similar set of individuals) can act as a substitute within this perspective.

## Study 2: The Effective Reliability of Enacted Support Judgments

An assumption underlying much interaction-based research is that, even in specific interactions, "perceptual biases" can influence the evaluation of "objective" behaviors (Collins & Feeney, 2000; Priem et al., 2009). For example, Priem et al. (2009) found that self and partners' (husbands' and wives') evaluations of supportiveness after a brief conversation were a function not only of enacted support in the conversation (what Priem et al. call "accuracy" represented by third-party and untrained observer ratings) but also of partners' relational satisfaction (sentiment override) and rating of

partners' own supportiveness during the conversation (projection). We are struck by the use of the researchers' operational definition of "accurate" enacted support. Similar to past research (Kenny & Acitelli, 2001), Priem et al. measured "accuracy of relational perceptions" (p. 539) by asking four third-party observers unacquainted with the couple to (1) review a broad definition of supportiveness, and then (2) judge husbands' and wives' supportiveness on a single-item measure. The four ratings were subsequently averaged to form a single accuracy measure.

But why should we view these judgments as accurate representations of enacted support evaluations? Third-party judgments consist of a wide range of evaluators who vary in training and relational history.<sup>4</sup> In short the assumption of much of the research on enacted support is that different decoding sources do in fact represent unique perspectives on support, which are then reflected directly in empirical data. Currently, the best available evidence that decoding sources provide different kinds of information comes from the few studies that actually report evaluation data from more than one type of judge (Afifi et al., 2013; High & Dillard, 2012). These findings provide circumspect yet compulsory evidence for the claim that different sources (e.g., listeners, disclosers, friends, and observers) represent different perspectives that, in turn, generate qualitatively different types of information about enacted support evaluations. A common argument for judgment studies is that certain classes of evaluators (e.g., different types of "friends"; Weiner & Hannum, 2013) are unique in their evaluations of targets. In Rosenthal's terms, we are interested in the effective reliability of a set of judges who rate a single person's enacted support. We once more relied on listening and immediacy cues to represent enacted support evaluations. Because the third-party perspective is the most commonly used perspective in supportive communication scholarship, we focus on that perspective here. In particular, we test whether untrained third-party observers reach similar conclusions of support enacted in a specific conversation. As in the design used by Priem et al. (2009), the judges were unacquainted with the dyad in the conversation. We used untrained judges because trained raters are usually highly consistent; indeed, consistency is a prerequisite for using trained rater data as the proxy for quality enacted support. Consistency (reliability) can, however, often outweigh thoughtful consideration of exactly what behaviors are constitutive of a supportive response. If untrained judges represent a unique perspective for enacted support evaluations, we should find a high degree of consistency among these third-party observers when they evaluate a supportive conversation.

### *Method*

*Participants.* Undergraduate students ( $N = 383$ ) from LSU A&M ( $n = 305$ ; 192 women and 109 men; four participants did not report sex) and the University of Minnesota, Twin Cities ( $n = 78$ ; 48 women and 29 men; one participant did not report sex) constituted the sample for Study 2. Participants either completed the study as part of a research requirement or received a modest amount of extra credit. The mean age of the participants from LSU A&M was 22.1 years ( $SD = 4.87$ ;

Range = 18–49); the average age of the participants from the University of Minnesota, Twin Cities, was comparable ( $M = 20.5$ ;  $SD = 3.01$ ; Range = 18–48). Most participants ( $n = 300$ ; 78.3%) self-reported as Caucasian.

*Procedures.* Identical procedures, each approved by the appropriate IRB, were used at both institutions. A research assistant unaffiliated with the project and blind to the hypotheses greeted participants and then asked them to view one randomly selected 5-minute recorded conversation on an individual computer, using headphones to control the audio output. After they watched the conversation, participants completed a computer-based survey.

*Stimulus conversations.* The conversations that participants evaluated were generated for an unrelated project.<sup>5</sup> Briefly, each of the originally collected conversations ( $N = 264$ ) featured a participant who was seemingly randomly assigned to discuss an emotionally upsetting event with the confederate, who was randomly assigned to exhibit more or less verbally person-centered and nonverbally immediate support on three levels (high, moderate, and low person-centeredness and immediacy). All confederates were trained to enact each of nine person-centered and immediate combinations. To ensure a reasonably representative stratified sample, we randomly selected eight conversations from each condition, resulting in a sample of 72 conversations from the original 264; confederate sex, person-centeredness, and immediacy were the three strata.<sup>6</sup>

*AELS.* The AELS was modified by instructing participants to assess “the person seated to the right” (always the confederate in the video) with respect to the 11 items measuring listening behaviors (7-point scaling). Fewer than 1% of cases were missing, and each missing point was replaced with the mean of surrounding items. Items were approximately normally distributed (skewness =  $-.317$ ,  $SE = .125$ ; kurtosis =  $-.764$ ,  $SE = .249$ ), and most item means hovered slightly above the midpoint ( $M = 4.57$ ,  $SD = 1.79$ ). The second-order model fit the data well (see Table 1), and the scale was internally consistent ( $\alpha = .93$ ).

*NIS.* Because we had operational difficulties with the 26-item NIS (the self-report version) in Study 1, we opted to use the scale employed for the behavioral coding in Study 1 developed by Andersen et al. (1979). Missing data constituted 1.2% of the data-set and were handled by mean imputation. Items were approximately normally distributed (skewness =  $.253$ ,  $SE = .125$ ; kurtosis =  $-.920$ ,  $SE = .250$ ), and most item means were slightly below the midpoint ( $M = 2.73$ ,  $SD = 1.31$ ). The data conformed to the measurement model (see Table 1), and the scale was internally consistent ( $\alpha = .93$ ).

## Results

Rosenthal (1987, 2005) suggested that *effective reliability* requires an assessment of both absolute agreement and consistency. To assess absolute agreement, we

computed difference scores representing the absolute discrepancy between any two judges for a single target. These difference scores represent the degree to which respondents are utilizing the same end points of a scale or are discrepant in their evaluations. To assess consistency, we computed an intraclass correlation coefficient (ICC) which represents the average consistency of all judges for one single participant. Because we had various judges rating each target, we used a one-way random analysis of variance (ANOVA) model, and we report the average measure of consistency (i.e., the reliability of the mean of the raters).

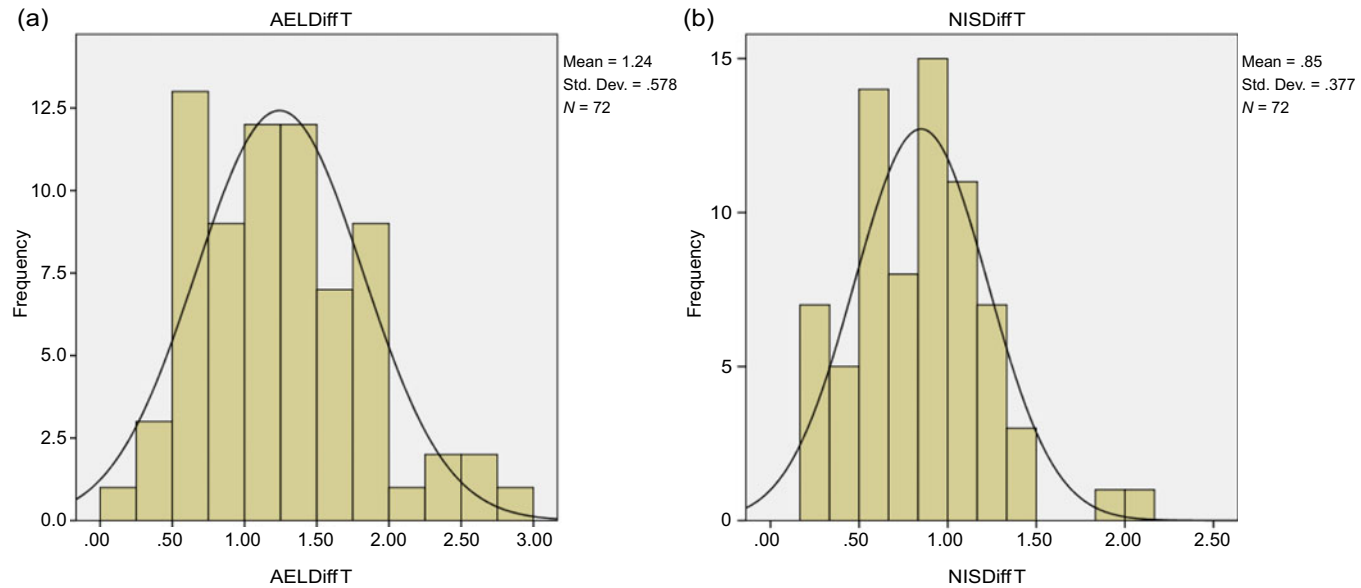
For the AELS, the average difference score was approximately one-and-a-quarter scale points ( $M = 1.24$ ,  $SD = .58$ ,  $Mdn = 1.23$ ,  $Mode = 1.24$ ) suggesting that raters were utilizing the same scale end points when assessing a target. Although not all targets had low difference scores (Range = 0–5.55), on average, scores clustered between half a scale point and less than two scale points (see Figure 2a). Similar results were found for the NIS ( $M = .85$ ,  $SD = .38$ ,  $Mdn = .86$ ,  $Mode = .60$ ). The minimum difference score was zero, while the maximum was 3.10, and scores tended to cluster at the low end (see Figure 2b). As a measure of the mean reliability of coders, both the AELS ( $r_i = .75$ ; 95%,  $CI = .62, .85$ ) and NIS ( $r_i = .86$ ; 95%,  $CI = .79, .92$ ) showed a high degree of consistency with tight confidence intervals.

### General Discussion

Social support is a multifaceted construct that refers to several support phenomena that contribute differently to health and human functioning. One of these facets is *enacted support*. Social psychologists frequently conceptualize enacted support as what is provided or received (Gable et al., 2012; Lakey et al., 2009), whereas enacted support in the communication sciences consists exclusively of verbal and nonverbal behaviors that are exchanged between two or more people in a supportive interaction when at least one person is trying to cope with distress (Goldsmith, 2004). Enacted support is central to social support. Although this point was acknowledged in early psychosomatic research on social support (Cassel, 1976), it was abandoned quickly in favor of examining facets that were more perceptually based, yet easier to access (see MacGeorge et al., 2011).

Of course, enacted support is perceived and evaluated, but who gets to do the evaluating and what are best practices for measuring evaluations of enacted support? Our studies are concerned with the construct validity of enacted support and tackle a methodological quandary, namely, the effects of different source evaluations on enacted support. Assessing the differences and similarities of a person's assessments of enacted support strikes us as an important question: If we attribute health benefits (or whatever the assessment outcomes may be) to enacted support, then understanding how and in what ways different data sources influence enacted support evaluations seems to be of vital importance.

We explored two construct validity issues. First, we examined source accuracy and bias in enacted support evaluations (Study 1). We used a multitrait-multimethod (MTMM) approach to examine differences in information provided by three



**Figure 2** (a) Histogram of average difference scores, AELS, Study 2; (b) Histogram of average difference scores, NIS, Study 2.

different person sources or facets (support recipients, providers, and strangers). Second, we examined the consistency of third-party evaluations of enacted support. We focused on untrained observers who were asked to evaluate the immediacy and listening behaviors of an unacquainted support provider in a prerecorded conversation. In both studies, we operationalized enacted support as consisting of behavioral manifestations best captured with listening and immediacy cues. Below, we first provide a summary of our findings and then move on to larger issues to which our studies speak.

### *Summary of Findings: Source Orientations Generate Different Judgments*

We found that the listening and immediacy measures (i.e., AELS and NIS) were internally consistent, yet featured relatively small relationships among source orientations. That is, the immediacy and listening judgments made by listeners (support providers), disclosers (support recipients), and strangers (raters) were largely unrelated with one another (Study 1). Even though the self-report measure was completed with respect to one's general tendencies and the partner-report and rating measures with respect to a specific conversation, we observed correlations of similar magnitude (see Table 2). Whereas noncorrespondence between self-reported general tendencies and perceptions of behavioral enactments within a specific conversation is not unprecedented (and is perhaps to be expected), one would be inclined to assume that people—whether observing participants or raters—would converge on judgments as concrete as the number of questions that a listener has asked in a conversation (an AELS indicator) or the number of touches a listener has used (an NIS item). That was not the case, however.

Notably, our data reflect aggregate measures only. Perhaps only *some* of the concrete behavioral immediacy indicators that make up the NIS (e.g., gestures), as well as those cues that tap active listening (e.g., summarizes points of agreement), are evaluated similarly. For example, Smythe and Hess (2005) found an average correlation of .15 for a set of seven nonverbal behaviors, but some retrospectively reported behaviors were more related to enacted behaviors than others: The item “shows a lot of facial expressiveness” generated a correlation of .55, whereas “smiles at the class” featured a correlation of .45. We investigated the possibility that our results might differ if individual verbal and nonverbal cues were analyzed separately. We computed correlations between individual self-reports and the behavioral ratings from Study 1 and featured these results in Table 3. Notably, all correlations are within sampling error of zero. In other words, it does not seem as though sources are more or less accurately reporting on specific behaviors. On the contrary, our results seem to support the general notion that self-judgments of *behavior* do not predict what people are actually likely to *do* in a conversation, at least when people are talking with a stranger. These supplemental results indeed suggest that “observations of single behaviors [...] do not correlate highly with anything, traits or otherwise” (Levine et al., 2012, p. 100).

**Table 3** Zero-order correlations between individual self-report items and behavioral coding categories.

Coded behavior	Self-report item	<i>r</i>
Express understanding verbally	I assure others that I am listening by using verbal acknowledgments.	.03
Uses exploring questions	I ask questions that show my understanding of others' positions.	-.10
Paraphrases information	I summarize points of agreement and disagreement when appropriate.	-.15
Smiling	I smile when I talk to people.	.09
Eye contact <sup>a</sup>	I look over or away from others while talking to them.	.04
	I avoid eye contact while talking to people.	.01
	I look directly at people while talking to them.	-.02
Facial pleasantness	I frown while talking to people.	.12
	I have a bland facial expression when I talk to people.	-.06
Forward body lean	I lean toward people when I talk to them.	-.10
	I lean away from people when I talk to them.	-.13
Body orientation	I have a relaxed body position when I talk to people.	.03
	I have a tense body position while talking to people.	.05
Open body position	I move away from others when they touch me while I am talking.	-.06
	I move closer to people when I talk to them.	.15
Gestural animation	I use my hands and arms to gesture while talking to people.	-.003
	I gesture when I talk to people.	.12
	I am animated when I talk to people.	.13
	I avoid gesturing while I am talking to people.	.18
Vocal pleasantness	I use a monotone or dull voice while talking to people.	-.11
	My voice is monotonous or dull when I talk to people.	-.08
	I use a variety of vocal expressions when I talk to people.	-.06
	I have a lot of vocal variety when I talk to people.	.05

Note: Relevant items were reverse-coded prior to analysis.  
<sup>a</sup>The item "I maintain eye contact with people when I talk to them" was inadvertently left off the SR version of the NIS.

With respect to the discrepancy between discloser (partner) reports of verbal and nonverbal listening behavior and the evaluations made by trained raters (for AELS:  $r = .30$ ; for NIS:  $r = .13$ ; see Table 2), disclosers may have (1) used different baseline information or (2) relied on different kinds of behaviors than trained raters when making their judgments. Of course, this may point to the utility of establishing behavioral baseline values, particularly in randomized control trials (RCTs) that manipulate some aspect of enacted support. In addition, disclosers may have had concrete emotional needs and objectives they needed to have addressed in the conversation, and that differed for raters. This point is captured in several theoretical models, such as the disclosure process model, which specifies that disclosers possess either approach goals that strengthen the relationship or avoidance goals that focus on avoiding negative relational outcomes (Chaudoir, Fisher, & Simoni, 2011). Furthermore, disclosers possess unique information about the behavioral patterns

of the listener that is directly accessible in the conversation and that may have influenced judgments. In addition, whereas raters are trained to look for specific cues and have the ability to watch the conversation multiple times to ensure consistent coding, disclosers may process the conversation in a *gestalt* fashion without much memory of the specific verbal or nonverbal behaviors that led to their impression. Indeed, the largest heterotrait–monomethod correlation was between partner-reported active listening and NI, suggesting near-isomorphic listening and immediacy partner assessments ( $r = .63/.75$ ; see Table 2). Such a finding is not necessarily surprising and reflects social psychological research showing that impressions of others, although based on specific behaviors, are driven primarily by superordinate attribute categories and that behaviors quickly dissipate from memory (Allen & Ebbesen, 1981); the same might well be true for listening and immediacy cues (Bodie et al., 2012).

### *What are the Implications of our Findings for Enacted Support?*

*What is the nature of enacted support?* Our results point to an ontological quandary for enacted support: Whose enacted support is it, anyway? If, as our results suggest, the nature of enacted support is in the eye of the beholder (i.e., the various decoder perspectives), does it make theoretical sense to determine which eye is more (or less) accurate? Viewed another way, is the theoretical segmentation of perceived support (which is an internalization of interpersonal experiences) and enacted support (which is observable behavior) a futile endeavor? The most cynical stance would be to maintain that we can never “truly” empirically ascertain enacted support because it is always a function of how it is perceived. If this is so, then there must be various kinds of enacted support: Consider that people assess enacted support through a host of lenses (e.g., relational, cultural, trait, and situational). The logical conclusion of this conjecture, then, is that all enacted support is perceived support. Such a view would certainly help us understand why perceived and enacted support are not related with one another in the real world even though these two constructs may exist in the context of research. A more moderate perspective might be that, although perception is crucial in generating meaning, we can certainly differentiate between what is perceived and what behaviors contribute to our perceptions. In other words, behavior and perception are inexorably linked, yet are easily differentiated when it comes to observation and analysis.

Contextual factors undoubtedly moderate a person's judgments of enacted support. This so-called situationism is prevalent in all empirical research and simply means that people's behaviors are influenced by circumstances (Funder, 2006). Circumstances vary a great deal in work on enacted support. We would argue, in line with Reis (2009), that situation indeed be given renewed emphasis in behavioral research by focusing particularly on interpersonal features. Reis (2009) provided a brief sample of studies, demonstrating the powerful effect of relationship contexts on social psychological phenomena, including social support (p. 321). We also point to recent research suggesting that several well-established biases, such as the actor-

observer bias, are perhaps more complex than believed and may need to be re-examined in the context of relationships-as-situations (Malle, 2006). Taken together, our studies recommend that researchers pay close attention to the kinds of interpersonal contexts they choose to use as proxy for or representation of enacted support.

Finally, our studies draw attention to the problem of representation. We have assumed thus far that internal representations of enacted support are a direct “read out” of the external world (i.e., what people actually do). To call behaviors enacted in the pursuit of support “objective” and any perceptions that match this standard as “accurate,” while labeling variables that increase or decrease behavioral evaluations as “biased,” is misleading. The perspective we adopt during a supportive encounter may matter more than the behaviors enacted. Consequently, evaluations of enacted support may differentially predict outcomes, such as health and well-being.

*Do trained coders do a better job than naive judges?* In Study 2, we examined multiple evaluations from third-party observers. Within-rater correlations were indicative of a high level of consistency (NIS  $r_i = .75$ ; Study 2) compared to correlations featuring ratings of trained coders (monotrait–monomethod NIS  $r = .81$ ; Table 2, Study 1). Within-rater correlations for listening evaluations were approximately equal among untrained observers (AELS  $r_i = .75$ , NVI  $r_i = .86$ ; Study 2) compared to those trained (monotrait–monomethod AELS  $r = .80$ , NVI  $r = .81$ ; Table 2, Study 1). These findings have important implications for design and research on enacted support because they behoove us to consider the advantages and disadvantages of using trained and untrained judges.

On the one hand, we may not need to spend time and money to train coders because naive judges seem to capture behavior equally reliably and consistently. In addition, using trained coders may accidentally introduce systematic measurement error. For example, coders may be mistakenly trained to rate forward lean incorrectly. Randomly selecting and assigning third-party observers would rule out such systematic measurement error. This point was made by Andersen (1979), who suggested that students could report their teachers’ immediacy as well as trained observers could. Indeed, the constrained variability among third-party observers seems to indicate a normative standard of “good support” that involves the enactment of specific verbal and nonverbal behaviors known to be indicative of “good listening” (Bodie, Vickery, et al., 2013), although we are well aware that our findings are likely context and relationship dependent (see Goldsmith, 2004, esp. p. 161).

On the other hand, we may want to use trained observers precisely because they have received several hours of instruction centered on a particular theoretical framework of enacted support. Deciding to train coders may be necessary if the construct to be coded is difficult to code (e.g., person-centeredness). The chief task of these coders is to code behavior reliably for etic categories and without any consideration of relationally idiosyncratic information because that information may actually be in the way of the study’s focus.

Past research examining certain aspects of enacted support has relied almost exclusively on the use of trained coders to generate “clean” reliabilities for different reasons (for an exception, see Priem et al., 2009). The argument of most scholars interested in supportive communication is that the behaviors we study affect lay observers in relatively stable ways primarily because these lay observers have similar implicit theories of support. That is, what constitutes supportive behavior for most people should be a relatively stable set of particular behaviors that, when enacted, can be evaluated as helpful, sensitive, and supportive (Goldsmith, McDermott, & Alexander, 2000). Based on our findings, we would stipulate that the decision about how behavioral coding is accomplished be viewed as a condition of the research design and research question at hand. Naïve third-party judges who belong to a unique subculture (e.g., Alzheimer’s disease patients, socially anxious people, and libertarians) may indeed be in a unique position to evaluate the behavior under study precisely because these lay observers possess similar implicit theories of support *for that subculture*. In short, the use of untrained observers may generate more ecologically valid data without having to forfeit consistency. The source of the coder may thus become a crucial component in a research design that can easily be captured with a series of measurable covariates.

### Limitations

As with any research, ours is not without its limitations. The first of these is that we captured only a narrow set of enacted support behaviors, namely, those behaviors associated with active listening and NI. Undoubtedly, listening and immediacy play crucial roles when supporting others. We also have to assume, however, that other support behaviors also play an important role when we comfort other people.

A second limitation of our study concerns the nature of NI in particular. The measure caused considerable consternation. Bluntly, the scale “fell apart” in Study 1 and simply did not confirm an *a posteriori* factor structure. That is problematic because it leads us to question the validity of the measure. For example, items did not converge onto a single immediacy dimension as posited; rather, NI seemed to be better represented by seven item parcels, which included nonverbal cues indicating (1) gestures; (2) touch; (3) voice; (4) eye behavior; (5) body position; (6) closeness; and, (7) facial expressions. To be sure, the problem also rests with ascertaining judgments about concrete behaviors. That is, a certain configuration of behaviors is conceptualized as generating a certain representation of warmth, care, and concern. Clearly, our results suggest that there are serious internal inconsistencies with respect to the set of nonverbal behaviors we used. Because NI is commonly explored as a nonverbal “representation” (see above) of emotional support, yet seems hard to capture accurately, we must examine what exactly immediacy is and how best to measure it.

A third limitation concerns the nature of our data. In Study 1 we asked participants to interact in a laboratory setting with a stranger, thus perhaps causing them to act differently than they would in a more informal setting or with a person

with whom they are familiar. Study 2 utilized similar data; the participants were third-party observers but observing strangers interacting in a laboratory (a “doubly stranger” design). Although effect sizes are lower in laboratory studies on supportive communication compared to pencil-and-paper designs, there are no differences in kind; that is, the direction of the effect for certain types of support is similar regardless of methodological differences (High & Dillard, 2012). Even so, there is still the chance that the anonymous environment of a lab setting influenced participant behavior.<sup>7</sup> Likewise, listeners might have engaged in a restricted set of behaviors due to the fact they were interacting with unknown others. Floyd and Morman (1998) found that self-reported affectionate communication, a construct similar to enacted support, predicted the display of immediacy behaviors in conversations among adult platonic friends. The degree to which interactions with friends and romantic partners, for instance, return similar results is simply an empirical question for which we do not have a satisfactory answer as yet. On the other hand, it might be that participants were motivated to “be on their best behavior” and to respond in ways consistent with the socially desirable nature of the activity. As politeness theory suggests, individuals are motivated to save face by putting their best self-face forward while also being mindful of the needs and desires of others. Thus, if an individual had it in his or her repertoire to be highly supportive, a conversation with a stranger is likely to draw out that behavior. This suggestion highlights a related limitation, namely, that a single conversation may not be the best method for detecting behavioral tendencies (Feeney et al., 2009). Perhaps a better method would place participants in multiple conversations, using the aggregate behavioral response as the measure of enacted support. Whatever form future work takes, this set of studies helps to clarify the different sources of data from which evaluations can be drawn when making judgments about enacted support.

### Conclusion

Limitations notwithstanding, this article highlights an important yet underappreciated assumption in work on enacted support. Virtually every study in supportive communication uses at least one of the three evaluation sources we investigated (self, partner, and coder), but few actually compare these sources or discuss differential patterns of association between different source evaluations and other outcomes. The AELS and the NIS assess two important constituents of enacted support. What our findings offer is cautionary evidence that when scholars examine evaluations of enacted support, support perspective matters and might ultimately contribute differently to well-being and health. If communication scientists are to continue to stress the importance of particular sets of behaviors to how we live our lives and relationships, future work should attend seriously to who is ultimately judging what counts as quality-enacted support.

### Acknowledgments

The authors would like to thank Tim Levine for serving as a consultant on Study 1. The authors also would like to thank Jonathan Denham, Michelle Pence, Trey Gibson, Logan Sacco, McCade McDaniel, Elizabeth McKee, Daniel Chapman, Lori Castano, Amanda Legrand, Nickole Hojnowski, Allison O'Neill, Dan Barberio, Billy Boland, and Kristin Carlson for their assistance with various aspects of data collection and coding. This study was supported by a Pilot Funding for New Research (Pfund) grant and a Research Competitiveness Subprogram (RCS) grant, both awarded to Graham D. Bodie from the Louisiana Board of Regents.

### Notes

- [1] Of these 171 dyads, 41 consisted of a trained confederate, whereas 130 paired two students unfamiliar with each other. This study reports on a subset of the 130 stranger dyads, namely, the 103 with full information to construct the MTMM matrix. Information about other aspects of this project is available upon request.
- [2] We considered as adequate models that exhibited a CFI value at or above .90 and a SRMR value at or below .08. Chi-square values were not considered in judging model fit. Although the approximation error (RMSEA) is reported for all scales, this value, when high, is not interpreted as an indication of poor model fit. Model fit is influenced by sample size and degrees of freedom (Kenny, Kaniskan, & McCoach, 2011), and our data featured small sample sizes for all measures and low degrees of freedom for many of the scales, which consequently attenuated model fit.
- [3] We also computed several other measures of internal consistency, each of which was comparable to the alpha values. Because most readers will be familiar with Cronbach's alpha, we report those values for convenience.
- [4] In the case of Priem et al., the observers were only minimally trained. They were provided a definition of supportiveness and were permitted to view the video multiple times before submitting evaluations. Other researchers, including the second author, have used more extensive training techniques (e.g., Jones & Guerrero, 2001).
- [5] The data presented for Study 2 have not been previously published, although the data-set under question has been used for other purposes.
- [6] We then randomly split the 72 conversations into two sets of 36. Although the initial plan was to collect data for 36 conversations at each research site, many students at the University of Minnesota, Twin Cities, who had signed up for time slots did not show up to the lab. Consequently, the majority of the data were collected at Louisiana State University A&M ( $n = 309$ ). All 72 conversations were utilized, and between two and seven participants viewed each video.
- [7] We asked listeners three questions regarding the normalcy of their behavior, and each of these three questions generated a mean above the midpoint ( $M = 5.69$ ,  $SD = 1.34$ ). Likewise, disclosers responded that they did not hide feelings ( $M = 2.36$ ,  $SD = 1.48$ ) and that they talked about their "true feelings" ( $M = 5.72$ ,  $SD = 1.37$ ). Nevertheless, we realize that the situation in which we placed participants is far from a "normal" conversational setting and, thus, has ramifications for the generalizability of our data.

### References

- Afifi, T., Afifi, W., Merrill, A. F., Denes, A., & Davis, S. (2013). "You need to stop talking about this!": Verbal rumination and the costs of social support. *Human Communication Research*, 39, 395–421. doi:10.1111/hcre.12012
- Allen, R. B., & Ebbesen, E. B. (1981). Cognitive processes in person perception: Retrieval of personality trait and behavioral information. *Journal of Experimental Social Psychology*, 17(2), 119–141. doi:10.1016/0022-1031(81)90010-X

- Almeida, D. M. (2005). Resilience and vulnerability to daily stressors assessed via diary methods. *Current Directions in Psychological Science*, 14(2), 64–68. doi:10.1111/j.0963-7214.2005.00336.x
- Anders, S. L., & Tucker, J. S. (2000). Adult attachment style, interpersonal communication competence, and social support. *Personal Relationships*, 7, 379–389. doi:10.1111/j.1475-6811.2000.tb00023.x
- Andersen, J. F. (1979). Teacher immediacy as a predictor of teaching effectiveness. *Communication Yearbook*, 3, 545–559.
- Andersen, J. F., Andersen, P. A., & Jensen, A. D. (1979). The measurement of nonverbal immediacy. *Journal of Applied Communication Research*, 7, 153–180. doi:10.1080/00909887909365204
- Andersen, P. A., & Andersen, J. F. (2005). Measurements of perceived nonverbal immediacy. In V. L. Manusov (Ed.), *The sourcebook of nonverbal measures: Going beyond words* (pp. 113–126). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bank, L., Dishion, T. J., Skinner, M., & Patterson, G. R. (1990). Method variance in structural equation modeling: Living with “glop”. In G. R. Patterson (Ed.), *Depression and aggression in family interaction* (pp. 247–279). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Barrera Jr, M., & Baca, L. M. (1990). Recipient reactions to social support: Contributions of enacted support, conflicted support, and network orientation. *Journal of Social and Personal Relationships*, 7, 541–551. doi:10.1177/0265407590074010
- Bernieri, F. J., & Gillis, J. S. (2001). Judging rapport: Employing Brunswick’s lens model to study interpersonal sensitivity. In J. A. Hall & F. J. Bernieri (Eds.), *Interpersonal sensitivity: Theory and measurement* (pp. 67–88). Mahwah, NJ: Lawrence Erlbaum Associates.
- Bodie, G. D. (2011). The Active-Empathic Listening Scale (AELS): Conceptualization and evidence of validity within the interpersonal domain. *Communication Quarterly*, 59, 277–295. doi: 10.1080/01463373.2011.583495
- Bodie, G. D., & Burleson, B. R. (2008). Explaining variations in the effects of supportive messages: A dual-process framework. *Communication Yearbook*, 32, 355–398.
- Bodie, G. D., Gearhart, C. C., Denham, J. P., & Vickery, A. J. (2013). The temporal stability and situational contingency of active-empathic listening. *Western Journal of Communication*, 77(2), 113–138. doi:10.1080/10570314.2012.656216
- Bodie, G. D., & Jones, S. M. (2012). The nature of supportive listening II: The role of verbal person centeredness and nonverbal immediacy. *Western Journal of Communication*, 76, 250–269. doi:10.1080/10570314.2011.651255
- Bodie, G. D., St. Cyr, K., Pence, M., Rold, M., & Honeycutt, J. (2012). Listening competence in initial interactions I: Distinguishing between what listening is and what listeners do. *International Journal of Listening*, 26(1), 1–28. doi:10.1080/10904018.2012.639645
- Bodie, G. D., Vickery, A. J., Cannava, K., & Jones, S. M. (in press). The role of “active listening” in informal helping conversations: Impact on perceptions of listener helpfulness, sensitivity, and supportiveness and discloser emotional improvement. *Western Journal of Communication*.
- Bodie, G. D., Vickery, A. J., & Gearhart, C. C. (2013). The nature of supportive listening I: Exploring the relation between supportive listeners and supportive people. *International Journal of Listening*, 27(1), 39–49. doi:10.1080/10904018.2013.732408
- Burleson, B. R. (1983). Social cognition, empathic motivation, and adults’ comforting strategies. *Human Communication Research*, 10, 295–304. doi:10.1111/j.1468-2958.1983.tb00019.x
- Burleson, B. R. (2003). Emotional support skill. In J. O. Greene & B. R. Burleson (Eds.), *Handbook of communication and social interaction skills* (pp. 551–594). Mahwah, NJ: Lawrence Erlbaum Associates.
- Burleson, B. R. (2007). What’s most important about communication? About communication and social support. *Communication Currents*, 2.
- Burleson, B. R., & Goldsmith, D. J. (1998). How the comforting process works: Alleviating emotional distress through conversationally induced reappraisals. In P. A. Andersen & L. K. Guerrero (Eds.), *Handbook of communication and emotion: Research, theory, applications, and contexts* (pp. 245–280). San Diego, CA: Academic Press.
- Burleson, B. R., & Samter, W. (1985). Individual differences in the perception of comforting messages: An exploratory investigation. *Central States Speech Journal*, 36(1–2), 39–50. doi:10.1080/10510978509363198
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multi-trait multi-method matrix. *Psychological Bulletin*, 52, 281–301. doi:10.1037/h0046016

- Cappella, J. N., & Street, R. L., Jr. (1989). Message effects: Theory and research on mental models of messages. In J. J. Bradac (Ed.), *Message effects in communication science* (pp. 24–51). Newbury Park, CA: Sage.
- Carlson, K. D., & Herdman, A. O. (2011). Understanding the impact of convergent validity on research results. *Organizational Research Methods*, 15, 17–32. doi:10.1177/1094428110392383
- Cassel, J. (1976). The contribution of the social environment to host resistance. *American Journal of Epidemiology*, 104, 107–123.
- Chaudoir, S. R., Fisher, J. D., & Simoni, J. M. (2011). Understanding HIV disclosure: A review and application of the disclosure process model. *Social Science & Medicine*, 72, 1618–1629. doi:10.1016/j.socscimed.2011.03.028
- Chen, Y., & Feeley, T. H. (2012). Enacted support and well-being: A test of the mediating role of perceived control. *Communication Studies*, 63, 608–625. doi:10.1080/10510974.2012.674619
- Collins, N. L., & Feeney, B. C. (2000). A safe haven: An attachment theory perspective on support-seeking and caregiving in adult romantic relationships. *Journal of Personality and Social Psychology*, 78, 1053–1073. doi:10.1037/0022-3514.78.6.1053
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302. doi:10.1037/h0040957
- Drollinger, T., Comer, L. B., & Warrington, P. T. (2006). Development and validation of the active empathetic listening scale. *Psychology and Marketing*, 23, 161–180. doi:10.1002/mar.20105
- Elliott, R., Bohart, A. C., Watson, J. C., & Greenberg, L. S. (2011). Empathy. *Psychotherapy*, 48(1), 43–49. doi:10.1037/a0022187
- Fassaert, T., van Dulmen, S., Schellevis, F., & Bensing, J. (2007). Active listening in medical consultations: Development of the active listening observation scale. *Patient Education and Counseling*, 68, 258–264. doi:10.1016/j.pec.2007.06.011
- Feeley, T. H. (2002). Comment on halo effects in rating and evaluation research. *Human Communication Research*, 28, 578–586. doi:10.1111/j.1468-2958.2002.tb00825.x
- Feeney, B. C., Cassidy, J., Lemay, Jr, E. P., & Ramos-Marcuse, F. (2009). Affiliation with new peer acquaintances during two initial social support interactions. *Personal Relationships*, 16, 489–506. doi:10.1111/j.1475-6811.2009.01236.x
- Feeney, B. C., & Collins, N. L. (2001). Predictors of caregiving in adult intimate relationships: An attachment theoretical perspective. *Journal of Personality and Social Psychology*, 80, 972–994. doi:10.1037/0022-3514.80.6.972
- Floyd, K. (2014). Empathic listening as an expression of interpersonal affection. *International Journal of Listening*, 28(1), 1–12. doi:10.1080/10904018.2014.861293
- Floyd, K., & Morman, M. T. (1998). The measurement of affectionate communication. *Communication Quarterly*, 46, 144–162. doi:10.1080/01463379809370092
- Fredrickson, B. L. (2013). *Love 2.0: How our supreme emotion affects everything we feel, think, do, and become*. New York, NY: Hudson.
- Funder, D. C. (2006). Towards a resolution of the personality triad: Persons, situations and behaviors. *Journal of Research in Personality*, 40(1), 21–34. doi:10.1016/j.jrp.2005.08.003
- Gable, S. L., Gosnell, C. L., Maisel, N. C., & Strachman, A. (2012). Safely testing the alarm: Close others' responses to personal positive events. *Journal of Personality and Social Psychology*, 103, 963–981. doi:10.1037/a0029488
- Gearhart, C. C., & Bodie, G. D. (2011). Active-empathic listening as a general social skill: Evidence from bivariate and canonical correlations. *Communication Reports*, 24(2), 86–98. doi:10.1080/08934215.2011.610731
- Goldsmith, D. J. (2004). *Communicating social support*. New York, NY: Cambridge University Press.
- Goldsmith, D. J., McDermott, V. M., & Alexander, S. C. (2000). Helpful, supportive, and sensitive: Measuring the evaluation of enacted support in personal relationships. *Journal of Social and Personal Relationships*, 17, 369–391. doi:10.1177/0265407500173004
- High, A. C., & Dillard, J. P. (2012). A review and meta-analysis of person-centered messages and social support outcomes. *Communication Studies*, 63(1), 99–118. doi:10.1080/10510974.2011.598208
- Jones, S. M. (2011). Supportive listening. *International Journal of Listening*, 25(1–2), 85–103. doi:10.1080/10904018.2011.536475

- Jones, S. M., & Guerrero, L. K. (2001). Nonverbal immediacy and verbal person-centeredness in the emotional support process. *Human Communication Research*, 27, 567–596. doi:[10.1111/j.1468-2958.2001.tb00793.x](https://doi.org/10.1111/j.1468-2958.2001.tb00793.x)
- Jones, S. M., & Wirtz, J. G. (2006). How does the comforting process work?: An empirical test of an appraisal-based model of comforting. *Human Communication Research*, 32, 217–243. doi:[10.1111/j.1468-2958.2006.00274.x](https://doi.org/10.1111/j.1468-2958.2006.00274.x)
- Jones, S. M., & Wirtz, J. G. (2007). “Sad Monkey See, Monkey Do”: Nonverbal matching in emotional support encounters. *Communication Studies*, 58(1), 71–86. doi:[10.1080/10510970601168731](https://doi.org/10.1080/10510970601168731)
- Kenny, D. A., & Acitelli, L. K. (2001). Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology*, 80, 439–448. doi:[10.1037/0022-3514.80.3.439](https://doi.org/10.1037/0022-3514.80.3.439)
- Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2011). *The performance of RMSEA in models with small degrees of freedom*. University of Connecticut, Unpublished paper.
- Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2nd ed.). New York, NY: Guilford.
- Kotowski, M. R., Levine, T. R., Baker, C. R., & Bolt, J. M. (2009). A multitrait-multimethod assessment of the verbal aggressiveness and argumentativeness scales. *Communication Monographs*, 76, 443–462. doi:[10.1080/03637750903300247](https://doi.org/10.1080/03637750903300247)
- Lakey, B., & Orehek, E. (2011). Relational regulation theory: A new approach to explain the link between perceived social support and mental health. *Psychological Review*, 118, 482–495. doi:[10.1037/a0023477](https://doi.org/10.1037/a0023477)
- Lakey, B., Orehek, E., Hain, K. L., & VanVleet, M. (2009). Enacted support’s links to negative affect and perceived support are more consistent with theory when social influences are isolated from trait influences. *Personality and Social Psychology Bulletin*, 36, 132–142. doi:[10.1177/0146167209349375](https://doi.org/10.1177/0146167209349375)
- Lakey, B., Ross, L. T., Butler, C., & Bentley, K. (1996). Making social support judgments: The role of similarity and conscientiousness. *Journal of Social & Clinical Psychology*, 15, 283–304. doi:[10.1521/jscp.1996.15.3.283](https://doi.org/10.1521/jscp.1996.15.3.283)
- Levine, T. R., Kotowski, M. R., Beatty, M. J., & Van Kelegom, M. J. (2012). A meta-analysis of trait-behavior correlations in argumentativeness and verbal aggression. *Journal of Language and Social Psychology*, 31(1), 95–111. doi:[10.1177/0261927X11425037](https://doi.org/10.1177/0261927X11425037)
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling*, 9, 151–173. doi:[10.1207/S15328007SEM0902\\_1](https://doi.org/10.1207/S15328007SEM0902_1)
- MacGeorge, E. L., Feng, B., & Burleson, B. R. (2011). Supportive communication. In M. L. Knapp & J. A. Daly (Eds.), *Handbook of interpersonal communication* (4th ed., pp. 317–354). Thousand Oaks, CA: Sage.
- MacGeorge, E. L., Feng, B., & Thompson, E. R. (2008). “Good” and “bad” advice: How to advise more effectively. In M. T. Motley (Ed.), *Studies in applied interpersonal communication* (pp. 145–164). Thousand Oaks, CA: Sage.
- Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin*, 132, 895–919. doi:[10.1037/0033-2909.132.6.895](https://doi.org/10.1037/0033-2909.132.6.895)
- McCroskey, J. C., Richmond, V. P., & Bennett, V. E. (2006). The relationships of student end-of-class motivation with teacher communication behaviors and instructional outcomes. *Communication Education*, 55, 403–414. doi:[10.1080/03634520600702562](https://doi.org/10.1080/03634520600702562)
- Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, 12, 531–544. doi:[10.1177/014920638601200408](https://doi.org/10.1177/014920638601200408)
- Priem, J. S., Solomon, D. H., & Steuber, K. R. (2009). Accuracy and bias in perceptions of emotionally supportive communication in marriage. *Personal Relationships*, 16, 531–552. doi:[10.1111/j.1475-6811.2009.01238.x](https://doi.org/10.1111/j.1475-6811.2009.01238.x)
- Reis, H. T. (2009). Relationships are situations, and situations involve relationships. *Journal of Research in Personality*, 43, 266. doi:[10.1016/j.jrp.2008.12.036](https://doi.org/10.1016/j.jrp.2008.12.036)
- Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods*, 12, 762–800. doi:[10.1177/1094428109332834](https://doi.org/10.1177/1094428109332834)

- Richmond, V. P., McCroskey, J. C., & Johnson, A. D. (2003). Development of the Nonverbal Immediacy Scale (NIS): Measures of self- and other-perceived nonverbal immediacy. *Communication Quarterly*, 51, 504–517. doi:10.1080/01463370309370170
- Rosenthal, R. (1987). *Judgment studies: Design, analysis, and meta-analysis*. Cambridge, UK: Cambridge University Press.
- Rosenthal, R. (2005). Conducting judgment studies: Some methodological issues. In J. A. Harrigan, R. Rosenthal, & K. R. Scherer (Eds.), *The new handbook of methods in nonverbal behavior research* (pp. 199–234). Oxford: Oxford University Press.
- Sarason, I. G., & Sarason, B. R. (2009). Social support [special issue]. *Journal of Social and Personal Relationships*, 26(1). Retrieved from <http://spr.sagepub.com/content/26/1.toc>
- Smythe, M.-J., & Hess, J. A. (2005). Are student self-reports a valid method for measuring teacher nonverbal immediacy? *Communication Education*, 54, 170–179. doi:10.1080/03634520500213389
- Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, 15(1), 72–101. doi:10.2307/1412159
- Street, R. L. (1985). Participant-observer differences in speech evaluation. *Journal of Language and Social Psychology*, 4(2), 125–130. doi:10.1177/0261927X8500400204
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Pearson Education.
- Tickle-Degnen, L., & Rosenthal, R. (1992). Nonverbal aspects of therapeutic rapport. In R. S. Feldman (Ed.), *Applications of nonverbal behavioral theories and research* (pp. 143–164). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Trees, A. R. (2000). Nonverbal communication and the support process: Interactional sensitivity in interactions between mothers and young adult children. *Communication Monographs*, 67, 239–261. doi:10.1080/03637750009376509
- Uchino, B. N., Carlisle, M., Birmingham, W., & Vaughn, A. A. (2011). Social support and the reactivity hypothesis: Conceptual issues in examining the efficacy of received support during acute psychological stress. *Biological Psychology*, 86(2), 137–142. doi:10.1016/j.biopsycho.2010.04.003
- Uno, D., Uchino, B. N., & Smith, T. W. (2002). Relationship quality moderates the effect of social support given by close friends on cardiovascular reactivity in women. *International Journal of Behavioral Medicine*, 9, 243–262. doi:10.1207/S15327558IJBM0903\_06
- Weger Jr, H. W., Castle, G. R., & Emmett, M. C. (2010). Active listening in peer interviews: The influence of message paraphrasing on perceptions of listening skill. *International Journal of Listening*, 24(1), 34–49. doi:10.1080/10904010903466311
- Weiner, A. S. B., & Hannum, J. W. (2013). Differences in the quantity of social support between geographically close and long-distance friendships. *Journal of Social and Personal Relationships*, 30, 662–672. doi:10.1177/0265407512465997