ORIGINAL ARTICLE

Parasocial Interaction and Parasocial Relationship: Conceptual Clarification and a Critical Assessment of Measures

Jayson L. Dibble[1], Tilo Hartmann[2], & Sarah F. Rosaen[3]

1 Department of Communication, Hope College, Holland, MI 49423, USA
2 Department of Communication Science, VU University Amsterdam, De Boelelaan 1081,
1081 HV Amsterdam, The Netherlands
3 Department of Communication and Visual Arts, University of Michigan-Flint, MI 48502, USA

Parasocial interaction and parasocial relationship are often conflated conceptually and methodologically, leaving researchers unclear as to which concept is being tapped. This research clarifies these concepts and experimentally compares the most common measure of parasocial interaction, the Parasocial Interaction Scale (PSI-Scale), with a newer measure, the Experience of Parasocial Interaction Scale (EPSI-Scale). Participants (N = 383) viewed a brief videorecording of a woman who either bodily addressed the viewer or not, then completed a questionnaire. The EPSI-Scale was a better measure of parasocial interaction, understood as a within-viewing experience of mutual awareness, whereas the PSI-Scale may measure short- or long-term liking, or something else. To avoid conceptual and empirical confusion, researchers must choose measures with greater care.

Parasocial interaction is a concept that remains popular among those who study media and media effects. Originally put forth by Horton and Wohl (1956), parasocial interaction referred to a media user's reaction to a media performer such that the media user perceives the performer as an intimate conversational partner. A sense of conversational give-and-take often emerges during viewing and is strongest when the media performer bodily addresses the viewer through the camera. This early work by Horton and Wohl also alluded to a related concept, the parasocial relationship, which is a more enduring relationship that a media user forms with a mediated performer. While parasocial interaction is restricted to the viewing episode, a parasocial relationship can extend beyond any single viewing episode.

Corresponding author: Jayson L. Dibble; e-mail: dibble@hope.edu

Since Horton and Wohl's foundational work, these concepts have found favor with researchers, and numerous measures have emerged for assessing parasocial interaction and parasocial relationship. Unfortunately, in the decades that followed, a fair amount of conceptual and empirical confusion has resulted, as measures purporting to measure either construct may not be assessing those constructs with the validity and precision they could be. Research conclusions about parasocial interaction and/or parasocial relationship hinge on the measures of those constructs, and any claims about how either construct relates with other constructs can never be more valid than the operationalization of the construct. Thus, researchers of parasocial phenomena must remain vigilant about the quality of their measures.

However, as researchers have noted (e.g., Levine & Kotowski, 2010), most published measures generally are not tested with sufficient rigor, and those that are rigorously tested tend not to fare well. We submit that most parasocial interaction measures have not been put through adequate tests of construct validation. Moreover, parasocial interaction and parasocial relationship have undergone much theoretical and conceptual refinement since the first measures of these concepts emerged, and we suggest that extant measures have not been evaluated sufficiently for their correspondence with the latest theoretic and conceptual developments. We contend that the measures used most frequently (e.g., Parasocial Interaction Scale [PSI-Scale]; Rubin, Perse, & Powell, 1985; Rubin & Perse, 1987) lack content validity for measuring parasocial interaction because they are neither consistent with Horton and Wohl's conceptual definition of parasocial interaction (understood as conversational give-and-take *during* viewing), nor with recent conceptual refinements. In summary, research on parasocial interaction is in need of conceptual clarification and additional validation of its relevant measures.

Our purpose for this research is to assess critically the measurement of parasocial interaction in light of seminal and contemporary approaches to these concepts. We will argue that certain measures (or specific items from measures) commonly taken to measure parasocial interaction are actually measuring something else, that parasocial interaction and parasocial relationship are neither conceptually interchangeable nor empirically interchangeable, and that researchers should think carefully about what they are trying to measure when they choose a measure of parasocial interaction.

In addition, we provide additional validity testing for the Experience of Parasocial Interaction Scale (EPSI-Scale; Hartmann & Goldhoorn, 2011), a relatively newer instrument for measuring parasocial interaction that captures Horton and Wohl's (1956) original ideas about parasocial interaction without also tapping the more enduring parasocial relationship.

Conceptualizing parasocial phenomena
Parasocial interaction
The term *parasocial interaction* was coined in a now-classic essay by Horton and Wohl (1956). At the time they wrote their article, Horton and Wohl were both

affiliated with the Department of Sociology at the University of Chicago. In the 1950s, this sociology department was strongly influenced by the theory of symbolic interactionism, and the ideas of George Herbert Mead and his successors at the university's Department of Philosophy (Gusfield, 2003). In the 1950s, Donald Horton frequently dabbled in the study of communication, for example, as a thesis advisor of Kurt Lang's MacArthur Day study (Lang & Lang, 1953), and of Erving Goffman's examination of conversational interaction (Peters & Simonson, 2004). Meanwhile, Richard Wohl was interested in understanding the processes of symbolic interaction between mass media performers and their users as a part of popular culture (Strauss, 1958).

Horton and Wohl's (1956) elaboration of the concept of parasocial interaction has to be understood in this context (Beniger, 1987; Ellis, Streeter, & Engelbrecht, 1983). They defined parasocial interaction as a "simulacrum of conversational give-and-take" (p. 215) that users experience as a response to a media performer (the "persona") in a media exposure situation. Horton and Wohl suggested that users perceive a parasocial interaction as an intimate reciprocal social interaction, despite knowing that it is only an illusion. Distinguishing parasocial interaction from other types of purely imaginative interactions with a media performer (e.g., Ellis et al., 1983), Horton and Strauss (1957) further stressed that parasocial interaction is initiated and directed by the media performer in exposure situations. According to Horton and Wohl (1956) and Horton and Strauss (1957), a parasocial interaction is triggered if media performers acknowledge the presence of the audience in their performance, adapt the conversational style of informal face-to-face gatherings, and bodily and verbally address their users.

Experience of parasocial interaction

In a recent approach, Hartmann and Goldhoorn (2011) linked back to Horton and Wohl's (1956) original ideas and enriched them with more recent social-psychological insights into social interaction. They also suggested defining parasocial interaction as an illusory user experience that takes place in an exposure situation. According to their view, users' parasocial interaction experience is "characterized by a felt reciprocity with a TV performer that comprises a sense of mutual awareness, attention, and adjustment" (p. 1107). Hartmann and Goldhoorn argued that certain cues provided by a media performer (e.g., eye-gazing, bodily addressing cues) effectively trigger automatic mindreading activities in users that, in turn, give rise to a parasocial experience, understood as a user's intuitive (gut) feeling of taking part in a normal social interaction. Two experimental studies provided preliminary evidence for the idea that particularly bodily addressing and eye-gazing effectively trigger users' parasocial interaction experience (Cummins & Cui, 2014; Hartmann & Goldhoorn, 2011).

Parasocial relationships

Several scholars suggest that parasocial interaction, understood as illusory experiences that are confined to the media exposure situation, should be distinguished from

more enduring (positive or negative) long-term parasocial relationships or socioe-motional bonds that users develop with media performers (e.g., Cummins & Cui, 2014; Dibble & Rosaen, 2011; Hartmann & Goldhoorn, 2011; Klimmt, Hartmann, & Schramm, 2006; Rosaen & Dibble, 2012; Tukachinsky, 2010). Such a distinction appears to be in line with Horton and Wohl's (1956) original conceptualization.

However, Horton and Wohl's use of the term "parasocial relationship" was somewhat ambiguous. On the one hand, being scholars interested in understanding the structure of (symbolic) social interactions, they used the term "parasocial relationship" to address the specific asymmetrical nature of the interaction relationship between a media performer and a user *during any single exposure situation* (see also Horton & Strauss, 1957). On the other hand, Horton and Wohl (1956) also referred to parasocial relationships to characterize the more enduring, long-term, and usually positive, one-sided intimacy at a distance that users develop toward media performers, based on repeated encounters.

Long-term social involvement: The approach by Rubin et al. (1985)

Maybe as a consequence of this ambiguity, potential differences between parasocial interaction and parasocial relationship that researchers acknowledge today were largely neglected in Rubin et al.'s (1985) widely influential adaptation of both concepts to uses-and-gratifications research. Rubin et al. (1985) proposed to understand parasocial interaction and relationship jointly as an overarching long-term social involvement, a conceptualization that arguably differs from the original approach proposed by Horton and Wohl (1956) and Horton and Strauss (1957). How did it come to this new reinterpretation of parasocial phenomena?

In general, uses-and-gratifications researchers in the 1970s and 1980s were less concerned with examining parasocial interaction as a specific type of social interaction, but treated parasocial interaction and relationship as a form of social involvement or gratification that users seek from the media (e.g., Babrow, 1987; Levy & Windahl, 1984; Wenner, 1983). Only a few adaptations in uses-and-gratifications research distinguished parasocial interaction from more enduring forms of parasocial attachment. For example, Rosengren, Windahl, Hakansson, and Johnsson-Smaragdi (1976) defined parasocial interaction as a phenomenon that exists mainly during the very act of TV viewing, thus explicitly distinguishing parasocial interaction from users' (short- and long-term) identification with a media character.

In contrast, in an influential approach, Nordlund (1978) nullified previously held conceptual distinctions between identification and interaction, respectively, between forms of short- and long-term social involvement with media characters. In his approach, he subsumed various forms of social involvement under the label of "media interaction," defined as a

> multifaceted phenomenon of audience involvement in certain portions of media fare. Such involvement may range from readiness to be preoccupied with and muse upon media characters, to talk about them with others, to relate their circumstances to one's

own, to feel as if in some relationship with them and, at the extreme, to get fully involved in their fate. (p. 151)

We find it noteworthy that Nordlund (1978) explicitly highlighted that his concept of media interaction "resembles, yet differs from, both the Horton and Wohl (1956) concept of parasocial interaction and the … measur[e] of parasocial interaction … promulgated by Rosengren et al. (1976)" (p. 151).

The approach by Rubin et al. (1985) bears striking similarities to Nordlund's conceptualization of media interaction. In their widely cited article, Rubin et al. (1985) provided the most influential conceptualization of parasocial interaction (and relationship), by laying out the theoretical basis of the PSI-Scale, the most popular measure to assess parasocial phenomena. Explicitly contrasting their approach to the idea of Horton and Wohl that "parasocial interaction exists only for the duration of the viewing experience" (p. 156), Rubin et al. understood various forms of social involvement, including "interaction, identification, and long-term identification with television characters" (p. 156) as parasocial interaction. Accordingly, they dropped the previously made distinction between parasocial interaction (in an exposure situation) and a (long-term) parasocial relationship. Instead, closely resembling Nordlund's (1978) overarching concept of media interaction, Rubin et al. defined parasocial interaction broadly as an "interpersonal involvement of the media user with what he or she consumes … including seeking guidance from a media persona, seeing media personalities as friends, imagining being part of a favorite program's social world, and desiring to meet media performers" (p. 156).

Having discussed parasocial interaction, parasocial relationship, and how the two have become confused, we clarify our current conceptual perspective. Parasocial interaction and parasocial relationship refer to related but distinct theoretical concepts. Parasocial interaction refers to a faux sense of mutual awareness that can only occur during viewing. In contrast, parasocial relationship refers to a longer-term association that may begin to develop during viewing, but also extends beyond the media exposure situation.

Alternatively, parasocial relationship can develop without any parasocial interaction. Such might be the case for a viewer who observes a character that does not break the "fourth wall" (Auter & Davis, 1991, p. 165) and directly address the viewer (e.g., most fictional characters). Thus, even though no illusory mutual awareness is occurring (no parasocial interaction), the viewer can still form a longer-term association with the character (parasocial relationship). Because clear theoretical conceptions are necessary to develop sound measures, researchers must take care not to conflate the two concepts. We discuss the measurement of parasocial phenomena in the next section.

Measuring parasocial phenomena
In the past, several measures have been suggested for assessing parasocial interaction phenomena, of which the PSI-Scale proposed by Rubin et al. (1985) has been by far the most popular. The complex and at times confusing history of conceptualizations

of parasocial phenomena is also reflected in the diversity of existing measures. Below we review some of the more prominent measures in use today. In the review, we follow the notion that a theoretical distinction between parasocial interaction and parasocial relationship has been laid out in Horton and Wohl's (1956) original conceptualization, and that assessing both concepts differently may prove fruitful for the field (e.g., Cummins & Cui, 2014). Accordingly, we pay particular attention to operational definitions, that is, the theoretical concept implied by each measure.

Rubin et al.'s (1985) PSI-Scale

To date, this is the most widely applied instrument in the field. The PSI-Scale is either applied in a 20-item version (Rubin et al., 1985) or a 10-item short version (Rubin & Perse, 1987), with the short version generally being more popular. The scale was originally developed to measure parasocial interaction as viewers' social involvement with local TV newscasters. Since then, it has been adapted (e.g., often by selectively dropping items or revising their original wording) to a multitude of other media characters, to include online avatars, movie characters, and politicians (for overviews see Giles, 2002; Schiappa, Allen, & Gregg, 2007; Tsao, 2004). In past studies, the PSI-Scale proved to be internally consistent and unidimensional (Rubin et al., 1985).

The operational definition implied by the PSI-Scale can be understood by inspecting the content of the scale's items (Schramm & Hartmann, 2008). To develop the PSI-Scale, Rubin et al. (1985) compiled items from several existing studies. Many of the 20 items are almost identical in wording to the seven items proposed by Levy (1979, p. 73) to measure parasocial involvement (e.g., "I like hearing the voice of my favorite newscaster in my home" or "The news program shows me what the newscasters are like"). In addition, some of the items bear close similarity to Houlberg's (1984, p. 426) 5-item measure of TV viewers' parasocial interaction to local newscasters. Interestingly, a couple of items are almost identical to Nordlund's (1978, p. 174) measures of long-term media interaction, that is, measures intended to assess a concept explicitly distinct from parasocial interaction (e.g., "I see my favorite newscaster as a natural, down-to-earth person," or "If my favorite newscaster appeared on another television program, I would watch that program").

Consistent with Rubin et al.'s (1985) notion of parasocial interaction as a long-term social involvement, only a few of the PSI-Scale's items tap into users' illusory experience of a conversational give-and-take in a media exposure situation (e.g., "When I'm watching the newscast, I feel as if I am part of their group"). Instead, most of the items on the PSI-Scale seem to assess users' *liking* of a media character, who they perceive to be a down-to-earth person and almost a friend and who they desire "to meet" in further encounters (Dibble & Rosaen, 2011; Klimmt et al., 2006; Schramm & Hartmann, 2008).

In an experimental examination of the PSI-Scale, Auter (1992) found that PSI-Scale scores were weakly but significantly higher among participants that watched a short episode of a TV comedy show from the 1950s in which they

were directly addressed by an unknown media character, as compared to participants that watched the same episode without being directly addressed. These higher scores may have resulted from the fact that users liked the TV character more if they were directly addressed (see also Auter & Davis, 1991). This notion is also consistent with findings by Dibble and Rosaen (2011) who observed much higher scores on the PSI-Scale for users rating liked as compared to disliked media characters. In summary, these findings suggest that the PSI-Scale's sensitivity to short-term media exposures could simply reflect a positive attitude toward the media persona, not the experience of mutual awareness that characterizes parasocial interaction.

Hartmann and Goldhoorn's (2011) Experience of Parasocial Interaction (EPSI) Scale
Based on the concern that the PSI-Scale does not measure parasocial interaction as originally conceptualized by Horton and Wohl (1956); Hartmann and Goldhoorn (2011) proposed an alternative 6-item measure, the EPSI-Scale. The scale intends to measure users' experience of parasocial interaction, defined as an intuitive feeling of mutual awareness, attention, and adjustment with a media character in an exposure situation. In three experimental examinations, the EPSI-Scale demonstrated strong reliability and validity (Cummins & Cui, 2014; Dibble & Rosaen, 2011; Hartmann & Goldhoorn, 2011). Parasocial interaction, if assessed by the EPSI-Scale, seems to be specifically triggered by forms of bodily addressing and eye-gazing of a media character (Cummins & Cui, 2014; Hartmann & Goldhoorn, 2011).

Confirming discriminant validity of the EPSI-Scale, the scale was only moderately correlated with the PSI-Scale in previous studies (Dibble & Rosaen, 2011; Hartmann & Goldhoorn, 2011), and the two scales also factor separately (Hartmann & Goldhoorn, 2011). In addition, whereas the PSI-Scale was much more sensitive to a deliberate variation of character liking, EPSI-Scale scores were only weakly influenced (Dibble & Rosaen, 2011).

Other scales
In the past, several other measures have been proposed to measure parasocial interaction phenomena. Next to the EPSI-Scale, at least three other scales have been published in an attempt to provide alternative measures that may overcome the potential limitations of the PSI-Scale. The 112-item *PSI-Process Scales* (Schramm & Hartmann, 2008; Schramm & Wirth, 2010) can be applied to measure parasocial interaction defined as parasocial processing, that is, users' psychological (cognitive, affective, and behavioral) involvement with a media character in an exposure situation. Furthermore, the 22-item *Audience–Persona Interaction* (*API*) *Scale* by Auter and Palmgreen (2000) was designed to measure parasocial interaction defined as a positive long-term involvement with a favorite media character. Finally, Tukachinsky's (2010) 24-item *Multiple-PSR Scale*, assesses parasocial relationship as (a) para-friendship, the imagined support and intimacy if the media character were a real person, and (b) para-love, perceiving a media character as physically attractive and desiring the media character.

Research propositions

As we have shown, parasocial interaction and parasocial relationship are related but conceptually distinct phenomena, and researchers should not interchange them haphazardly. Of critical importance to instantiating parasocial interaction as an intuitive feeling of mutual awareness, attention, and adjustment with a media character in an exposure situation is the sense that viewers feel they are being directly addressed by the media persona. Auter and Davis (1991) referred to this as breaking the invisible fourth wall. Researchers have consistently found stronger parasocial interaction when the performer broke the fourth wall by bodily addressing the viewer and looking into the camera than when the performer did not (e.g., Cummins & Cui, 2014; Hartmann & Goldhoorn, 2011; Mancini, 1988).

Measures of parasocial interaction and parasocial relationship might be differentiated by exposing viewers to a brief videorecorded message of a person who either bodily addresses the viewer (through body orientation and looking into the camera) or not. Hartmann and Goldhoorn (2011) followed this basic process. The videorecording should be long enough to permit in-the-moment experience of parasocial interaction, but short enough to prevent development of substantial forms of the more enduring parasocial relationship (e.g., intense friendship). In general, an experimental manipulation designed to instantiate within-the-moment parasocial interaction should associate more strongly with measures of parasocial interaction (e.g., EPSI-Scale; Hartmann & Goldhoorn, 2011) than with measures designed to tap long-term social involvement (e.g., PSI-Scale; Rubin et al., 1985). Therefore, we predict the following:

> H1: The effect of viewers being bodily addressed versus not bodily addressed by a media performer will result in greater variation on the EPSI-Scale than on a measure of long-term social involvement, specifically, the PSI-Scale.

Because both parasocial interaction and parasocial relationship fall under the broader category of parasocial phenomena, and because both concepts can develop during the viewing episode, we expect some overlap among measures of parasocial interaction and parasocial relationship. That is, certain items from both measures might be to some extent sensitive to our experimental induction (see also Auter & Davis, 1991). However, because past research using these measures only reports on the items in aggregate, establishing predictions for any specific item is difficult. Therefore, we pose the following research question:

> RQ1: At the item level, how do indicators of the EPSI-Scale compare with indicators of the PSI-Scale in terms of how sensitive they are to the experimental induction?

The process of construct validation involves evaluating the extent to which a measure assesses whatever it is purported to assess and nothing else, according to an a priori conceptual definition (Cronbach & Meehl, 1955; Pedhazur & Schmelkin, 1991). Construct validity (CV) is in part demonstrated by noting whether measures of the test construct correlate with measures of other constructs according to theory

(nomological network; Cronbach & Meehl, 1955). For nomological network validation, we selected measures purported to tap identification, wishful identification, and relationship closeness. The former two are well-established variables in the media effects literature, but the latter may not be as obvious. We chose closeness because a parasocial relationship in many ways resembles a real-life relationship, and many of the cognitive processes that govern the formation of parasocial relationships resemble those used in real-life relationships. For example, parasocial relationships and real-life relationships both can serve similar functions (e.g., providing social support, Finn & Gorr, 1988; Schiappa et al., 2007; Tsao, 1996) and both can trigger distress if broken up (e.g., parasocial breakups, Cohen, 2004). Theoretically, if people would have a strong parasocial relationship with a mediated persona, they should also feel something of a close relationship to the persona. Indeed, Horton and Wohl (1956) referred to a parasocial relationship as intimacy at a distance, and intimacy is often synonymous with relationship closeness (Dibble, Levine, & Park, 2012; Parks & Floyd, 1996). Thus, we reasoned that measures of relationship closeness might serve as loose substitute measures of parasocial relationship.

Westen and Rosenthal (2003) provided an initial means for quantifying CV by way of nomological network. The $r_{\text{alerting-CV}}$ is an effect size correlation between (a) and (b), where (a) is the pattern of correlations predicted between the test construct and the network variables, and (b) is the pattern of correlations that is actually obtained. Because the patterns are theorized to be linear, $r_{\text{alerting-CV}}$ is interpreted simply as Pearson's r. A statistically significant $r_{\text{alerting-CV}}$ provides a preliminary alert to potential trends of interest and, in this way, supports CV.

This research afforded us an opportunity to utilize Westen and Rosenthal's (2003) procedure, as one of our aims is to subject the EPSI-Scale (Hartmann & Goldhoorn, 2011) to nomological network validation. The $r_{\text{alerting-CV}}$ technique is concerned more with the relative magnitudes in the pattern of correlations than with absolute values of the correlations. As noted previously, the EPSI-Scale is predicated on the conception of parasocial interaction as a sense of mutual awareness with the media performer that occurs during viewing (brought on by bodily and verbally addressing the viewer through the camera). By contrast, parasocial relationship refers to a more enduring sense of liking or long-term social involvement that can start to form during viewing, but that ultimately transcends the viewing episode.

When inspecting the operational definitions implied by how each scales' items were worded, we reasoned the EPSI-Scale would correlate most strongly with measures that apply explicitly to in-the-moment viewing (PSI-Process Scales; Schramm & Hartmann, 2008), followed by measures of parasocial relationship (PSI-Scale, Rubin et al., 1985; PSR Scales, Tukachinsky, 2010), and would correlate least with measures that tap other concepts like identification (Auter & Palmgreen, 2000; Tal-Or & Cohen, 2010); and relationship closeness (Unidimensional Relationship Closeness Scale [URCS], Dibble et al., 2012; Inclusion of Other in the Self Scale [IOS], Aron, Aron, & Smollan, 1992; Relationship Closeness Inventory [RCI], Berscheid, Snyder, & Omoto, 1989). This is because concepts like identification and closeness generally

require the perceiver to know the target persona to some extent, which suggests these concepts will not be as salient during a brief first encounter, hence, the correlations with the EPSI-Scale would be lower.

To fine-tune our nomological net, we further expected that the EPSI-Scale (a measure of parasocial interaction) will correlate more strongly with the PSI-Scale (a measure of parasocial relationship) than with the closeness measures. If relationship closeness, or something like it, is part of a parasocial relationship, as we suggested earlier, how can this be so? When looking at the operational definitions implied by the exact item wording of each scale, we realize that PSI-Scale items often refer to plausible future relationship scenarios, for example, "I *would* like to meet X in person," or, "I think X *is like* an old friend." Because of this wording, and because it is reasonable to *want* to meet an individual after only a brief first encounter, respondents might agree to some extent with these items. That is, even after a short video exposure, respondents can register a *desire* to meet the performer, which would drive some agreement with the PSI-Scale items and promote a positive correlation with the EPSI-Scale. By contrast, the applied relationship closeness measures appeared to tap the strength of the bond with the performer that was *actually felt* by the respondent, for example, "My relationship with X *is* close." (Such wording probably drives the similarities between closeness and parasocial relationship in the first place). This bond is not likely to be strong after a brief video encounter, even if a desire to meet the performer is present. Thus, we reasoned that the EPSI-Scale should correlate more strongly with the PSI-Scale than with the closeness measures, even if the two latter can be taken as alternate measures of parasocial relationship.

As a starting point, we based our predictions on Hartmann and Goldhoorn's (2011) prior work on the EPSI-Scale. For example, the EPSI-Scale correlated with the PSI-Scale (Rubin & Perse, 1987) at $r = .43$, and with the PSI-Process Scales (Schramm & Hartmann, 2008) at $r = .48$. Therefore, we used .40 and .50, respectively, as starting points for this research and adjusted the magnitudes of the other predictions relative to these, according to the aforementioned basic pattern (see Table 3). Cohen (1988) considered r values of .10, .30, and .50 to be small, medium, and large, respectively. If the obtained nomological network correlations correspond with our predictions, then $r_{\text{alerting-cv}}$ should be significant and large ($r > .50$). These predictions reflect the following hypothesis:

> H2: The pattern of predicted correlations between the EPSI-Scale and the nomological network variables will correlate with the pattern of obtained correlations such that $r_{\text{alerting-CV}}$ will be significant and large.

The CV of a measure also hinges, in part, on the extent to which the measure correlates more strongly with alternate measures of the same construct (convergent validity), than with measures of different constructs (discriminant validity). If the PSI-Scale (Rubin et al., 1985) better reflects parasocial relationship than parasocial interaction, then we expect the PSI-Scale to correlate more highly with other measures designed to measure parasocial relationship, for example, Tukachinsky's (2010)

parasocial relationship scales, than it does with the EPSI-Scale (Hartmann & Gold-hoorn, 2011).

> H3: The PSI-Scale will correlate more highly with Tukachinsky's (2010) PSR Scales than with the EPSI-Scale.

Method

We conducted a basic independent-groups experiment to test the performance of measures of parasocial interaction and parasocial relationship, as well as to perform further construct validation testing for the EPSI-Scale (Hartmann & Goldhoorn, 2011). Participants were exposed to a brief video where the media performer either bodily addressed the participant or not, prior to answering an online questionnaire. The research was internal review board approved.

Participants

Participants ($N = 446$) were recruited using the Mechanical Turk (MTurk) feature of Amazon.com, and were each paid $0.25 (United States) for their participation. From this original data set, we dropped 46 cases for failing to complete the survey and 1 due to technical problems with viewing and/or hearing the video. We reasoned further that anyone staying on the video page for less than the run time of the video (41 seconds) may not have paid sufficient attention to the video, and that prolonged exposure (>5 minutes) to the video might encourage the development of a parasocial relationship and thus contaminate our induction of parasocial interaction. Therefore, we dropped 10 cases for the former and 6 cases for the latter. In total we dropped 63 cases, resulting in a final sample of $N = 383$ valid cases.

The sample was 52% female ($n = 200$) and 48% male ($n = 183$), ranging in age from 18 to 70 years ($M = 34.74$, $SD = 11.73$). The majority of the participants were Caucasian ($n = 301$, 79%), and the remaining participants were African American ($n = 27$, 7%), Asian ($n = 22$, 6%), Hispanic ($n = 20$, 5%), Native American ($n = 3$, 1%), Pacific Islander ($n = 3$, 1%), and other ($n = 7$, 2%). Participants reported a variety of careers with management (13%), sales (12%), student (15%), and unemployed (11%) being the most popular. Less than 2% of the participants did not graduate from high school, 10% had a high school diploma, 42% had some college, 33% had a 4-year college degree, and 13% reported graduate and professional education. About 46 of 50 U.S. states were represented in the sample, with California representing the greatest proportion of participants (14%).

Our study was listed among the Human Interest Tasks (HITs) available for registered users of MTurk ("workers") to complete in return for pay. Research shows samples from MTurk to be much more diverse than typical American college samples (Buhrmester, Kwang, & Gosling, 2011), more representative of U.S. populations than convenience samples (Berinsky, Huber, & Lenz, 2012), and the conclusions generated are quite comparable to samples generated through other means (e.g., Mason & Suri, 2012). MTurk enables the restriction of samples according to various

demographic variables. We required that participants be 18 years of age and reside within the United States (according to the address the worker self-reports when registering through Amazon.com). After selecting our HIT and reading a brief description of our study and some consent information, workers were directed to the survey that was hosted on Qualtrics.com. Upon successful completion of the survey, and in accordance with MTurk protocol, we credited $0.25 (United States) to each worker's MTurk account.

Procedure

First, using an experimental procedure very similar to that of Hartmann and Goldhoorn (2011), participants viewed a 41-second video clip of a college-aged, Caucasian woman named "Amy." The video was designed to emulate a video introduction Amy might use for an online dating profile. Thus, Amy discussed the qualities of her ideal romantic partner. Participants were randomly assigned to view one of two versions of this video (see Figure 1). In the first condition ($n = 194$, 51%), Amy looked into the camera and oriented her body so as to bodily address the viewer. This is in line with research suggesting that direct bodily addressing and eye-gaze facilitate parasocial interaction (Auter & Davis, 1991; Cummins & Cui, 2014; Hartmann & Goldhoorn, 2011; Horton & Wohl, 1956; Mancini, 1988).

Figure 1 Illustration of the manipulation of media performer's bodily addressing style. Top panel = bodily addressing and looking into the camera; bottom panel = no bodily addressing or looking into the camera.

In the second condition ($n = 189$, 49%), the video showed Amy from a side profile. When creating the video clips, two cameras were used to capture simultaneously two different angles of the same performance. Thus, the only difference between the two conditions was the camera angle (bodily addressing vs. side profile). After participants viewed the video, they answered the questionnaire items that followed. All survey items were worded so that the participant would report on his or her reactions to Amy.

Measures
All measures took the form of Likert-type summated ratings scales whose response sets vary by construct and are described under each construct. In all cases, higher scores reflect greater magnitudes of the construct. Internal consistency was assessed using Cronbach's (1951) alpha. All means, standard deviations, and alphas appear in Table 2.

PSI-Scale
Parasocial phenomena were assessed in multiple ways. First, we employed Rubin et al.'s (1985) classic PSI-Scale. The PSI-Scale consists of 20 Likert items to which responses range from 1 (*strongly disagree*) to 5 (*strongly agree*). The original PSI-Scale was oriented to capture reactions toward television newscasters. Therefore, we adjusted the wording to reflect the video clips used for this study. In addition, we eliminated five items from the 20-item scale because they did not apply to the experimental manipulation central to this study. We analyzed the remaining PSI-Scale items according to both their long (15 items; Rubin et al., 1985) and short forms (8 items; Rubin & Perse, 1987), less each version's nonapplicable items. Table 3 lists the items for both forms of the PSI-Scale.

EPSI-Scale
We also measured parasocial phenomena using the EPSI-Scale (Hartmann & Goldhoorn, 2011). The EPSI-Scale was designed to measure the experience of parasocial interaction as we have presented it in this article, and consists of six Likert items with a response set anchored by 1 (*do not agree at all*) and 5 (*totally agree*). Items appear in Table 3.

PSI-Process Scales
We used shortened, English-language versions of the PSI-Process Scales (Hartmann & Goldhoorn, 2011; Schramm & Hartmann, 2008) to tap specific cognitive (e.g., "I carefully followed the behavior of Amy"), affective (e.g., "If Amy felt bad, I felt bad as well; if Amy felt good, I felt good as well"), and behavioral (e.g., "Sometimes I felt like speaking out on Amy") aspects of parasocial reactions that occur while viewing. The cognitive, affective, and behavioral components were assessed with six, three, and two Likert items, respectively. Two affective items and one behavioral item were dropped to improve internal consistency. All items feature a response set anchored by 1 (*not at all*) and 5 (*very much*).

Tukachinsky's (2010) PSR Scales

We measured parasocial relationship using Tukachinsky's (2010) Multiple-PSR Scales. These scales tap four dimensions: parasocial friendship-communication (five items, e.g., "Sometimes I wish I could ask Amy for advice"); parasocial friendship-support (seven items, e.g., "I could have a warm relationship with Amy"); parasocial love-physical attraction (four items, e.g., "I find Amy very attractive physically"); and parasocial love-emotional response (seven items, e.g., "For me, Amy could be the perfect romantic partner"). All items featured a Likert-type response set anchored by 1 (*strongly disagree*) and 5 (*strongly agree*). We analyzed each dimension separately.

Relationship closeness

To the extent that a parasocial relationship resembles a real-life social relationship, viewers can be more or less "para-close" to the media performer. Therefore, we included three different measures of relational closeness. First, we utilized the URCS (Dibble et al., 2012). The URCS consists of 12 Likert items (e.g., "My relationship with Amy is close") to which responses were anchored by 1 (*strongly disagree*) and 5 (*strongly agree*).

Second, we utilized Aron et al.'s (1992) IOS. The IOS measures self-concept overlap, and consists of a single item in the form of seven pairs of circles that overlap to varying degrees (1 = *no overlap*, 7 = *almost complete overlap*). Respondents simply choose the pair of circles that best describes their relationship with their imagined other.

Finally, we also measured relationship closeness using the first item of Berscheid et al.'s (1989) RCI, worded to reflect Amy from our experimental induction: "Relative to all your other relationships (both same and opposite sex), how would you characterize your relationship with Amy?" Respondents indicated agreement on a 7-point Likert scale anchored by 1 (*not at all close*) and 7 (*extremely close*).

Identification

Identification was measured using the six items of the identification factor taken from Auter and Palmgreen's (2000) API scale, for example, "I can identify with Amy." All items featured a 5-point Likert response set anchored by 1 (*strongly disagree*) and 5 (*strongly agree*).

We also measured identification using five items from Tal-Or and Cohen (2010), for example, "I think I understood Amy well." Responses were on a 7-point Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

Wishful identification

Wishful identification was measured using five items from Hoffner and Buchanan (2005). Sample items include, "I'd like to do the kinds of things Amy does on the show," "Amy is the sort of person I want to be like myself," and "I wish I could be more like Amy." Responses were on a 5-point Likert scale ranging from 1 (*disagree a lot*) to 5 (*agree a lot*).

Results

Given that we utilized only one female performer in the video clips, we wanted to determine whether gender was distributed evenly between the two experimental conditions. That is, with respect to gender, did we violate the assumption of initially equivalent groups? To test this assumption, we conducted a chi-square analysis to test the null hypothesis that gender and experimental condition were not dependent. This chi-square value was not significant, $\chi^2 (1, n = 383) = 1.18$, $p = .28$, $\phi = .06$, suggesting that gender composition was not significantly different across the experimental conditions, and that any within-group variation owing to gender would not confound our experimental induction.

Our first hypothesis (H1) predicted that the EPSI-Scale would be more sensitive to the experimental induction than would the PSI-Scales. Fisher's r to Z tests comparing differences between correlations appear in Table 1. As the table shows, the EPSI-Scale reflected more variance across experimental conditions ($r = .20$) than did the PSI-8 ($r = .07$), but not the PSI-15 ($r = .10$). Therefore, the data were partially consistent with H1 in that an experimental induction of parasocial interaction (by way of bodily addressing) instantiated greater variance on the EPSI-Scale (a measure carefully designed to tap parasocial interaction) than on the PSI-Scale short form, but not the PSI-Scale long form. None of the other scales responded to our experimental induction (all rs n.s., see Table 1).

Our research question (RQ1) asked, at the item level, how indicators purported to tap parasocial interaction compare with indicators purported to tap parasocial relationship. Although the effect sizes were not large, every item of the EPSI-Scale reliably discriminated between the experimental conditions (directly addressing vs. side profile), $p < .05$. In contrast to this, all but 3 of the 15 PSI-Scale items were not sensitive to the experimental induction (see Table 2). Thus, despite outperforming its short form counterpart in the test of H1, the PSI-Scale long form still did not seem to tap the same construct as the EPSI-Sale.

Our second hypothesis (H2) attempted to quantify the correspondence between the predicted and observed correlations in a more refined way, as part of nomological net construct validation. An $r_{\text{alerting-CV}}$ that is significant and large indicates a correspondence between the nomological net predictions and the pattern of obtained correlations. According to conventions outlined by Cohen (1988), $r \geq .50$ is considered large. As seen in Table 3, $r_{\text{alerting-CV}} = .64$, which suggests that the obtained correlations generally trended in their predicted ways, thus enabling greater faith in the CV of the EPSI-Scale. Because our predictions were based on the EPSI-Scale as being sensitive to bodily and verbally addressing during viewing (parasocial interaction), these results support the EPSI-Scale as measuring parasocial interaction as we have conceptualized it.

Finally, if the PSI-Scale were actually measuring parasocial interaction, then it should correlate highly with an alternate measure constructed to tap parasocial interaction (e.g., the EPSI-Scale), and less so with measures constructed to tap parasocial

Table 1 Zero-Order Correlations and Descriptive Statistics

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. Condition (1 = direct)															
2. EPSI	.20**														
3. PSI-15 item	.10*	.54**													
4. PSI-8 item	.07	.44**	.96**												
5. PSI process	.08	.50**	.73**	.65**											
6. PSR communication	.05	.43**	.75**	.73**	.68**										
7. PSR support	.06	.40**	.71**	.71**	.59**	.84**									
8. PSR attraction	.02	.20**	.56**	.62**	.43**	.59**	.62**								
9. PSR emotion	.07	.42**	.71**	.68**	.68**	.83**	.71**	.62**							
10. URCS	.01	.38**	.54**	.49**	.53**	.62**	.50**	.39**	.76**						
11. Closeness IOS	.02	.30**	.50**	.44**	.43**	.50**	.41**	.34**	.61**	.74**					
12. Closeness RCI	−.01	.37**	.56**	.50**	.49**	.57**	.48**	.37**	.65**	.81**	.79**				
13. Identify Auter	.003	.34**	.58**	.55**	.51**	.68**	.61**	.39**	.56**	.49**	.39**	.48**			
14. Identify Tal-Or	.04	.40**	.63**	.59**	.56**	.66**	.65**	.44**	.56**	.47**	.41**	.46**	.70**		
15. Wishful identify	−.02	.31**	.63**	.60**	.50**	.67**	.59**	.47**	.54**	.47**	.42**	.45**	.71**	.69**	
M		2.28	2.67	2.88	2.46	2.51	2.84	3.03	1.98	1.69	1.58	1.87	2.58	3.65	2.69
SD		1.10	0.79	0.88	0.71	0.92	0.94	1.04	0.96	0.94	1.25	1.45	0.93	1.48	0.90
α		.92	.92	.91	.84	.88	.91	.90	.94	.98			.91	.91	.86

Note: $N = 383$, All correlation differences $(r2 − r1) \geq .12$ were significantly different according to Fisher's Z-transformation test at $p < .05$. EPSI = experience of parasocial interaction; IOS = Other in the Self Scale; PSI = parasocial interaction; PSR = parasocial relationship; RCI = Relationship Closeness Inventory; URCS = Unidimensional Relationship Closeness Scale.
$*p < .05$, two-tailed. $**p < .01$, two-tailed.

Table 2 Per Item Comparisons by Experimental Condition for the Experience of Parasocial Interaction Scale (EPSI-Scale) and the Parasocial Interaction Scale (PSI-Scale)

Item	Indirect		Direct			
	M	SD	M	SD	t	r
EPSI	2.06	1.07	2.49	1.09	3.92**	.20
While watching the video clip, Amy						
1. … was aware of me.	1.88	1.16	2.39	1.24	4.13**	.21
2. … knew I was there.	1.90	1.19	2.51	1.33	4.70**	.23
3. … knew I was aware of her.	2.49	1.47	2.88	1.40	2.70**	.14
4. … knew I paid attention to her.	2.43	1.41	2.92	1.40	3.40**	.17
5. … knew that I reacted to her.	2.05	1.23	2.40	1.28	2.70**	.14
6. … reacted to what I said or did.	1.58	1.01	1.83	1.12	2.31*	.12
PSI-15 item	2.58	0.82	2.75	0.76	2.02*	.11
1. The video clip showed me what Amy is like.	3.28	1.12	3.47	0.98	1.86	.09
2. While watching the video clip, I felt as if I were part of a group.	2.03	1.15	2.16	1.14	1.15	.06
3. I liked to compare my ideas with what Amy said in the video clip.	2.60	1.19	2.78	1.15	1.53	.08
4. Amy made me feel comfortable, as if I was with a friend.[+]	2.82	1.21	3.11	1.07	2.47*	.13
5. I see Amy as a natural, down-to-earth person.[+]	3.67	0.98	3.78	0.92	1.20	.06
6. I like hearing the voice of Amy in my home.	2.69	1.21	2.84	1.13	1.21	.06
7. Amy kept me company while the video clip was running.	2.04	1.18	2.42	1.20	3.13**	.16
8. I look forward to watching Amy in another video clip.[+]	2.56	1.25	2.59	1.18	0.22	.01
9. If Amy would appear in another video clip, I would watch that clip.[+]	2.76	1.20	2.92	1.17	1.37	.07
10. I sometimes made remarks to Amy during the video clip.	1.54	1.01	1.66	1.06	1.16	.06
11. If there were a story about Amy in a newspaper or magazine, I would read it.[+]	2.61	1.22	2.72	1.09	0.94	.05
12. I would miss Amy when she was on vacation.[+]	2.00	1.14	2.07	1.08	0.64	.03
13. I would like to meet Amy in person.[+]	2.79	1.27	2.93	1.17	1.17	.06
14. I think Amy is like an old friend.	2.03	1.11	2.29	1.21	2.18 *	.11
15. I find Amy to be attractive.[+]	3.34	1.09	3.45	1.00	0.95	.05

Notes: Independent samples *t*-tests were conducted to compare items on Hartmann and Gold-hoorn's (2011) EPSI-Scale with Rubin et al.'s (1985) PSI-Scale by experimental condition. Items with the + superscript are items on the 8-item PSI-Scale.
*$p < .05$. **$p < .01$.

Table 3 Predicted and Observed Correlations Between Experience of Parasocial Interaction (EPSI) and Nomological Network Variables

Measure	Predict r	Predict Z_1	Observe r	Observe Z_1 $r_{\text{alerting-CV}} = .64^{\star}$
PSI-15 item	.40	.42	.54	.60
PSI process	.50	.55	.50	.55
PSR communication	.30	.31	.43	.46
PSR support	.30	.31	.40	.42
PSR attraction	.30	.31	.20	.20
PSR emotion	.30	.31	.42	.45
URCS	.20	.20	.38	.40
Closeness I	.20	.20	.30	.31
Closeness II	.20	.20	.37	.39
Identify Auter	.20	.20	.34	.35
Identify Tal-Or	.20	.20	.40	.42
Wishful identify	.18	.18	.31	.32

Note: PSI = parasocial interaction; PSR = parasocial relationship; URCS = Unidimensional Relationship Closeness Scale.

$^{\star}p < .05$.

relationship. However, to observe the PSI-Scale correlating less strongly with the EPSI-Scale, and more strongly with explicit parasocial relationship indicators (e.g., Tukachinsky's (2010) PSR Scales) would suggest that the PSI-Scale better reflects something akin to parasocial relationship and less parasocial interaction (H3).

Our data were consistent with this hypothesis in that the correlations between the EPSI-Scale and the PSI-Scale in either its short or long form were $r = .44$ and $r = .54$, respectively (see Table 2). By contrast, the PSI-Scale correlated more strongly with Tukachinsky's (2010) PSR dimensions ($rs = .56 - .75$). Moreover, the closeness measures correlated higher with the PSI-Scale ($rs = .44 - .56$) than with the EPSI-Scale ($rs = .30 - .38$); and the identification and wishful identification scales correlated higher with the PSI-Scale ($rs = .55 - .63$) than with the EPSI-Scale ($rs = .31 - .40$). Taken together, these data suggest that the PSI-Scale does not assess the same construct as the EPSI-Scale, and that the PSI-Scale corresponds better with measures that tap aspects of a parasocial (and social) relationship.

Discussion

The broad purpose of this research was to improve the measurement of parasocial interaction. Toward that end, we reviewed the seminal and subsequent literature to define parasocial interaction as an immediate experience whereby the media viewer feels as if he or she is in a reciprocal social encounter with the media performer (Hartmann & Goldhoorn, 2011). This experience is especially triggered when the

media performer addresses the viewer via body orientation and looking into the camera. Unfortunately, the concept of parasocial interaction has acquired additional meaning over the decades, having become conflated with the more enduring parasocial relationship. While parasocial interaction occurs during viewing, a parasocial relationship can transcend the immediate viewing exposure and more resembles a sense of liking for, or short- or long-term involvement with, the media performer. Consequently, the measures of parasocial interaction leave much to be desired in terms of what they are actually measuring.

To establish a "purer" measure of parasocial interaction as a perception of reciprocal awareness, attention, and adjustment with a media performer, Hartmann and Goldhoorn (2011) developed the EPSI-Scale. We interpret our data to suggest that the EPSI-Scale may be a more suitable measure of parasocial interaction than the PSI-Scale, particularly the short version of the PSI-Scale. According to Horton and Wohl (1956), attention to how the media persona bodily and verbally addresses the viewer is key to instantiating parasocial interaction. The EPSI-Scale showed significantly more variance than did the short PSI-Scale (albeit not significantly more than the long PSI-Scale), which partially supported H1. Perhaps more importantly, whereas all six items of the EPSI-Scale significantly differed in response to our experimental induction of bodily addressing, only 3 of 15 (long form) and 1 of 8 (short form) items of the PSI-Scales showed significant sensitivities to the induction (RQ1).

The EPSI-Scale also performed according to expectations in a nomological network, indexed quantitatively by calculating the validity coefficient known as $r_{\text{alerting-CV}}$ (Westen & Rosenthal, 2003; H2). Given that our theoretical predictions were based on the EPSI-Scale as measuring mutual awareness, attention, and adjustment to the media performer, the significant and large value for $r_{\text{alerting-CV}}$ suggests the observed pattern of correlations was consistent with theory and lends additional credence to the EPSI-Scale's CV. In contrast, we found the PSI-Scales to correlate higher with measures that tap the more enduring parasocial relationship (e.g., PSR Scales, Tukachinsky, 2010), and with social (real life, not para) relationship-oriented concepts such as closeness and identification, than did the EPSI-Scale (H3). The EPSI-Scale correlated with the PSI-Scale at $r = .54$, further suggesting that these two measures are not assessing the same construct.

Taken together, this preliminary evidence suggests that the EPSI-Scale may be more suitable to assess parasocial interaction, whereas the PSI-Scale (particularly the frequently applied short version), despite its name, may rather tap into users' short- or long-term liking of media characters. To the extent this assumption receives further support, some of the findings of past studies employing the popular PSI-Scale may need to be reinterpreted as informing about the role of users' liking of media characters, but not of parasocial interaction understood as users' illusory experience of participating in a social interaction with a media performer (see also Rosaen & Dibble, 2012).

From our perspective, future research should further explore the empirical and theoretical implications of the PSI-Scales. Based on the present approach and these

preliminary findings, and granting that additional research is necessary, we are doubt-ful that either version of the PSI-Scale measures parasocial interaction as we believe it ought to be conceptualized (separate from the more enduring parasocial relation-ship). Accordingly, we caution researchers to think carefully about their research goal and the questions they are trying to answer before mindlessly employing any of the existing scales. As we see it, for projects that require sensitivity to the illusory "con-versational give-and-take" that comes during viewing (Horton & Wohl, 1956, p. 215), existing evidence favors applying the EPSI-Scale. As other researchers have found when revisiting commonly used self-report measures (e.g., Kotowski, Levine, Baker, & Bolt, 2009; Levine & Kotowski, 2010), the name given a scale does not always corre-spond to the construct being measured by that scale. Because the conclusions drawn can be no more valid than the measures used to derive those conclusions, we remind researchers to stay mindful about their measures.

Limitations and conclusion

As with all research, this study is not without limitations. Most obviously, we acknowl-edge that self-report data feature their own limitations, and that these are well known (e.g., Fredrickson, 2000; Fredrickson & Kahneman, 1993; Kahneman, 1999; Ross & Simonson, 1991; Schwarz, 1999). However, to the extent that such limitations apply, they should apply to all of our study's measures equally. Moreover, despite such limita-tions, we observed many differences between the performances of the EPSI-Scale and the PSI-Scales, and these differences favored the EPSI-Scale over the PSI-Scales for measuring parasocial interaction. Nevertheless, we agree with Hartmann and Gold-hoorn (2011) that additional methodologies be employed (e.g., thinking aloud while viewing, behavioral measures) to cross-examine the findings from the EPSI-Scale.

Second, we cannot ascertain the generality of our findings. However, in contrast to much social scientific research, we did utilize a noncollege student sample that drew from most of the U.S. states and included a range of occupations. Although studies to date are favorable toward the conclusions drawn from samples gathered through Amazon's Mechanical Turk, there simply has not been enough research with which to compare the generality. Our results were largely consistent with those from a noncollege student sample from the Netherlands (Hartmann & Goldhoorn, 2011), however, we look to future research to determine the extent to which our findings generalize.

Third, we recognize that our experiment featured only one video performer, Amy. This is a limitation for at least two reasons. First, Amy was female. Consequently, male and female participants could respond differently to Amy. For example, men more than women might have deemed Amy to be very beautiful, which might differentially influence their desires to meet Amy. However, our study employed an experimental design that randomly assigned participants either to be bodily addressed by Amy or not. Thus, any such effects would have been equivalent in both experimental condi-tions, and they could not explain our core findings related to between-group differ-ences. Although gender was not the focus of our study, we do believe the effects of

gender on media responses are interesting, and we look to future research to take up this issue directly.

Second, using a single stimulus as we did invites the language-as-fixed-effect fallacy (see Jackson & Jacobs, 1983), if Amy is taken to represent a larger class of media performers. To overcome this fallacy and enhance generalizability, future studies should attempt to replicate these findings by simultaneously applying a variety of media performers when manipulating bodily addressing. At present, we hoped to instantiate an effect of our induction for purposes of discriminating between the PSI-Scales and the EPSI-Scale. Thus, we prioritized control of extraneous factors. We look to future research to determine the extent to which our findings generalize.

Finally, we acknowledge the relatively small effect sizes registered by the EPSI-Scale items in response to our experimental induction. Our induction featured what might be only a modest variation of bodily addressing (camera captures performer looking straight ahead vs. camera captures side profile of performer). Future research should strive for stronger effects by applying equally plausible, but hopefully more powerful, experimental manipulations. For example, comparing bodily addressing (e.g., performer looks into camera lens) to zero bodily addressing (e.g., performer has back turned to the camera) might result in stronger inductions, hence, larger effects.

In summary, we believe it is useful to conceptualize parasocial interaction as a viewer's sense of mutual awareness, attention, and adjustment to a media performer that occurs during viewing, and to distinguish parasocial interaction from the more enduring parasocial relationship. This definition of parasocial interaction is consistent with Horton and Wohl's (1956) original theorizing, and we have shown that the scale most commonly used to assess parasocial interaction, the PSI-Scale (Rubin et al., 1985), does not perform well given our understanding of parasocial interaction. At present, we recommend the EPSI-Scale (Hartmann & Goldhoorn, 2011) in situations where parasocial interaction is prudent. Additional research is necessary to continue to validate the EPSI-Scale, as well as to determine the future of the PSI-Scale.

References

Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, **63**, 596–612. doi:10.1037/0022-3514.63.4.596

Auter, P. J. (1992). Psychometric: TV that talks back: An experimental validation of a parasocial interaction scale. *Journal of Broadcasting & Electronic Media*, **36**, 173–181. doi:10.1080/08838159209364165

Auter, P. J., & Davis, D. M. (1991). When characters speak directly to viewers: Breaking the fourth wall in television. *Journalism Quarterly*, **68**, 165–171. doi:10.1177/107769909106800117

Auter, P. J., & Palmgreen, P. (2000). Development and validation of a parasocial interaction measure: The audience-persona interaction scale. *Communication Research Reports*, **17**, 79–89. doi:10.1080/08824090009388753

Babrow, A. S. (1987). Student motives for watching soap operas. *Journal of Broadcasting & Electronic Media*, **31**, 471–487. doi:10.1080/08838158709386666

Beniger, J. R. (1987). Personalization of mass media and the growth of pseudo-community. *Communication Research*, **14**, 352–371. doi:10.1177/009365087014003005

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis*, **20**, 351–368. doi:10.1093/pan/mpr057

Berscheid, E., Snyder, M., & Omoto, A. M. (1989). The Relationship Closeness Inventory: Assessing the closeness of interpersonal relationships. *Journal of Personality and Social Psychology*, **57**, 792–807. doi:10.1037/0022-3514.57.5.792

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality data? *Perspectives on Psychological Science*, **6**, 3–5. doi:10.1177/1745691610393980

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (2004). Parasocial break-up from favorite television characters: The role of attachment styles and relationship intensity. *Journal of Social and Personal Relationships*, **21**, 187–202. doi:10.1177/0265407504041374

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, **16**, 297–334.

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, **56**, 81–105. doi:10.1037/h0040957

Cummins, R. G., & Cui, B. (2014). Reconceptualizing address in television programming: The effect of address and affective empathy on viewer experience of parasocial interaction. *Journal of Communication*, **64**, 723–742. doi:10.1111/jcom.12076

Dibble, J. L., Levine, T. R., & Park, H. S. (2012). The Unidimensional Relationship Closeness Scale (URCS): Reliability and validity evidence for a new measure of relationship closeness. *Psychological Assessment*, **24**, 565–572. doi:10.1037/a0026265

Dibble, J. L., & Rosaen, S. F. (2011). Parasocial interaction as more than friendship: Evidence for parasocial interactions with disliked media figures. *Journal of Media Psychology*, **23**, 122–132. doi:10.1027/1864-1105/a000044

Ellis, G. J., Streeter, S. K., & Engelbrecht, J. D. (1983). Television characters as significant others and the process of vicarious role taking. *Journal of Family Issues*, **4**, 367–384. doi:10.1177/019251383004002007

Finn, S., & Gorr, M. B. (1988). Social isolation and social support as correlates of television viewing motivations. *Communication Research*, **15**, 135–158. doi:10.1177/009365088015002002

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and Emotion*, **14**, 577–606. doi:10.1080/026999300402808

Fredrickson, B. L., & Kahneman, D. (1993). Duration neglect in retrospective evaluations of affective episodes. *Journal of Personality and Social Psychology*, **65**, 45–55. doi:10.1037/0022-3514.65.1.45

Giles, D. (2002). Parasocial interaction: A review of the literature and a model for future research. *Media Psychology*, **4**, 279–305. doi:10.1207/S1532785XMEP0403_04

Gusfield, J. R. (2003). A journey with symbolic interactionism. *Symbolic Interaction*, **26**, 119–139.

Hartmann, T., & Goldhoorn, C. (2011). Horton and Wohl revisited: Exploring viewers' experience of parasocial interaction. *Journal of Communication*, **61**, 1104–1121. doi:10.1111/j.1460-2466.2011.01595.x

Hoffner, C., & Buchanan, M. (2005). Young adults' wishful identification with television characters: The role of perceived similarity and character attributes. *Media Psychology*, **7**, 325–351. doi:10.1207/S1532785XMEP0704_2

Horton, D., & Strauss, A. L. (1957). Interaction in audience-participation shows. *The American Journal of Sociology*, **62**, 579–587.

Horton, D., & Wohl, R. (1956). Mass communication and para-social interaction: Observations on intimacy at a distance. *Psychiatry*, **19**, 215–229.

Houlberg, R. (1984). Local television news audience and the para-social interaction. *Journal of Broadcasting*, **28**, 423–429. doi:10.1080/08838158409386551

Jackson, S., & Jacobs, S. (1983). Generalizing about messages: Suggestions for design and analysis of experiments. *Human Communication Research*, **9**, 169–191. doi:10.1111/j.1468-2958.1983.tb00691.x

Kahneman, D. (1999). Objective happiness. In D. Kahneman, E. Diener & N. Schwarz (Eds.), *Well-being: The foundations of hedonic psychology* (pp. 3–25). New York, NY: Sage.

Klimmt, C., Hartmann, T., & Schramm, H. (2006). Parasocial interactions and relationships. In J. Bryant & P. Vorderer (Eds.), *Psychology of entertainment* (pp. S291–S313). Mahwah, NJ: Lawrence Erlbaum.

Kotowski, M. R., Levine, T. R., Baker, C. R., & Bolt, J. M. (2009). A multitrait-multimethod validity assessment of the verbal aggressiveness and argumentativeness scales. *Communication Monographs*, **76**, 443–462. doi:10.1080/03637750903300247

Lang, K., & Lang, G. E. (1953). The unique perspective of television and its effect: A pilot study. *American Sociological Review*, **18**, 3–12.

Levine, T. R., & Kotowski, M. R. (2010). Measuring argumentativeness and verbal aggressiveness: Psychometric concerns and advances. In T. A. Avtgis & A. S. Rancer (Eds.), *Arguments, aggression, and conflict: New directions in theory and research* (pp. 67–81). New York, NY: Taylor & Francis.

Levy, M. R. (1979). Watching TV news as para-social interaction. *Journal of Broadcasting*, **23**, 177–187.

Levy, M., & Windahl, S. (1984). Audience activity and gratifications: A conceptual clarification and exploration. *Communication Research*, **11**, 51–78. doi:10.1177/009365084011001003

Mancini, P. (1988). Simulated interaction: How the television journalist speaks. *European Journal of Communication*, **3**, 151–166. doi:10.1177/0267323188003002003

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavioral Research Methods*, **44**, 1–23. doi:10.3758/s13428-011-0124-6

Nordlund, J. E. (1978). Media interaction. *Communication Research*, **5**, 150–175. doi:10.1177/009365027800500202

Parks, M. R., & Floyd, K. (1996). Meanings for closeness and intimacy. *Journal of Social and Personal Relationships*, **13**, 85–107. doi:10.1177/0265407596131005

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.

Peters, J. D., & Simonson, P. (Eds.) (2004). *Mass communication and American social thought: Key texts, 1919–1968*. Oxford, England: Rowman & Littlefield.

Rosaen, S. F., & Dibble, J. L. (2012, November). *Parasocial relationships with liked and disliked characters*: *Does the need for companionship*, *attachment style*, *and the gender of the character predict strength*? Paper presented at the 98th Annual Convention of the National Communication Association, Orlando, FL.

Rosengren, K. E., Windahl, S., Hakansson, P. A., & Johnsson-Smaragdi, U. (1976). Adolescents' TV relations: Three scales. *Communication Research*, **3**, 347–366. doi:10.1177/009365027600300401

Ross, W. T., & Simonson, I. (1991). Evaluations of pairs of experiences: A preference for happy endings. *Journal of Behavioral Decision Making*, **4**, 273–282.

Rubin, A. M., & Perse, E. M. (1987). Audience activity and soap opera involvement: A uses and effects investigation. *Human Communication Research*, **14**, 246–268. doi:10.1111/j.1468-2958.1987.tb00129.x

Rubin, A. M., Perse, E. M., & Powell, R. A. (1985). Loneliness, parasocial interaction, and local television news viewing. *Human Communication Research*, **12**, 155–180. doi:10.1111/j.1468-2958.1985.tb00071.x

Schiappa, E., Allen, M., & Gregg, P. B. (2007). Parasocial relationships and television: A meta-analysis of the effects. In R. Preiss, B. Gayle, N. Burrell, M. Allen & J. Bryant (Eds.), Mass media effects*:* Advances through meta-analysis *(S* (pp. 301–314). Mahwah, NJ: Lawrence Erlbaum.

Schramm, H., & Hartmann, T. (2008). The PSI-Process Scales: A new measure to assess the intensity and breadth of parasocial processes. *Communications*, **33**, 385–401. doi:10.1515/COMM.2008.025

Schramm, H., & Wirth, W. (2010). Testing a universal tool for measuring parasocial interactions cross different situations and media. *Journal of Media Psychology*, **22**, 26–36. doi:10.1027/1864-1105/a000004

Schwarz, N. (1999). Self-reports: How the questions shape the answers. *American Psychologist*, **54**, 93–105. doi:10.1037/0003-066X.54.2.93

Strauss, A. L. (1958). In memoriam: R. Richard Wohl, 1921–1957. *The American Journal of Sociology*, **63**, 533–534.

Tal-Or, N., & Cohen, J. (2010). Understanding audience involvement: Conceptualizing and manipulating identification and transportation. *Poetics*, **38**, 402–418. doi:10.1016/j.poetic.2010.05.004

Tsao, J. (1996). Compensatory media use: An exploration of two paradigms. *Communication Studies*, **47**, 89–109. doi:10.1080/10510979609368466

Tsao, C. (2004). Research on parasocial involvement: An overview. *Journal of Hsuan Chuang Information and Communication*, **1**, 1–21.

Tukachinsky, R. (2010). Para-romantic love and para-friendships: Development and assessment of a multiple parasocial relationships scale. *American Journal of Media Psychology*, **3**, 73–94.

Wenner, L. A. (1983). Political news on television: A reconsideration of audience orientations. *Western Journal of Speech Communication*, **47**, 380–395. doi:10.1080/10570318309374132

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology*, **84**, 608–618. doi:10.1037/0022-3514.84.3.608