Journal of Language and Social Psychology

A Meta-Analysis of Trait–Behavior Correlations in Argumentativeness and Verbal Aggression

Timothy R. Levine, Michael R. Kotowski, Michael J. Beatty and Martijn J. Van Kelegom

The online version of this article can be found at:

Published by:

SAGE

Additional services and information for **Journal of Language and Social Psychology** can be found at:

Email Alerts: http://jls.sagepub.com/cgi/alerts

Subscriptions: http://jls.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://jls.sagepub.com/content/31/1/95.refs.html

>> Version of Record - Feb 6, 2012

OnlineFirst Version of Record - Nov 7, 2011

What is This?

A Meta-Analysis of Trait–Behavior Correlations in Argumentativeness and Verbal Aggression

Timothy R. Levine[1], Michael R. Kotowski[2],
Michael J. Beatty[3], and
Martijn J. Van Kelegom[2]

Abstract

This article addresses controversy over the validity of two popular scales used to measure trait argumentativeness and verbal aggressiveness. The first half of the article offers a rejoinder to the Infante, Rancer, and Wigley article. It is argued that original conceptualizations of the scales are logically incoherent and lack empirical correspondence with research findings. The second part of the article offers a meta-analysis of scale–behavior and –nonbehavioral associations. The results show that research testing scale–behavior convergence is sparse and that the little research that currently exists is inconsistent with convergent and predictive validity. The Infante scales correlate consistently and to a greater extent with self-reported communication than with actual behavior, suggesting that the scales assess cognitive–affective rather than communication behavior tendencies.

Keywords

[1]Michigan State University, East Lansing, MI, USA
[2]University of Tennessee, Knoxville, TN, USA
[3]University of Miami, Coral Gables, FL, USA

Corresponding Author:
Timothy R. Levine, Department of Communication, Michigan State University,
482 Communication Arts and Sciences Building, East Lansing, MI 48824, USA
Email: levinet@msu.edu

The construct validity and factor structures of the Argumentativeness Scale (ARG; Infante & Rancer, 1982) and Verbal Aggressiveness Scale (VAS; Infante & Wigley, 1986) are in dispute. The current authors published a series of studies showing that the ARG is unidimensional after removing problematic items (Kotowski, Levine, Baker, & Bolt, 2009) and that the VAS has two dimensions with aggressively worded items on one dimension and the reflected, reverse-scored items measuring supportive communication on a second factor (Beatty, Rudd, & Valencic, 1999; Kotowski et al., 2009; Levine et al., 2004). The two-factor interpretation of the VAS is further corroborated by (a) the observation that the reliability coefficients computed separately for each 10-item factor are approximately the same as the reliability coefficient for the 20-item single-factor treatment and (b) the correlation between the two factors is too small to sustain a conclusion of unidimensionality (Beatty et al., 1999).

In a subsequent multitrait–multimethod validation study, Kotowski et al. (2009) found that scores on the ARG did not correlate strongly with observations of argumentative behavior and that scores on the VAS did not correlate substantially with observed verbal aggression. Based on these findings, Kotowski et al. question the predictive and convergent validity of the ARG and VAS as measures of communication behavior traits. They speculated that the ARG and VAS instead measure cognitive and affective orientations. Simply put, the ARG and VAS appear to measure "want-to-be" argumentativeness and verbal aggressiveness rather than the tendency to engage in actual argumentative and verbal aggressive behaviors. This conclusion is based on the observation that the scales associate more strongly with self-report measures than with observed communication.

Infante, Rancer, and Wigley (2011) contested the results and conclusions of our previous work. They assert that *all* published work supports the validity of their scales as originally proposed. Infante et al. contend that (a) a large amount of research shows validity-consistent association between their scales and actual behavior, (b) their scales factor as intended, (c) original scales should be preferred because of "argumentative presumption," (d) a fair test requires consideration of four situation moderators, (e) Kotowski et al.'s (2009) findings result from measuring single behaviors, and (f) Kotowski et al. engaged in the statistically invalid practice of accepting null hypotheses.

In this article, we offer refutation of each of Infante et al.'s (2011) assertions.[1] We also provide a new meta-analysis of the associations between (a) scores on the ARG and argumentative behavior, (b) scores on the ARG and other paper-and-pencil measures of argumentativeness, (c) scores on the VAS and verbally aggressive behavior, and (d) scores on the VAS and other paper-and-pencil measures of verbal aggressiveness. Infante et al.'s assertions are shown to lack both logical coherence and correspondence with the results of scientific investigation. Infante et al.'s claims are therefore not scientifically defensible because valid scientific theory and measurement must be logically coherent and correspond with data.

Does Research Show Strong Scale–Behavior Correspondence?

Verbal aggressiveness is defined by Infante and Wigley (1986) as "a personality trait that predisposes people to attack the self-concepts of others" (p. 61). Argumentativeness is defined by Infante and Rancer (1982) as, a "trait that predisposes people to advocate positions on controversial issues while attacking verbally the positions which other people take on these issues" (p. 72). Explicit in both of these conceptual definitions is communication behavior. The higher someone is on verbal aggressiveness, the more likely he or she is to actually verbally aggress on others. Similarly, the more highly argumentative a person, the more likely that person is to argue with others. Thus, evidence essential to the validity portfolios of these scales are tests of correspondence between scores on the scale and the behavioral manifestations that define the respective constructs. Predictive and convergent validity requires empirical evidence of strong scale–behavior correspondence, and absent such evidence, validity cannot be presumed.

Infante et al. (2011) state that they are puzzled by our 2009 claim that relatively few studies have examined the link between their two scales and behavior. They cite a dozen studies that they claim provide empirical evidence that the two scales have predictive validity. They also assert that the construct validity of the ARG and VAS have been tested in "a very large number of studies" (p. 146) and "all [previous studies] support the validity of those original items" (p. 146). We do not dispute that the ARG and VAS sometimes perform well with self-report criterion variables and that a vast number of studies exist linking scores on the ARG and VAS to scores on other questionnaire-based self-report variables. The regional communication journals contain many such studies. It is important, however, to distinguish self-report research and behavioral research. For us, the two are not the same.

We believe that our disagreement with Infante et al. (2011) stems from the reliance on different definitions of behavior. For us, behavior refers to overt actions. Behaviors are things people *do* and that we can see them *doing*. Communication behaviors involve people communicating with real other people in ways that are directly observable. Put differently, measuring communication behaviors involves observation of messages exchanged between people. So, argumentative behavior involves someone actually refuting another's assertions and verbal aggression involves someone actually saying something with the intent of harming the message recipient's self-concept. We exclude things such as memory and projection from our definition of behavior. We see imagined interactions and recalled communication as more cognition than behavior. So, we do not consider making a rating of a hypothetical message to a hypothetical other person in an imagined situation argumentative or aggressive *behavior*.

It is clear from the citations provided by Infante et al. (2011) that they count projection and recall questionnaire studies as behavior measures. Infante et al.'s argument

has empirical merit if and only if self-reported projected and recalled communication are considered behavior and if and only if the results of observational research studying behavior as we defined behavior are ignored. To provide evidence for this point, the results of a new meta-analysis involving tests of convergent validity of the scales with observed behavior as well as self-reported pencil-and-paper measures are reported. The current meta-analytic findings sharply disconfirm Infante et al.'s claim of universal support. Strong associations are limited to self-report research. Validity coefficients involving behavioral criterion are small at best.

The Factor Structure of the ARG and VAS

We agree with Infante et al. (2011) that the factor structures of the scales are unequivocal in the sense that previous findings consistently replicate. The preponderance of evidence unequivocally shows that the ARG scale is a single dimension and the VAS has two factors. The problem for Infante et al. is that the findings are inconsistent with their conceptualizations. According to Infante et al., the VAS is a single dimension and ARG has two factors. Thus, Infante et al.'s claims do not correspond with empirical findings and are therefore not supported.

With regard to ARG, Infante et al. (2011) state, "Clearly, two dimensions of argumentativeness were predicted" (p. 147). Most researchers, however, score the ARG not only as two dimensions but also as if it were unidimensional. ARG produces *one* total score. Scores on the ARGav (tendency to avoid) scale are subtracted from scores on the ARGap (tendency to approach) scale when scoring total argumentativeness. Of course, adding or subtracting scores to obtain a total score presumes unidimensionality. Hypothesizing and interpreting two dimensions while scoring the scale as one dimension makes Infante et al.'s theoretical view logically and psychometrically incoherent. Infante et al.'s (2011) defense is that this scoring is consistent with their theory. However, such a theory cannot be accepted because asserting two dimensions and subtracting one from the other to form a total are logically contradictory and result in a lack of theoretical coherence.

Research shows that the ARG is unidimensional (Hamilton & Mineo, 2002; Kotowski et al., 2009). Since Infante et al. (2011) say ARG is bidimensional and it is not, they are wrong. Research findings as well as logic refute Infante et al.

In contrast, the VAS is supposed to be unidimensional (Infante et al., 2011; Infante & Wigley, 1986). Infante and Wigley did not propose a theory that verbal aggressiveness has two dimensions like argumentativeness. Rather, the unidimensional interpretation was post hoc after finding but dismissing a second orthogonal factor (Infante & Wigley, 1986). The problem with this is that all the research, including Infante's own original validation studies, finds that the VAS indeed has two factors (Beatty et al., 1999; Infante & Wigley, 1986, Kotowski et al., 2009; Levine et al., 2004, Suzuki & Rancer, 1994). Infante et al. (2011) misinterpreted the second factor as a methodological artifact based on a very small mean difference between the two sets of items. They asserted that because the second mean was slightly higher than the first, the second

factor was a social desirability artifact and could be ignored. The reported *t* value and the mean difference between positively and negatively worded items indicate that Infante and Wigley erroneously (1986) used a *t* test for independent samples rather than the appropriate *t* test for correlated samples. As a result, the error term was grossly reduced and the illusion of a statistically significant difference emerged when in fact the difference between positively and negatively worded items was less than one tenth of a point per item, far too small to indicate wording bias or any other meaningful interpretation.

There are other empirical problems with interpreting the second factor as a social desirability artifact. There is no affirmative evidence that the second factor is an artifact and there is affirmative evidence to the contrary—namely that scores on the second dimension function as meaningful construct (Levine et al., 2004). It should be noted, however, that if Infante et al. (2011) are right about the social desirability artifact, then the existence of an artifact strong enough to create the appearance of a second factor means that the VAS is confounded and therefore not construct valid.

Argumentative Presumption and Science

Infante et al. (2011) argue that their scales should be accepted because they are the status quo. This strikes us as a decidedly unscientific criterion because in our understanding of science, theory and measures are not presumed valid but must be consistent with data over time. As an epistemological method, tenacity is less desirable than science. Moreover, reliance on tenacity (steadfastly adhering to the status quo) impedes theoretical and methodological development. The moderate to weak validity coefficients in the literature are not sufficient rationale for rejecting scale improvement or reconceptualizations that more closely coincide with empirical findings.

Situational Variables

Infante et al. (2011) accuse Kotowski et al. (2009) of not understanding traits. They assert that traits do not correlate highly with single behaviors and that their theory says that the trait–behavior link is moderated by four situational variables. Our definition of a trait is a relatively stable tendency or predisposition to respond in a particular manner over time and across situations. If the behavioral effects of argumentativeness and verbal aggressiveness require a specific combination of four situational variables to obtain, then, by our definition, they are not traits. We have no doubt that both traits and situations guide behavior and that traits and situations interact to affect behavior. However, both argumentativeness and verbal aggressiveness are conceptually defined as behavioral traits. If they are simultaneously specified to be situation dependent, then this is more proof that Infante et al.'s (2011) theory is incoherent. Infante et al.'s position is once again self-contradictory and therefore can be rejected. A construct cannot be both trait and situation dependent.

Traits and Single Behaviors

Infante et al. (2011) dismiss the small trait–behavior correlations observed by Kotowski et al. (2009) as the results of correlating traits with momentary single behaviors. They cite a 40+-year-old claim that traits do not correlate with single behaviors. In our view, Infante et al. are correct about the predictability ceiling with single behaviors. But this has nothing to do with traits per se. Observations of single behaviors (e.g., an eyeblink, an isolated utterance, etc.) do not correlate highly with anything, traits or otherwise. Observations of single behaviors come with large random response errors, and thus their associations with other variables are substantially attenuated. Research designs that measure constellations of behavior over time and use large samples can address such issues.

As a criticism of Kotowski et al. (2009), the single behavior claim is specious. Kotowski et al. did not study single behaviors. Instead, they assessed global judgments of argumentation and verbal aggression over whole interactions in a situation likely to provoke argument and aggression.

Statistical Null Hypotheses

Infante et al. (2011) accuse Kotowski et al. (2009) of the practice of accepting a null hypothesis based on a nonsignificant result. This claim is also specious. Kotowski et al. did not infer that their nonsignificant trait–behavior correlations meant a zero association affirming the statistical null hypothesis. Instead, Kotowski et al.'s point was that because the observed trait–behavior correlations were small and sometimes in the wrong direction, they were too small to be acceptable as convergent validity coefficients. Here is some of what Kotowski et al. actually wrote,

> Most noteworthy, however, is the finding that the scale-behavior correlations for all three constructs were within sampling error of zero. The statistical power for these tests was in excess of .995 . . . These findings are inconsistent with the validity claims for the VAS and the ARG scale because the constructs are conceptualized as behavioral traits. Validity coefficients of the minimum magnitude recommended by Campbell and Fiske are outside the 95% and 99% confidence intervals. (p. 445).

A New Meta-Analysis

Key to both our disagreement with Infante et al. (2011) and to the validity of the ARG and VAS as originally conceptualized is Infante et al.'s assertion that the ARG and VAS are substantially related to argumentative and verbally aggressive behavior, respectively. Infante et al.'s position rests critically on their claim that an abundance of supportive evidence exists linking their scales to actual relevant

behaviors. Researchers have been skeptical of this claim since it does not match the literature.

To ascertain the empirical merits of Infante et al.'s (2011) claims relative to our claims, we conducted a literature search and meta-analysis to address six critical issues:

> *Research Question 1:* How abundant are studies linking ARG scores with actual argumentative behavior?
>
> *Research Question 2:* How abundant are studies linking VAS scores with actual verbally aggressive behavior?
>
> *Research Question 3:* What is the magnitude of association between scores on the ARG and actual argumentative behavior observed in the literature?
>
> *Research Question 4:* What is the magnitude of association between scores on the ARG and other survey reports of arguing in the literature?
>
> *Research Question 5:* What is the magnitude of association between scores on the VAS and actual verbally aggressive behavior observed in the literature?
>
> *Research Question 6:* What is the magnitude of association between scores on the VAS and other survey reports of verbal aggression in the literature?

Infante et al. (2011) assert an abundance of behavioral studies documenting substantial positive associations, whereas we expect that few such studies will be found and that the few existing studies will report effect sizes well below acceptable levels for predictive and/or convergent validity coefficients. Our argument is not for a zero or null association. Instead, our argument is that the essential scale–behavior correlations are small and that minimally acceptable predictive and convergent validity coefficients are outside the confidence limits around the estimates. Therefore convergent validity is statistically improbable.

Method

Meta-Analytic Procedure and Literature Search

The meta-analytic procedures used aggregated the strength of association between the ARG and VAS scales and their respective corresponding behavior across studies. This provided a clearer picture of the scale–criterion correlations. The Hunter and Schmidt (2004) variance-centered approach to meta-analysis was used.

An extensive literature search was conducted to obtain an exhaustive collection of relevant prior studies for use in the meta-analysis. First, all studies cited as evidence by Infante et al. (2011) were obtained. Next, two searches were performed using Google Scholar. "Cited by" searches on the original Infante and Rancer (1982) and Infante and Wigley (1986) scale development articles were conducted. These two searches were conducted through Google Scholar while it was linked to other databases in a university's library. These additional databases included PsychInfo,

Communication & Mass Media Complete, ProQuest Dissertations, and so on. Because of this linkage, the studies returned were from traditional academic databases as well as the web at large. To be thorough, the reference sections of the studies returned by these search criteria were then searched for manuscripts that the "cited by" searches might have missed.

Next, inclusion criteria were applied to the search results. For a study to be included in this meta-analysis, it had to meet four criteria. First, the study had to examine the relationship between (a) the ARG and argumentative behaviors, (b) the ARG and other survey argumentativeness measures, (c) the VAS and verbal aggressiveness behaviors, or (d) the VAS and other survey verbal aggressiveness measures. Second, the unit of analysis had to be the ARG or VAS respondent. Third, the study had to be peer or expert reviewed through publication in a peer-reviewed journal, through presentation as a peer-reviewed conference paper, or through the committee review process present in dissertation and thesis projects. Fourth, the study had to report sufficient information for an effect size to be computed or that information needed to be otherwise obtainable.

Hunter and Schmidt's (2004) meta-analytic procedures were applied to studies meeting the inclusion criteria. First, quantitative information regarding the relationship between the focal variables in each relevant study were transformed into *r*, the Pearson product–moment correlation coefficient. Second, the effects were weighted by sample size and grouped according to the four variable combinations under study. Third, the variability in effect sizes across studies (i.e., homogeneity) within each group was examined to determine if there was more heterogeneity in effect sizes than what could be attributable to sampling error. The final step involved attempting to account for heterogeneity if present.

Coding

All studies identified through the search procedures that administered either ARG or VAS were coded into one of eight categories based on three variables, each with two levels. The first variable coded was whether or not the ARG or VAS was used in the study. The second variable coded was whether or not the study examined the relationship between the ARG or VAS and a behavior (e.g., observed conversation) or a survey measure (e.g., hypothetical message generation). This variable was referred to as the measurement of the ARG or VAS correlate. The third variable coded was whether or not the study examined the relationship between the ARG or VAS and the same construct (e.g., a different way of measuring argumentativeness or verbal aggressiveness, respectively) or a different construct (e.g., instructor immediacy). This variable was referred to as the nature of the ARG or VAS correlate. Two coders worked independently, each coding approximately 75% of the identified studies on these three variables. Thus, approximately 25% of the total studies were double-coded. Intercoder reliability on these double-coded studies was high ($\kappa = .86$).

Table 1. Number of Studies in the Extant Literature Using the ARG or VAS

ARG			VAS		
	Measurement			Measurement	
Nature	Survey	Behavior	Nature	Survey	Behavior
Argumentativeness	9	4	Verbal aggressiveness	9	2
Other	99	7	Other	107	7

Note: ARG = Argumentativeness Scale; VAS = Verbal Aggressiveness Scale.

Results

Abundance of Behavioral ARG and VA Studies

Research Questions 1 and 2 addressed the prevalence of behavioral ARG or VAS studies. The search procedures identified 119 studies using the ARG and 125 studies employing the VAS. Table 1 presents the number of studies in each of the eight categories. Only a small percentage of the extant literature using either the VAS or ARG examined the relationships between either scale and behaviors of any sort. The number of studies that provide behavior-based evidence of convergent validity is two for VAS and four for ARG. References and coding for ARG studies can be found in Appendix A and the same information for VAS studies can be found in Appendix B.

Associations

Research Questions 3 through 6 addressed the associations between the ARG and VAS and other ways of measuring argumentativeness and verbal aggression. Table 2 presents an overview of the studies examined in this meta-analysis along with sample sizes and computed effect sizes, where possible broken down by group. The results are organized similarly by group.

ARG–ARG behavior. Four of 119 studies with a combined sample of 294 participants were found examining the ARG–ARG behavior relationship. All 4 studies contained enough information to compute r. The weighted mean correlation for these studies was .20 (unweighted mean correlation = .21) with a weighted standard deviation (*SD*) of .03 (unweighted *SD* = .03). Variability expected due to sampling error alone for this set of four studies with 294 participants was *SD* = .11. Consequently, 100% of the variability in the distribution of observed effects was accounted for by sampling error, $\chi^2(3, N = 294) = 0.21$, *ns*. These results were consistent with a homogenous ARG–ARG behavior effect across the 4 studies, with the 95% confidence interval ranging from .16 to .24.

ARG–ARG survey. Of the 119 studies using the ARG in the literature, 9 considered the relationship between the scale and some other survey measure of argumentativeness. Of those 9 studies, 2 did not present the information necessary to compute r. This left 7

Table 2. Effects and Sample Sizes for Studies Meeting the Inclusion Criteria by Group

Study	Sample	Effect	Note
ARG			
ARG behavior			
Infante (1981)	100	0.20	Will-to-argue effect
Kotowski, Levine, Baker, and Bolt (2009)	103	0.17	
Levine and Boster (1996)	60	0.21	
Semic and Canary (1997)	31	0.26	
ARG pencil-and-paper			
Infante (1985)	Insufficient information to compute *r*		
Infante (1987)	108	0.26	
Infante and Rancer (1982), Study 8	50	0.36	
Infante and Rancer (1993)	181	0.26	
Johnson, Becker, Wigley, Haigh, and Craig (2007)	102	0.50	Average effect of two conceptual replications within study
Kotowski, Levine, Baker, and Bolt (2009)	103	0.29	Average effect of two conceptual replications within study
Rancer, Baukus, and Infante (1985)	138	0.76	
Rancer, Kosberg, Baukus (1992)	Insufficient information to compute *r*		
Suzuki and Rancer (1994)	716	0.64	
VAS			
VA behavior			
Chory-Assad (2002)	87	0.13	Obtained directly from author
Kotowski, Levine, Baker, and Bolt (2009)	103	−0.10	
VA pencil-and-paper			
Beatty, Zelley, Dobos, and Rudd (1994)	74	0.44	
DiCioccio (2008)	219	0.57	
Infante and Rancer (1993)	175	0.42	
Infante and Wigley (1986), Study 3	104	0.43	
Infante and Wigley (1986), Study 4	86	0.69	
Johnson, Becker, Wigley, Haigh, and Craig (2007)	100	0.69	Average effect of two conceptual replications within study
Kotowski, Levine, Baker, and Bolt (2009)	103	0.44	Average effect of two conceptual replications within study
Levine et al. (2004), Study 1	194	0.43	Average effect of two conceptual replications within study
Levine et al. (2004), Study 2	177	0.51	
Sutter and Martin (1998)	401	0.37	
Suzuki and Rancer (1994)	716	0.40	

Note: ARG = Argumentativeness Scale; VAS = Verbal Aggressiveness Scale.

studies with a combined sample of 1,398 participants meeting the inclusion criteria for which *r* could be computed. The weighted mean correlation across these studies was .53 (unweighted mean correlation = .44). The weighted *SD* for this effect distribution was .18 (unweighted *SD* = .20). Variability expected due to sampling error for this set of 7 studies with 1398 participants was *SD* = .04. Thus, variability in the distribution of effects unaccounted for by sampling error was .14. Put differently, 22% of the variability in the effect distribution was attributable to sampling error. This difference between observed and expected variability was statistically significant, $\chi^2(6, N = 1,398) = 44.50$, $p < .05$. Thus, there appeared to be heterogeneity in the ARG–ARG nonbehavior effect across the 7 studies. The 95% confidence interval ranged from .38 to .68.

A review of the measure reliabilities and descriptive statistics in each of the seven studies revealed that differential measurement error or range restriction did not appear to account for the ARG–ARG survey effect heterogeneity. Therefore, potential moderator variables were considered with two confounded moderators emerging as likely candidates. The ARG survey measures were quite diverse with almost each study using a different method to measure a different conceptual outcome. Suzuki and Rancer (1994) used a semantic differential attitude toward arguing measure; Kotowski et al. (2009) relied on hypothetical message generation and selection tasks; Johnson, Becker, Wigley, Haigh, and Craig (2007) used a recalled argumentativeness instrument; Infante (1987) used a motivation to argue measure; and so on. Assuming each of these different measures assessed slightly different aspects of the argumentativeness construct, it would be expected that slightly different correlations between each different measure and the ARG would be observed. Because the different ARG survey measures used across these seven studies could not be grouped according to common features of their method, it was not possible to test this proposition in the usual meta-analytic way by partitioning the effects according to the proposed moderator to test if accounting for the moderator removed the effect heterogeneity.

VAS–VA behavior. Only 2 of the 125 extant VAS studies in the literature with a combined sample of 190 participants fit in this category. Although one of the two studies did not present information sufficient to compute *r* between respondent scores on the VAS and VA behaviors, it was obtained from the study's author. The weighted mean correlation for these two studies was .01 (unweighted mean correlation = .02). The weighted *SD* of the distribution was .11 (unweighted *SD* = .16). Variability expected due to sampling error for this set of 2 studies with 190 participants was *SD* = .10. Thus, variability in the distribution of effects unaccounted for by sampling error was .01, or 91% of the variability in the effect distribution was attributable to sampling error. This difference between observed and expected variability was within sampling error of zero, $\chi^2(1, N = 190) = 2.50$, *ns*. These results were consistent with a homogenous VAS–VA behavior effect across the two studies. The 95% confidence interval ranged from −.22 to .24.

VAS–VA survey. There were 11 studies with a combined sample of 2,349 participants examining the relationship between responses on the VAS and verbal aggressiveness measured through survey means. All 11 studies presented sufficient information for *r* to be computed. The weighted mean correlation for these studies was .45 (unweighted mean

correlation = .49. The weighted *SD* for this distribution of effects was .08 (unweighted *SD* = .10). Variability expected due to sampling error alone for this set of 11 studies with 2,349 participants was *SD* = .05. Thus, variability in the distribution of effects unaccounted for by sampling error was .03. Put differently, 63% of the variability in the effect distribution was attributable to sampling error. This difference between observed and expected variability was within sampling error of zero, $\chi^2(10, N = 2,349) = 17.40$, *ns*. Consequently, these results were indicative of a homogenous VAS–VA survey effect across the 11 studies with a 95% confidence interval of .39 to .51.

Discussion

The purpose of this article was to provide a reply to Infante et al. (2011). The current authors had previously published a series of articles finding that (a) the ARG and VAS have factor structures different than those originally specified by Infante and his colleagues, (b) both scales lack strong correlations with observations of verbally aggressive and argumentative behavior, and (c) both scales systematically correlate more strongly with self-reported behaviors than actual behaviors. From these findings, we had inferred that the original conceptualizations of the scales as measures of communication behavior traits were incompatible with research findings and that reconceptualization was necessary to obtain an acceptable degree of conceptual–empirical correspondence. Infante et al. (2011) disputed our interpretation and conclusions advocating instead for their original conceptualizations. The current reply refutes the assertions of Infante et al. (2011) and provides new meta-analytic findings consistent with the propositions that the ARG and VAS correlate substantially with self-reported communication but correlate less strongly with actual argumentative or verbally aggressive behavior.

One point of disagreement concerned the sheer quantity of existing research correlating scores on the ARG and VAS with actual argumentative and verbally aggressive behavior. Infante et al. (2011) maintained that a plethora of such studies exist. Our extensive literature search, which included all studies cited as evidence by Infante et al. (2011), yielded only 4 for the ARG and 2 for the VAS. Given that both scales (a) are defined as behavioral traits, (b) have been in use since the 1980s, and (c) have been used in more than a hundred studies each, the dearth of research actually assessing scores on the scales and corresponding behavior is surprising and troubling. Nevertheless, as far as we can tell, less than a handful of behavior-based convergent and predictive validity studies exist for the ARG and VAS. Evidence for Infante et al.'s (2011) claim of a large literature base of behavioral studies is nowhere to be found.

A second and more important point of disagreement concerned the magnitude of association between scores on the ARG and VAS and actual argumentative and verbally aggressive behavior, respectively. For the ARG, that correlation is $r = .20$ with a 95% confidence interval from .16 to .24. Furthermore, although this outcome is based on relatively few studies, the studies that exist are consistent in findings. Although the association is statistically larger than zero, we do not believe that a validity coefficient smaller than .25 is acceptable or desirable. Infante et al. (2011) claim that because the

ARG measures a trait, it should not correlate highly with behavior. Their claim, however, ignores that argumentativeness is defined as a behavioral trait.

The trait–behavior association for the VAS is approximately zero ($r = .01$, P [$-.22 \leq \rho \leq .24$] = .95). There is no evidence in the literature that scores on the VAS correlate with actual verbally aggressive behavior, although there is evidence that the association is within sampling error of zero. Note that we are not saying the correlation is exactly zero. Instead, we are saying that what little evidence exists indicates a very weak association, and this is inconsistent with convergent and predictive validity.

A third critical finding from the current meta-analysis is for the association of the ARG and VAS with self-reported projections or recollections of behavior. The meta-analysis shows that studies with self-report criteria are both more plentiful and yield larger effects. The lower end of the confidence intervals approach a correlation of $r = .40$ for both scales and extend to .68 for the ARG and .51 for the VAS. Thus, the ARG and VAS perform much better when the outcome is measured by self-report. That is, neither the ARG or VAS is strongly predictive of what people actually do, but both are more predictive of what people remember themselves doing or imagine themselves doing. This finding is fully consistent with our previous arguments and our current reply to Infante et al. (2011). The supportive findings to which Infante et al. point involve intention to behave, imagined behavior, or recollections of behavior rather than actual behavior.

There are at least two non–mutually exclusive explanations for why the scales correlate more strongly with self-reports than behaviors. The first is method variance. Studies using the scales and a self-reported criterion have both the construct and method in common because the ARG and VAS are also self-report measures. Behavioral studies involve criteria that are similar in construct but not method. Thus, the self-report outcome studies confound construct and method, whereas behavior observation studies do not involve construct–measure confounding. The current findings are consistent with the idea that correlations between self-report method might be inflated by a common method confound. If this is the case, the evidence Infante et al. (2011) cite in support of construct validity is contaminated by confounding and is therefore dubious. When such confounding is absent, so too is evidence on which Infante et al.'s (2011) position rests.

A second explanation is that the ARG and VAS measure affective and cognitive orientations rather than behavioral tendencies. If this is the case, then the larger correlations with self-reported outcomes make sense because the self-report criterion variables also assess cognitive–affective states such as attitudes, intentions, memories, projections, and desires. This explanation allows for the scales to be somewhat valid measures, but of constructs different than those originally specified. The pragmatic advantage of this explanation is that it makes sense of the findings without dismissing previous findings as pure artifact.

When the evidence is taken together, the conclusion is that the original conceptualizations of the ARG and VAS are not consistent with the results of research and are therefore flawed. This said, we continue to believe that argumentativeness and verbal

aggression are important communication constructs that merit research and that trait-like individual differences exist in argumentative and verbally aggressive behaviors. Furthermore, we do not believe that the ARG and VAS are universally invalid. Instead, it is our more modest conclusion that the ARG and VAS do not measure behavioral predispositions or tap the constructs as originally specified.

Appendix A

References and Coding of the Meta-Analyzed ARG Studies

Reference	Group
Infante, D. A. (1981). Trait argumentativeness as a predictor of communicative behavior in situations requiring argument. *Communication Studies, 32*, 265-272.	ARG–ARG behavior
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.	ARG–ARG behavior
Levine, T. R., & Boster, F. J. (1996). The impact of self and others' argumentativeness on talk about controversial issues. *Communication Quarterly, 44*, 345-358.	ARG–ARG behavior
Semic, B. A., & Canary, D. J. (1997). Trait argumentativeness, verbal aggressiveness, and minimally rational argument: An observational analysis of friendship discussions. *Communication Quarterly, 45*, 355-378.	ARG–ARG behavior
Infante, D. A. (1985). Inducing women to be more argumentative: Source credibility effects. *Journal of Applied Communication Research, 13*, 33-44.	ARG–ARG pencil-and-paper
Infante, D. A. (1987). Enhancing the prediction of response to a communication situation from communication traits. *Communication Quarterly, 35*, 308-316.	ARG–ARG pencil-and-paper
Infante, D. A., & Rancer, A. S. (1982). A conceptualization and measure of argumentativeness. *Journal of Personality Assessment, 46*, 72-80.	ARG–ARG pencil-and-paper
Infante, D. A., & Rancer, A. S. (1993). Relations between argumentative motivation, and advocacy and refutation on controversial issues. *Communication Quarterly, 41*, 415-426.	ARG–ARG pencil-and-paper
Johnson, A. J., Becker, J. A. H., Wigley, S., Haigh, M. M., & Craig, E. (2007). Reported argumentativeness and verbal aggressiveness levels: The influence of type of argument. *Communication Studies, 58*, 189-205.	ARG–ARG pencil-and-paper
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.	ARG–ARG pencil-and-paper
Rancer, A. S., Baukus, R. A., & Infante, D. A. (1985). Relations between argumentativeness and belief structures about arguing. *Communication Education, 34*, 37-47.	ARG–ARG pencil-and-paper
Rancer, A. S., Kosberg, R. L., & Baukus, R. A. (1992). Beliefs about arguing as predictors of trait argumentativeness: Implications for training in argument and conflict management. *Communication Education, 41*, 375-387.	ARG–ARG pencil-and-paper
Suzuki, S., & Rancer, A. S. (1994). Argumentativeness and verbal aggressiveness: Testing for conceptual and measurement equivalence across cultures. *Communication Monographs, 61*, 256-179.	ARG–ARG pencil-and-paper

Note: ARG = Argumentativeness Scale.

Appendix B

References and Coding of the Meta-Analyzed VAS Studies

Reference	Group
Chory-Assad, R. M. (2002). The predictive value of the verbal aggressiveness scale. *Communication Research Reports, 19*, 237-245.	VAS–VA behavior
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.	VAS–VA behavior
Beatty, M. J., Zelley, J. R., Dobos, J. A., & Rudd, J. E. (1994). Fathers' trait verbal aggressiveness and argumentativeness as predictors of adult sons' perceptions of fathers' sarcasm, criticism, and verbal aggressiveness. *Communication Quarterly, 42*, 407-415.	VAS–VA pencil-and-paper
DiCioccio, R. L. (2008). The development and validation of the teasing communication scale. *Human Communication, 11*, 255-272.	VAS–VA pencil-and-paper
Infante, D. A., & Rancer, A. S. (1993). Relations between argumentative motivation, and advocacy and refutation on controversial issues. *Communication Quarterly, 41*, 415-426.	VAS–VA pencil-and-paper
Infante, D. A., & Wigley, C. J. (1986). Verbal aggressiveness: An interpersonal model and measure. *Communication Monographs, 53*, 61-69.	VAS–VA pencil-and-paper
Johnson, A. J., Becker, J. A. H., Wigley, S., Haigh, M. M., & Craig, E. (2007). Reported argumentativeness and verbal aggressiveness levels: The influence of type of argument. *Communication Studies, 58*, 189-205.	VAS–VA pencil-and-paper
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.	VAS–VA pencil-and-paper
Levine, T. R., Beatty, M. J., Limon, S., Hamilton, M. A., Buck, R., & Chory-Assad, R. M. (2004). The dimensionality of the verbal aggressiveness scale. *Communication Monographs, 71*, 245-268.	VAS–VA pencil-and-paper
Sutter, D. L., & Martin, M. M. (1998). Verbal aggression during disengagement of dating relationships. *Communication Research Reports, 15*, 318-326.	VAS–VA pencil-and-paper
Suzuki, S., & Rancer, A. S. (1994). Argumentativeness and verbal aggressiveness: Testing for conceptual and measurement equivalence across cultures. *Communication Monographs, 61*, 256-179.	VAS–VA pencil-and-paper

Note: VAS = Verbal Aggressiveness Scale.

Declaration of Conflicting Interests

Funding

Note

1. Levine was one of two reviewers for the Infante et al. (2011) submission to *Communication Quarterly*. In his review, he stated his disagreement with Infante et al.'s assertions and his opinion that their assertions were not scientifically defensible. Nevertheless, his opinion was that in the spirit of an open scientific process, the Infante et al. submission might be published so long as replies were invited allowing both sides to make their cases, and so long as the editor invited the opinion of an independent, unbiased psychometric expert to publish an evaluation of the relative scientific merits of each side. Infante et al. objected to a reply from the current authors and objected to the independent assessment. The editor of *Communication Quarterly* published Infante et al. (2011) without inviting a reply or independent assessment. As a consequence, we are grateful to Howard Giles and Mark Hamilton for inviting this reply in *Journal of Language and Social Psychology*. Infante et al. were invited but declined the opportunity to participate in the current venue.

References

Beatty, M. J., Rudd, J. E., & Valencic, K. M. (1999). A re-examination of the verbal aggressiveness scale: One factor or two? *Communication Research Reports, 16*, 10-17.

Hamilton, M. A., & Mineo, P. J. (2002). Argumentativeness and its effect on verbal aggressiveness: A meta-analytic review. In M. Allen, R. W. Preiss, B. M. Gayle, & N. Burrell (Eds.), *Interpersonal communication research: Advances through meta-analysis* (pp. 281-314). Mahwah, NJ: Lawrence Erlbaum.

Hunter, J. E., & Schmidt, F. L. (2004). *Methods of meta-analysis: Correcting for error and bias in research findings*. Thousand Oaks, CA: SAGE.

Infante, D. A. (1987). Enhancing the prediction of response to a communication situation from communication traits. *Communication Quarterly, 35*, 308-316.

Infante, D. A., & Rancer, A. S. (1982). A conceptualization and measure of argumentativeness. *Journal of Personality Assessment, 46*, 72-80.

Infante, D. A., Rancer, A. S., & Wigley, C. J. (2011). In defense of the argumentativeness and verbal aggressiveness scales. *Communication Quarterly, 59*, 145-154.

Infante, D. A., & Wigley, C. J. (1986). Verbal aggressiveness: An interpersonal model and measure. *Communication Monographs, 53*, 61-69.

Johnson, A. J., Becker, J. A. H., Wigley, S., Haigh, M. M., & Craig, E. (2007). Reported argumentativeness and verbal aggressiveness levels: The influence of type of argument. *Communication Studies, 58*, 189-205.

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.

Levine, T. R., Beatty, M. J., Limon, S., Hamilton, M. A., Buck, R., & Chory-Assad, R. M. (2004). The dimensionality of the verbal aggressiveness scale. *Communication Monographs, 71*, 245-268.

Suzuki, S., & Rancer, A. S. (1994). Argumentativeness and verbal aggressiveness: Testing for conceptual and measurement equivalence across cultures. *Communication Monographs, 61*, 256-279.

Bios

Timothy R. Levine (PhD, Michigan State University) is a professor in the Department of Communication at Michigan State University. Levine's research interests include deception, interpersonal communication, personal relationships, persuasion and social influence, intercultural communication, communication traits, and measurement validation. Levine has published more than 100 journal articles.

Michael R. Kotowski (PhD, Michigan State University) is an assistant professor in the School of Communication Studies at the University of Tennessee, Knoxville. His primary area of research interest is social influence processes, with particular attention given to interpersonal persuasion and compliance gaining. He also maintains a strong research interest in the study of research methodology, with a particular emphasis on measurement.

Michael J. Beatty (PhD, Ohio State University) is a professor in the School of Communication at the University of Miami. He has been named among the top three most productive scholars in his field and recognized for his pioneering communication research in the area of bio-communication theory. He has published numerous books and scholarly articles in *Communication Monographs*, *Quarterly Journal of Speech*, and *Communication Quarterly*, among others.

Martijn J. Van Kelegom (MS, University of Tennessee, Knoxville) is a doctoral student in the College of Communication and Information at the University of Tennessee, Knoxville. He is interested in interpersonal communication. His research focus emphasizes processes related to goals and planning, such as imagined interactions.