International
Communication
Association

OXFORD

The reconstructability of persuasive message variables affects the variability of experimental effect sizes: evidence and implications

Hans Hoeken [iD] **[1]*, Daniel J. O'Keefe** [iD] **[2]**

[1]Department of Languages, Literature, and Communication, Utrecht University, Utrecht, the Netherlands
[2]Department of Communication Studies, Northwestern University, Evanston, IL, USA

*Corresponding author: Hans Hoeken; e-mail: j.a.l.hoeken@uu.nl

Abstract

Whereas the persuasive impact of message variables such as weaker versus stronger threat appeals, vivid versus pallid messages, and one-sided versus two-sided messages has received much research attention, more abstract properties of such message variables have gone largely unexamined. This article reports an analysis of one such property, reconstructability: the degree to which one of the two messages in an experimental pair can be deduced from the other. Evidence is offered from research on persuasive communication that as message variables become less reconstructable, the variability of the associated effect sizes increases—which creates distinctive challenges for theoretical progress and practical message design. Attention to message-variable properties such as reconstructability promises to shed light on how and why effects differ across message variables.

Keywords: Persuasion, Heterogeneity, Meta-analysis

Much experimental research has taken up the question of the effects of message variables on persuasive outcomes—strong versus weak fear appeals, narrative versus non-narrative formats, gain-framed versus loss-framed appeals, and so on. Results from such studies can be described using effect sizes (ESs); for a given study, the ES describes the difference in persuasiveness between the two message forms being compared.

Meta-analytic treatments of such research have naturally focused on the mean ES associated with a given message variable—the average persuasive advantage conferred by a given message design choice. But attention is now also being given to effect-size *variability*. A meta-analytic mean ES is based on some number of ESs from individual studies, and those individual ESs vary in magnitude. That is, the ESs are not uniform but rather heterogeneous. (For some recent discussions of ES heterogeneity, see Bryan et al., 2021; Kenny & Judd, 2019; Levine & Weber, 2020; Linden & Hönekopp, 2021.)

One question that naturally arises is: For a given message variable, how can the observed variability in persuasion ESs be explained? Answering this question has commonly involved examination of potential moderating variables. For example, the relative persuasiveness of gain-framed and loss-framed appeals has been hypothesized to be influenced by the perceived riskiness of the advocated behavior (see, e.g., Rothman et al., 2006; for an assessment of this hypothesis, see Van 't Riet et al., 2016). Similarly, the relative persuasiveness of one-sided and two-sided messages has been hypothesized to be influenced by whether the two-sided message is refutational or nonrefutational (Allen, 1991; for an assessment, see O'Keefe, 1999a).

Questions about how to account for effect-size variability for specific message variables will naturally have different answers for different message variables. The factors that moderate the relative persuasiveness of narrative and non-narrative formats will not necessarily be the same as those that moderate the relative persuasiveness of metaphorical and non-metaphorical messages. But the abstract idea is the same: For a given message variable, the presence of effect-size variability can be a consequence of having included ESs representing different levels of some relevant moderator. Within the levels of such a moderator, ESs are expected to be more consistent.

This article takes up a related, but subtly different, research question about effect-size variability: What influences the *amount* of persuasion effect-size variability for message variables? The suggestion to be advanced is that some message variables are more likely than others to yield persuasion ESs that vary considerably (quite independent of any influence of moderating factors). Thus the present interest is not with the variability associated with this or that message variable in particular (and so not with explaining the variation within the ESs associated with any specific message variable), but rather with examining differences *between* message variables in the amount of effect-size variability they exhibit.

The specific purpose of this article is to explore the hypothesis that the amount of persuasion effect-size variability associated with a message variable will be influenced by the nature of the message variable itself. That is, the expectation is that because of the properties of the message variables

involved, different message variables will display different amounts of effect-size variability. The implication is that, for example, some message variables will naturally be more likely than other to exhibit apparent "replication failures," just because the nature of the message variable is such as to create a greater amount of variation in ESs. What on the surface seems like a failure-to-replicate might instead be an understandable consequence of the nature of the message variable being studied—which would suggest some caution before concluding that a given effect is not in fact genuine.

In what follows, first some greater attention is given to how to describe the amount of persuasion effect-size variation associated with a given message variable. One key message-variable property of interest is then discussed: the reconstructability of experimental messages, that is, whether seeing one experimental message (in a pair) permits one to reconstruct the other message. Then relevant empirical evidence is reviewed to assess the degree to which message reconstructability is associated with differences in the amount of persuasion effect-size variability.

Describing ES variation

The present project's interest is in seeing how and why different message variables might have systematically different amounts of variability among the persuasion ESs associated with that variable. The purpose of this section is to provide some greater clarification about the idea of ES variation, and specifically to provide a means of describing ES variation.

As an entry point: Imagine two message variables A and B (variables such as narrative vs. non-narrative, strong vs. weak threat appeals, gain-framed vs. loss-framed appeals, and so on) that have been studied for their effects on message persuasiveness. Meta-analyses of the research on each variable reveal that the two message variables have the same mean ES; that is, the meta-analytic mean difference in persuasiveness between the two versions of message variable A is the same as the meta-analytic mean difference in persuasiveness between the two versions of message variable B.

However, the dispersion of the ESs is quite different. The ESs for message variable A are all tightly clustered around the mean effect; there's not much difference from study to study in the observed ES. But for message variable B, the ESs are quite dispersed; there's a lot of variation from study to study in the observed ES. That difference—in the absolute amount of variability in ESs—is the focus of the present article. This focus is potentially open to misunderstanding, so this section tries to be clear about the present interest.

The focus here is not on the *heterogeneity* of ESs, at least under some definitions of heterogeneity. For example, the present interest is not represented by familiar meta-analytic "measures of heterogeneity" such as I^2, Q, H, and Birge's R (Birge, 1932; Card, 2012, pp. 184–191; Higgins & Thompson, 2002; Higgins et al., 2003; Huedo-Medina et al., 2006). The reason is that those indices do not describe the absolute amount of variation in a set of ESs. Those indices describe the relative amount of variation—relative to the amount expected on the basis of human sampling variation.

As Borenstein et al. (2017, p. 11) have explained concerning I^2, this point is often misunderstood: "A small value of I^2 is interpreted as meaning that the effect size is comparable across studies. A large value of I^2 is interpreted as meaning that the effect size varies substantively across studies. In fact,

I^2 does not tell us how much the effect size varies." The reason is that "I^2 is a proportion and not an absolute value. As such, it cannot tell us how much the effects vary" (p. 7; see also Rücker et al., 2008).

Similarly, the present interest is not heterogeneity as defined by Levine and Weber (2020, p. 343): "Heterogeneity of effects exists when effects vary from primary study to primary study more than would be expected by sampling error alone." In that usage, heterogeneity does not exist if the amount of variability is relatively small (that is, relative when compared to, and hence potentially explainable by, human sampling variation); heterogeneity would be said to exist when the value of a heterogeneity index is statistically significant, because such a result indicates a greater amount of ES-to-ES variation than would be expected on the basis of sampling variation alone. But the present project is interested in the absolute amount of variability of ESs, regardless of the relationship between the amount of observed variability and the amount of variability expected on the basis of sampling variation alone.

So the present article's concern is specifically with the absolute amount of ES variability associated with persuasive message variables.[1] Two (related) ways of describing that property are used here: prediction intervals and T (the estimate of tau).

A prediction interval specifies the plausible range of ESs to be observed in the next application (e.g., the next study). Thus "a prediction interval provides a description of the plausible range of effects if the treatment is applied in a new study or a new population similar to those included in the meta-analysis" (Partlett & Riley, 2017, p. 302). "A 95% prediction interval estimates where the true effects are to be expected for 95% of similar (exchangeable) studies that might be conducted in the future" (IntHout et al., 2016, p. 2). So the more dispersed a given set of ESs is, the greater uncertainty there is about the location of the ES in the next study, and hence the wider the corresponding prediction interval will be. (For computational details, see Borenstein, 2019, pp. 337–345; Borenstein et al., 2021, pp. 121–122.)

Prediction intervals (PIs) are not to be confused with confidence intervals (CIs) as these address different questions (see Borenstein, 2019, pp. 94–96). The 95% CI around a meta-analytic mean ES gives the range of plausible population (mean) values; the 95% PI gives the range of plausible future individual ESs. The relationship of PIs and CIs is sometimes misunderstood because of a mistaken belief that the CI describes the dispersion of effects in individual studies; the CI instead describes the range within which the *mean* effect (the population effect) is likely to be found. The CI answers the question "where is the mean effect likely to be?" whereas the PI answers the question "where is the effect size in an individual study likely to be?"

The width of a prediction interval thus provides a straightforward representation of the absolute variability of the associated ESs. As the dispersion of ESs increases, so will the width of the corresponding PI. And hence comparing PI widths between message variables will give an indication of how variables might differ in the variability of their ESs.

A second way of describing ES variation uses T, the estimate of the standard deviation of the true ESs. T is the square root of T^2, the estimate of the variance of the true ESs (the estimate of the between-study variance). (For computational details, see Borenstein, 2019, p. 336; Borenstein et al., 2021,

pp. 106–109.)[2] T thus provides another representation of the absolute variability of the associated ESs. As the dispersion of ESs increases, so will the value of T. And hence comparing values of T between message variables will give an indication of how variables might differ in the variability of their ESs.

Prediction intervals and T are related, because the width of the prediction interval—the plausible range of ESs in future individual studies—is affected by (inter alia) the observed variability of ESs, which is reflected in T. (In fact, T figures in the computation of the associated prediction interval.) But because values of T are likely to be opaque to many whereas prediction intervals are more readily understandable, both are reported here.

The reconstructability of message variable manipulations

In experimental studies of the effects of different messages on persuasive outcomes, message variables are manipulated so as to create pairs of messages differing in some specified way (e.g., vivid versus pallid messages, or messages with strong versus weak arguments). But if the ways in which a given message variable is manipulated varies widely between studies, there might naturally be variation in ESs as well.

For example, Hoeken et al. (2020) showed that research on the contrast between "strong arguments" and "weak arguments" has used very different operationalizations of "weak arguments": arguments that point to less desirable consequences than the strong arguments, arguments that present information that is not relevant to the issue at hand, or even counterarguments to the advocated view (e.g., arguments describing undesirable consequences of the advocated action). Such variations in operationalization could contribute to variation in observed ESs.

Message variables might naturally differ in the extent to which there are such differences in operationalizations. To assess how and to what extent manipulations of a given message variable differ between studies, one might ideally collect and analyze all the different message versions that have been used in the studies of that message variable; this could permit comparisons between different message variables with respect to variation in the extent to which operationalizations differ. However, for many—especially, but not exclusively, older—studies the experimental message materials are unavailable, making this an unworkable way to proceed. The question is whether there is some alternative way of systematically describing such difference between message variables.

The proposal offered in this article is that an estimate of the variation in operationalizations can be found in the *reconstructability* of the message versions. Concretely put, the question is whether the definition of the message variable enables a researcher who is provided with one version of the message to reconstruct the other version of the message. Three categories of message-variable reconstructability can be distinguished: fully reconstructable, semi-reconstructable, and unreconstructable.

A *fully reconstructable* message variable is one in which either version of the message can be deduced from the other. The message variable labeled "But you are free" provides an example. Carpenter (2013, p. 6) describes this variable as follows: "In the control condition, the experimenter made a simple direct request: 'Sorry, Madam/Sir, would you have some coins to take the bus, please?' In the experimental condition,

the experimenter added: 'But you are free to accept or to refuse.'" If one is provided with the message for the experimental condition, one can reconstruct what the message in the control condition must have looked like by deleting the 'But you are free' phrase. And if provided with the control-condition message, one can reconstruct what the experimental-condition message must have looked like by adding the 'But you are free' phrase.

For message variables to be termed "fully reconstructable" here, it is necessary that the reconstructed message be quite constrained, even if the exact wording cannot be deduced. For instance, the contrast between rhetorical questions and statements is a fully reconstructable variable even though (e.g.) the rhetorical version of a statement such as "Carbon dioxide emissions should be reduced" could read either "Shouldn't carbon dioxide emissions be reduced?" or "Carbon dioxide emissions should be reduced, shouldn't they?"

A *semi-reconstructable* message variable is one in which there is an asymmetry in reconstructability: given the variable's definition, one version of the message can be deduced from the other, but not vice versa. That is, one can reconstruct version A based on version B but not the other way around. As an example, consider research on the effects of adding visual material to text, where the message contrast is text-only versus text-plus-visual. One can reconstruct the text-only version by deleting the visual of the text-plus-visual version; however, one cannot reconstruct the text-plus-visual version based on the text-only version.

An *unreconstructable* message variable is one in which neither version of the message can be deduced from the other. There are two kinds of unreconstructable message variables, ones involving categorical message variables and ones involving continuous message variables.

For *unreconstructable categorical* message variables, the message versions to be compared differ on various unidentified dimensions. For instance, Walter et al. (2018) describe their inclusion criteria for studies on the effect of humor as follows: "a more realistic inclusion criteria will require relevant studies to provide a direct comparison between a humor message and a humorless message on the same topic, rather than two identical messages that differ only with respect to the inclusion of humor" (Walter et al., 2018, p. 352). As a result, when provided with a humorous message, it is not possible to reconstruct the humorless message or the other way around. Note, however, that for such categorical message variable, one can presumably tell which message type is represented by a given message; that is, one can tell whether the message is (e.g.) the humor message or the humorless message.

For *unreconstructable continuous* message variables, the message variable varies along a continuum. For such message variables, not only can one not reconstruct either message given knowledge of the other, one cannot even be sure which message type is represented by a given message. Studies of speech-rate variation (e.g., as reviewed by Preiss et al., 2014) provide an example. Such studies compare the persuasiveness of a faster-spoken message to a slower-spoken message (expressed in wpm: words per minute). But very different points on the wpm continuum can be chosen to create a rate contrast. Hence, for example, a speech rate classified as relatively "slow" in one study (207 wpm; Vann et al., 1987) can be considerably higher than that of the relatively "fast" message in another study (140 wpm; Miller et al, 1976). And

such studies also differ in the size of difference between the two conditions; for example, in one study it was 17 wpm (173 vs. 190; Vann et al., 1987) but in another was 151 wpm (145 vs. 296; Wheeless, 1971).

Continuous message design variables can thus lead to additional variation in ESs for two reasons. First, studies may differ in which part of the continuum they exploit. The two versions differing in, for instance, speech rate might both be located toward the lower end of the continuum, or toward the higher end, or somewhere in between. Especially when a message version appears in the middle of the continuum, it may be impossible to decide whether the message is meant to represent the high or low version of the comparison. Second, the difference between the two versions may be relatively small in one study but quite large in another study.

To summarize: Message variables differ in reconstructability—and hence in the extent to which the manipulation of the message variable in different studies is likely to be quite similar or very different. For fully reconstructable message variables, the definition of the message variable determines how to design message version A based on version B, and how to design version B provided with version A; this should lead to very similar manipulations across studies. Semi-reconstructable message variables enable designing version A when given version B, but not the other way around; such manipulations seem likely to vary more widely compared to the fully reconstructable variables. And unreconstructable message variables should show considerable variation in operationalization across studies, because the message variable is based either on unreconstructable message categories or on a continuum.

Methods

The hypothesis to be examined is that message-variable reconstructability is related to the variability of persuasion ESs. As a way of obtaining evidence relevant to that hypothesis, we exploited the dataset of O'Keefe and Hoeken (2021). Their dataset consisted of 30 meta-analyses of the effect of different message variables on persuasive outcomes. A number of previous papers have reviewed persuasion meta-analyses and so might potentially have been sources of relevant data (Dillard, 1998; Rains et al., 2018; Weber & Popova, 2012). But O'Keefe and Hoeken's (2021) review has three features that made it especially attractive as a data source.

One is its broad sweep of message variables. O'Keefe and Hoeken's (2021) review concerned meta-analyses of studies in which two versions of a persuasive message were compared, that is, meta-analyses of persuasion message variables; they included 30 meta-analyses. And their review appears to be rather inclusive, as suggested by the diversity of the message variables included: familiar well-studied variables (such as gain-framed versus loss-framed and narrative versus non-narrative), but also other variables that don't figure so prominently in the research literature (e.g., speaking rate). They identified, and included as appropriate, unpublished meta-analyses. They excluded persuasion reviews that did not concern the persuasive effects of message variations, such as reviews of the effects of communication campaigns and reviews of the effects of psychological states such as guilt or anger. Thus their review concerned message variables

specifically—the focus of the present report—and it identified a large number of relevant meta-analyses.[3]

Second—and especially relevant to the present project—is their analysis of ES variability. Other reviews have focused on the magnitude of the mean ESs associated with message variables, but have not reported data concerning ES variability.[4]

Third, O'Keefe and Hoeken's (2021) underlying data are publicly available, unlike the data for some previous reviews of persuasion meta-analyses (e.g., Weber & Popova, 2012). Their data—and thus the data for the present report—can be scrutinized or re-analyzed by other researchers.

We classified each of O'Keefe and Hoeken's (2021) 30 message variables as fully reconstructable, semi-reconstructable, or unreconstructable (the Supplementary Appendix provides details, including information about intercoder reliability). We then examined the variability of the ESs associated with each message variable by (a) calculating the width of its 95% prediction interval as reported by O'Keefe and Hoeken (2021) and (b) calculating T, the estimate of the standard deviation of the true ESs.[5] Wider prediction intervals and larger values of T indicate greater absolute variability in a set of ESs.

Six of the message variables were fully reconstructable: appeal framing (gain vs. loss; meta-analytic evidence from O'Keefe & Jensen, 2006), "but you are free" (included vs. omitted; Carpenter, 2013), conclusion (included vs. omitted; O'Keefe, 2002), disrupt-then-reframe (vs. reframe-only; Carpenter & Boster, 2009), legitimizing paltry contributions (included vs. omitted; Bolkan & Rains, 2017), and rhetorical questions (vs. statements; Gayle et al, 1998).

Five of the message variables were semi-reconstructable: information-source identification (included vs. omitted; O'Keefe, 1998), metaphorical messages (vs. non-metaphorical; Brugman et al., 2019), sidedness (one-sided vs. two-sided; O'Keefe, 1999a), "that's not all" (included vs. omitted; Lee et al., 2019), and visual material (text-plus-visual vs. text-only; Seo, 2020, and Seo & Kim, 2018).

Nineteen of the message variables were unreconstructable, of which thirteen were continuous variables and six were categorical variables. The unreconstructable continuous variables were: argument explicitness (explicit vs. implicit; meta-analytic evidence from O'Keefe, 1998), argument strength (strong vs. weak; Carpenter, 2015), depicted response efficacy (high vs. low; Witte & Allen, 2000), depicted self-efficacy (high vs. low; Witte & Allen, 2000), depicted threat severity (high vs. low; De Hoog et al., 2007), depicted threat vulnerability (high vs. low; De Hoog et al., 2007), evidence amount (more vs. less; Stiff, 1985, 1986), language intensity (high vs. low; Hamilton & Hunter, 1998), political advertising tone (positive vs. negative; Lau et al., 2007), recommendation specificity (specific vs. general; O'Keefe, 2002), speaking rate (faster vs. slower; Preiss et al., 2014), threat appeal strength (strong vs. weak; White & Albarracín, 2018), and vividness (vivid vs. pallid; Blondé & Girandola, 2016). The unreconstructable categorical variables were: cultural tailoring (deep-tailored vs. not tailored; Hornikx & O'Keefe, 2009), evidence type (statistical vs. narrative; Allen & Preiss, 1997), humor (humorous vs. non-humorous; Walter et al., 2018), narrative (vs. non-narrative; Shen et al., 2015), sexual content (vs. non-sexual; Lull & Bushman, 2015), and victim description (identifiable vs. non-identifiable; Lee & Feeley, 2016).

Table 1. Prediction intervals, prediction interval widths, and Ts for message variables

Message variable	k	95% PI	95% PI width	T
Appeal framing (gain vs. loss)	165	−0.149, 0.180	0.329	0.084
Sidedness (two-sided vs. one-sided)	107	−0.277, 0.273	0.550	0.141
Metaphorical (vs. non-metaphorical)	91	−0.119, 0.254	0.373	0.095
Cultural tailoring (deep-tailored vs. not-tailored)	67	−0.223, 0.357	0.580	0.148
Humor (humorous vs. non-humorous)	58	−0.300, 0.500	0.800	0.212
Depicted threat severity (high vs. low)	55	−0.106, 0.327	0.433	0.110
Threat appeal strength (strong vs. weak)	48	−0.396, 0.551	0.947	0.255
Speaking rate (faster vs. slower)	44	−0.387, 0.495	0.882	0.232
"But you are free" (included vs. omitted)	42	−0.008, 0.346	0.354	0.089
Victim description (identifiable vs. non-identifiable)	41	−0.174, 0.273	0.447	0.110
Vividness (vivid vs. pallid)	37	−0.168, 0.440	0.608	0.155
Narrative (narrative vs. non-narrative)	34	−0.107, 0.237	0.344	0.084
Legitimizing paltry contributions (included vs. omitted)	34	−0.013, 0.434	0.337	0.114
Depicted threat vulnerability (high vs. low)	32	−0.245, 0.516	0.761	0.197
Evidence amount (high vs. low)	31	0.040, 0.395	0.355	0.089
Political advertising tone (positive vs. negative)	27	−0.423, 0.406	0.829	0.210
Depicted response efficacy (high vs. low)	24	−0.148, 0.501	0.649	0.164
Depicted self-efficacy (high vs. low)	21	−0.138, 0.495	0.633	0.158
Visual material (text-plus-visual vs. text-only)	20	−0.240, 0.342	0.582	0.138
Rhetorical questions (vs. statements)	18	−0.103, 0.218	0.321	0.071
Recommendation specificity (specific vs. general)	18	−0.158, 0.347	0.505	0.118
Argument explicitness (explicit vs. implicit)	18	−0.113, 0.372	0.485	0.114
"That's not all" (included vs. omitted)	18	−0.198, 0.477	0.675	0.161
Conclusion (included vs. omitted)	17	−0.185, 0.373	0.558	0.130
Evidence type (statistical vs. narrative)	16	−0.307, 0.385	0.692	0.161
Language intensity (high vs. low)	15	−0.195, 0.230	0.425	0.095
Disrupt-then-reframe (vs. reframe-only)	14	0.225, 0.347	0.122	0.000
Information-source identification (included vs. omitted)	13	−0.112, 0.251	0.363	0.077
Argument strength (strong vs. weak)	13	−0.084, 0.437	0.521	0.118
Sexual content (sexual vs. non-sexual)	11	−0.293, 0.329	0.622	0.138

k: number of effect sizes; PI: prediction interval; *T*: estimate of the standard deviation of the true effect sizes.

Results

The PI widths and *T* values for each message variable are given in Table 1. As indicated in Table 2, both PI widths and *T* values varied depending on the reconstructability of the message variable. Effect size variability was smallest for fully reconstructable variables (mean PI width of 0.34, mean *T* of 0.08), intermediate for semi-reconstructable variables (mean PI width of 0.51, mean *T* of 0.12), and largest for unreconstructable variables (mean PI width of 0.61, mean *T* of 0.15). Among unreconstructable variables, there was not much difference in ES variability between categorical variables (mean PI width of 0.58, mean *T* of 0.14) and continuous variables (mean PI width of 0.62, mean *T* of 0.16).[6]

Discussion

There is considerable variability in the ESs associated with persuasive message variables. All of the prediction intervals are rather wide (mean PI width = 0.54), and most of them (28/30, 93%) include both positive and negative values; the *T* values also suggest considerable variation in ESs (mean *T* = 0.13).

But the reconstructability of a message variable influences the variability of the associated ESs. Specifically, as reconstructability diminishes, effect-size variability increases: prediction intervals widen (mean PI width of 0.34 for fully reconstructable variables, 0.61 for unreconstructable variables) and *T* increases (mean *T* of 0.08 for fully reconstructable variables, 0.15 for unreconstructable variables).

Consequences of effect-size variability

Large effect-size variability creates at least three challenges. First, apparent replication failures are made more likely. Two studies of a given message variable are more likely to produce divergent results if the underlying population of ESs has more intrinsic variability.

Second, establishing the existence of a non-zero mean (population) effect is made more difficult. If there is little ES variability from study to study, then a relatively smaller number of studies might end up providing sufficient evidence to underwrite the claim of a non-zero effect; that is, it's easier to find a statistically significant mean meta-analytic effect if the ESs being synthesized consistently find similar positive effects. (And this is the case both for establishing main effects and for establishing contingent effects, that is, effects when limiting conditions are specified.)

Third, guidance for message designers becomes less certain. The larger the ES variability for a message variable, the less predictable is the effect of that variable for any given future application—and hence the less confidence one can have in making message design recommendations.

But, as the present results indicate, these challenges are likely to be more substantial for some message variables than for others. Specifically, the less reconstructable a message variable is, the more significant these challenges will be. Researchers and message designers should plan accordingly. For example, when studying an unreconstructable message, researchers should expect to need more studies before being in a position to reach dependable conclusions. And message

Table 2. Prediction interval widths and Ts as a function of reconstructability

Message variable reconstructability	N	Mean 95% PI width (SD)	Median 95% PI width	Mean T (SD)	Median T
Fully reconstructable	6	0.337 (0.14)	0.333	0.081 (0.05)	0.087
Semi-reconstructable	5	0.509 (0.14)	0.550	0.122 (0.03)	0.138
Unreconstructable	19	0.606 (0.18)	0.608	0.151 (0.05)	0.148
Categorical	6	0.581 (0.16)	0.601	0.142 (0.04)	0.143
Continuous	13	0.618 (0.19)	0.608	0.155 (0.05)	0.155

N: number of cases (mean effect sizes); PI: prediction interval; T: estimate of the standard deviation of the true effect sizes; SD: standard deviation.

designers should expect that the effects associated with an unreconstructable message variable will be less predictable than those associated with other variables.

Contextualizing the observed effect-size variability

To better understand the ES variability reported above, it will be useful to compare that variability to other reports of ES variability. Two such reports are Linden and Hönekopp's (2021) analysis of T values in psychological research and O'Keefe and Hoeken's (2021) analysis of prediction-interval widths in persuasive message effects research.

Analysis of T values

The variation estimates reported by Linden and Hönekopp (2021) for close replications and for conceptual replications provide particularly informative comparisons. Close replications aim to replicate studies as much as possible, using the exact same intervention, measuring dependent variables and instructing participants in the exact same way, and taking care that the samples are as similar as possible; close replications thus functionally set a lower limit on what variation in ESs might be expected. Conceptual replications, on the other hand, are the sorts of studies commonly reviewed in a meta-analysis; although every study addresses the same abstract research question, the concrete experimental materials, samples, and so on vary considerably.

Linden and Hönekopp's (2021, p. 364) mean T for close replications was 0.05; that for conceptual replications was 0.16.[7] The Ts reported here fall within this range; the Ts for semi-reconstructable (0.12) and unreconstructable (0.16) variables are toward the upper end of the range, and the T for reconstructable message variables (0.08) is at the lower end.

But for all the message variables reviewed here, the studies included in the meta-analyses were conceptual replications; thus one might have expected that all the message variables would have displayed ES variability similar to Linden and Hönekopp's (2021) value for conceptual replications. However, the observed variability for fully reconstructable message variables is noticeably smaller than that. In fact, the variability for fully reconstructable message variables is much closer to the variability seen for close replications than to that seen for conceptual replications.

So compared against the T values seen in psychological research generally, the variability of persuasive message-effect ESs seems quite typical for unreconstructable message variables—but the variability is considerably smaller for fully reconstructable message variables, and indeed approaches what one might think of as the lower bound of variability.

Analysis of prediction-interval widths

O'Keefe and Hoeken (2021) reported prediction-interval widths for persuasive message variables, both for simple main effects and for various contingent effects, that is, effects under specified moderating conditions. The PI widths of interest here are the ones observed when taking one or two moderating conditions into account. Because these subsets of effects come from studies that share some properties (e.g., where the moderating variable is the nature of the participant sample, the studies will have similar samples), the ESs can be expected to be more consistent.

More specifically, the PI widths of interest are those that O'Keefe and Hoeken reported as the *narrowest* PI widths observed under moderating conditions. These provide a realistic estimate of the narrowest PI widths that might be expected (just as Linden and Hönekopp's mean T for close replications provides a realistic estimate of the lower bound to be expected for T values). O'Keefe and Hoeken (2021, p. 7) reported that the mean PI width for the narrowest PI widths was 0.38 when taking one moderator into account and was 0.32 when taking two moderators into account.

All of the mean ESs analyzed here represented simple main effects (without moderating conditions being specified), and so one might have expected the ES variability to be considerably larger than those values. And indeed that is the case for semi-reconstructable message variables (mean PI width of 0.51) and for unreconstructable message variables (0.61). However, the mean PI width for fully reconstructable message variables (0.34) is very much in line with the narrowest widths seen when moderating conditions are specified. That is, the PI widths for fully reconstructable message variables— even when no moderating conditions are specified—are rather like those seen more generally for the narrowest PI widths when moderators are taken into account.

Summary

When a message variable is fully reconstructable, the observed variability in ESs is noticeably small—close to the effective lower bound for variability. That's apparent both for T values (where the mean T for fully reconstructable message variables is in the neighborhood of the mean T for close replications in psychological research generally) and for PI widths (where the mean PI width for fully reconstructable message variables is in the neighborhood of the smallest PI widths observed for message variables when moderators are taken into account). In short, fully reconstructable message variables naturally produce relatively small amounts of ES variation compared to that generated by less reconstructable message variables.

Understanding the causes of ES variation is important as it directly influences accurate prediction. In the case of

persuasive message variables, one wants to be able to predict whether making a given message design choice will yield a more persuasive message. Especially if the true ES is small, large variations in ESs makes that a problematic undertaking. Such variations decrease the chances of replicating previous findings (even if there is a genuine effect in the population) and makes it more difficult to establish a true non-zero mean (population) effect. It also limits the ability to provide sound advice to message designers as the application of a message variable may backfire. One way forward is to identify moderating conditions under which the effect of the message variable is larger and/or more consistent. But the findings in this paper point to another promising avenue, namely, defining message variables more clearly in order to reduce the variation in how a given message variable has been manipulated.

Studying unreconstructable message variables

Most of the message variables reviewed here are unreconstructable (19/30, 63%). Given that as effect-size variability increases as reconstructability diminishes, one might be tempted to conclude that unreconstructable message variables are intrinsically defective, or at least somehow less desirable than more reconstructable variables. After all, as noted above, unreconstructable message variables pose special challenges for developing dependable generalizations and useful message design guidance.

But it would be a mistake to think that there is something inherently wrong with defining a message variable in a way that makes it unreconstructable. An unreconstructable message variable can be perfectly sensible—and manifestly relevant to practical message-design decisions. For example, "Should our ad be funny or serious?" is a question that might naturally arise during a message design process, and hence it will be valuable to have some understanding of the effects associated with the humorous-versus-non-humorous message contrast—even though that contrast is unreconstructable.

At the same time, researchers should understand the consequences of studying unreconstructable message variables. With such variables, the door is open to many different concrete realizations of the message contrast. Studies of message vividness provide a convenient example. As Blondé and Girandola (2016, p. 112) explained, "in experimental studies, vividness has been manipulated using many operationalizations including the presence (vs. absence) of pictures, concrete (vs. abstract) words, concrete (vs. abstract) pictures, narratives (vs. statistical) evidences, and direct (vs. indirect) transmission of information."

So, faced with some unreconstructable message variable, one way in which researchers might try to make some progress in sorting out message effects is by decomposing the abstract variable into its more concrete (and reconstructable) operationalizations—and then focusing on understanding the effects of those more specific message elements. In the case of vividness, for example, it might not be possible to provide a precise definition given its motley set of operationalizations. But it could be feasible to provide a more precise definition of, for instance, what concrete language entails, and to subsequently assess the difference in persuasive impact of a more concretely worded message versus a more abstractly worded one. Pursuing such a course will require close attention to the particular experimental realizations of the abstract message contrast and correspondingly careful analysis of the varieties thereof (for an example, see Saucier & Walter, 2021). This

will not always be easy, but will be worth the effort—for two reasons: It will provide the basis for a more fine-grained understanding of persuasion processes, and it will make it easier to provide guidelines for message designers (because advice of the form "make your message more vivid" can be replaced by more concrete suggestions).

A second strategy for making progress in understanding unreconstructable message variables is to assess the strength of the manipulation (see also Linden & Hönekopp, 2021, p. 371). Typically, message manipulations are presented as a dichotomous variable, such as vivid versus pallid. However, studies of such message contrasts can vary in the strength of the manipulation; for example, the vivid message may be only slightly more, or much more, vivid than the pallid one.

Assessing the strength of the manipulation may bring two important advantages. First, the size of the effect of a message variable (the size of the difference in persuasiveness between the two message forms) can be expected to be sensitive to the strength of the message manipulation. If vividness has a persuasive effect, a larger vividness contrast may have a larger effect than a smaller vividness contrast. Assessing differences in manipulation strength between studies might provide a basis for explaining at least some of the variability in ESs.

Second, variability in mediating states can be expected to be sensitive to the strength of the message manipulation. Consider, for example, studies of threat appeals where the messages differ in the depicted severity of the negative consequences. The degree to which aroused fear (a mediating state) varies between such message conditions may vary depending on the strength of message manipulation. Stronger manipulations thus may ease the task of identifying relevant mediating states and thereby accelerate progress toward explanations of observed effects.

In any case, the larger point to be appreciated is that an unreconstructable message variable does not represent a dead end for analysis. On the contrary, such message variables represent an open invitation to further study.

Theorizing message variables
Message variables and theories of message effects

When thinking about the theory-related aspects of studying persuasive message variables, attention naturally first turns to theories about the *effects* of message variables. For example, the elaboration likelihood model (ELM; Petty & Cacioppo, 1986) offers a theoretical framework for understanding how the effects of argument strength can vary depending on the level of moderating factors such as involvement. Similarly, the extended parallel process model (EPPM; Witte, 1992) offers a treatment of the effects of fear-appeal variations.

But such theories of message effects require evidence, and one contribution of the present analysis is to offer some clarification about some properties of that evidence. Developing sound theory about an unreconstructable message variable presents a different set of evidentiary challenges than those for a fully reconstructable message variable. Where theoretical claims are advanced about unreconstructable message variables, one should expect the relevant research evidence to be rather "noisy," because of the greater associated variability in ESs. Noise obscures underlying causal mechanisms, thus impairing theoretical progress. But with a clearer understanding of how and why some message variables might naturally generate such noise, researchers will have a better understanding of what will be required in order to make theoretical

progress in understanding message effects, especially where unreconstructable message variables are of interest.

And although the focus of the present analysis has been message variables studied for their roles in *persuasive* communication, there is no reason to assume that the value of concepts such as reconstructability is limited to that domain. So, for example, studies that focus on the effects of news-media message variables or entertainment-media message variables might usefully consider the degree to which the variables being examined are reconstructable. Where a given message variable is unreconstructable, one should expect to see greater variability of ESs, with correspondingly greater challenges in developing sound generalizations that can inform theories of message effects.

Message variables as objects of theoretical attention

The analysis offered here also is a tentative first step along a new theoretical avenue, perhaps best described by contrast to the approach represented by familiar theories of message effects such as the ELM and EPPM. These sorts of theories are theoretical treatments of why this or that message variable might be expected to have this or that effect; these theories concern the effects of various specific message variables.

In contrast, the approach taken here does not involve theorizing about the effects of any specific message variable, but rather theorizing about the nature of message variables themselves: the properties of message variables, how message variables differ from one another, and so on. The property discussed here—reconstructability—is a dimension along which all message variables vary; thus reconstructability potentially represents one piece of a general theoretical analysis of message-variable attributes.

As another—speculative—example of a dimension along which persuasive message variables might differ, consider (what might be called) directionality. For some message variables, the direction of difference between the two forms might be expected to be relatively constant. For example, strong-argument messages might be expected to consistently be more persuasive than weak-argument messages; the size of the persuasive advantage of strong-argument messages might vary, but the direction of difference could be constant. In contrast, for other message variables, the direction of difference between the two forms might be expected to vary. Consider, for example, argument explicitness, variation in the extent to which the components of the message's arguments (e.g., premises) are made explicit. It might be the case that greater explicitness enhances persuasiveness under some conditions (perhaps when the arguments are strong) but diminishes it in other cases (as when the arguments are weak).

In a sense, the nature of message variables is a remarkably undertheorized domain. Communication scholars have extensive theoretical equipment for describing *messages*. A question such as "How are these two persuasive messages similar and different?" can be answered with familiar conceptual apparatus: both messages are narratives, one is gain-framed and the other loss-framed, both are two-sided, and so on.

But there is not such theoretical equipment for describing message *variables*. A question such as "How are these two message variables similar and different?" does not immediately suggest possible answers. Individual message variables are commonly treated as distinctive objects, sui generis, not objects to be compared. Extant theoretical treatments of the effects of specific message variables are not conceptual

frameworks for thinking generally about the properties of message variables.

The approach taken here focuses on message variables as themselves objects of theoretical attention; it represents a first-pass effort at unpacking properties of message variables that might be helpful in illuminating differences between message variables. Thus the present analysis points to the possibility that a broader theoretical analysis of message variables themselves might be valuable. For all that communication researchers are interested in message variables, there has been curiously little attention given to abstract analyses of the general properties of message variables—properties such as reconstructability.

Conclusion

Message variables differ in their reconstructability—the degree to which one of the two messages in a pair can be deduced from the other. Examination of ESs associated with persuasive message variables reveals that as message variables become less reconstructable, the variability of the associated ESs increases—with consequent challenges for research progress and practical message design. Unreconstructable message variables thus require special attention in future research. More broadly, attention to message-variable properties such as reconstructability may shed light on how and why effects differ across message variables.

Supplementary material

Supplementary material is available online at *Human Communication Research*

Data availability

The data underlying this article are available in the article and in its online supplementary material.

Author bio

Hans Hoeken is a professor of Communication & Information Studies in the Department of Languages, Literature, and Communication at Utrecht University; ORCID 0000-0002-7535-1273. Daniel J. O'Keefe is the Owen L. Coon Professor in the Department of Communication Studies at Northwestern University; ORCID 0000-0003-1594-8892.

Notes

1. The present focus is thus on what Linden and Hönekopp (2021) call heterogeneity: "the variability in population effect sizes" (p. 361). As they emphasize, this is something different from the property captured in "heterogeneity" measures such as I^2.

2. T is the estimate of tau (τ), the (unobservable) standard deviation of the true ESs. T^2 is the sample estimate of tau-squared (τ^2), the (unobservable) variance of the true ESs.
3. In contrast, Rains et al.'s (2018) review of 37 "persuasion" meta-analyses included meta-analyses of (e.g.) the effects of forewarning, the effects of distraction, and factors influencing judgments of communicator credibility. These are certainly relevant to research questions about persuasion, but they do not concern the persuasive effects of message variables.
4. O'Keefe's (1999b) review addressed persuasion ES heterogeneity, but it relied on I^2. As explained above, I^2 is not an index of the absolute variability of ESs.
5. T is expressed in a metric corresponding to the metric used for the ESs (Borenstein et al., 2021, p. 112). The ES index used by O'Keefe and Hoeken (2021) was r (correlation), so T is reported here in Fisher Z units.
6. We considered conducting significance tests (to compare the values of T and PI for different reconstructability categories) but realized such tests would be otiose. With so few cases (e.g., only five message variables classified as semi-reconstructable) any significance tests would be guaranteed to return a nonsignificant result. Candor thus compels us to say that one cannot rule out randomness as an explanation for the observed numerical differences.
7. The T values in the present report are expressed in Fisher Z units. Linden and Hönekopp (2021) reported T values expressed in d units (standardized mean differences), so their Ts have been converted to Fisher Z units for the present comparisons.

References

Allen, M. (1991). Meta-analysis comparing the persuasiveness of one-sided and two-sided messages. *Western Journal of Speech Communication*, 55(4), 390–404. https://doi.org/10.1080/10570319109374395

Allen, M., & Preiss, R. W. (1997). Comparing the persuasiveness of narrative and statistical evidence using meta-analysis. *Communication Research Reports*, 14(2), 125–131. https://doi.org/10.1080/08824099709388654

Birge, R. T. (1932). The calculation of errors by the method of least squares. *Physical Review*, 40 (2nd ser.), 207–227. https://doi.org/10.1103/PhysRev.40.207

Blondé, J., & Girandola, F. (2016). Revealing the elusive effects of vividness: A meta-analysis of empirical evidences assessing the effect of vividness on persuasion. *Social Influence*, 11(2), 111–129. https://doi.org/10.1080/15534510.2016.1157096

Bolkan, S., & Rains, S. A. (2017). The legitimization of paltry contributions as a compliance-gaining technique: A meta-analysis testing three explanations. *Communication Research*, 44(7), 976–996. https://doi.org/10.1177/0093650215602308

Borenstein, M. (2019). *Common mistakes in meta-analysis and how to avoid them*. Biostat.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2021). *Introduction to meta-analysis* (2nd ed.). Wiley.

Borenstein, M., Higgins, J. P. T., Hedges, L. V., & Rothstein, H. R. (2017). Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Research Synthesis Methods*, 8, 5–18. https://doi.org/10.1002/jrsm.1230

Brugman, B. C., Burgers, C., & Vis, B. (2019). Metaphorical framing in political discourse through words vs. concepts: A meta-analysis. *Language and Cognition*, 11(1), 41–65. https://doi.org/10.1017/langcog.2019.5

Bryan, C. J., Tipton, E. & Yeager, D. S. (2021). Behavioural science is unlikely to change the world without a heterogeneity revolution. *Nature Human Behaviour*, 5, 980–989. https://doi.org/10.1038/s41562-021-01143-3

Card, N. A. (2012). *Applied meta-analysis for social science research*. Guilford.

Carpenter, C. J. (2013). A meta-analysis of the effectiveness of the "but you are free" compliance-gaining technique. *Communication Studies*, 64(1), 6–17. https://doi.org/10.1080/10510974.2012.727941

Carpenter, C. J. (2015). A meta-analysis of the ELM's argument quality × processing type predictions. *Human Communication Research*, 41(4), 501–534. https://doi.org/10.1111/hcre.12054

Carpenter, C. J., & Boster, F. J. (2009). A meta-analysis of the effectiveness of the disrupt-then-reframe compliance gaining technique. *Communication Reports*, 22(2), 55–62. https://doi.org/10.1080/08934210903092590

De Hoog, N., Stroebe, W., & de Wit, J. (2007). The impact of vulnerability to and severity of a health risk on processing and acceptance of fear-arousing communications: A meta-analysis. *Review of General Psychology*, 11(3), 258–285. https://doi.org/10.1037/1089-2680.11.3.258

Dillard, J. P. (1998). Evaluating and using meta-analytic knowledge claims. In M. Allen & R. W. Preiss (Eds.), *Persuasion: Advances through meta-analysis* (pp. 257–270). Hampton Press.

Gayle, B. M., Preiss, R. W., & Allen, M. (1998). Another look at the use of rhetorical questions. In M. Allen & R. W. Preiss (Eds.), *Persuasion: Advances through meta-analysis* (pp. 189–201). Hampton Press.

Hamilton, M. A., & Hunter, J. E. (1998). The effect of language intensity on receiver evaluations of message, source, and topic. In M. Allen & R. W. Preiss (Eds.), *Persuasion: Advances through meta-analysis* (pp. 99–138). Hampton Press.

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21, 1539–1558. https://doi.org/10.1002/sim.1186

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, 327, 557–560. https://doi.org/10.1136/bmj.327.7414.557

Hoeken, H., Hornikx, J., & Linders, Y. (2020). The importance and use of normative criteria to manipulate argument quality. *Journal of Advertising*, 49(2), 195–201. https://doi.org/10.1080/00913367.2019.1663317

Hornikx, J., & O'Keefe, D. J. (2009). Adapting consumer advertising appeals to cultural values: A meta-analytic review of effects on persuasiveness and ad liking. *Annals of the International Communication Association*, 33(1), 39–71. https://doi.org/10.1080/23808985.2009.11679084

Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or I^2 index? *Psychological Methods*, 11, 193–206. https://doi.org/10.1037/1082-989X.11.2.193

IntHout, J., Ioannidis, J. P. A., Rovers, M. M., & Goeman, J. J. (2016). Plea for routinely presenting prediction intervals in meta-analysis. *BMJ Open*, 6, e010247. https://doi.org/10.1136/bmjopen-2015-010247

Kenny, D. A., & Judd, C. M. (2019). The unappreciated heterogeneity of effect sizes: Implications for power, precision, planning of research, and replication. *Psychological Methods*, 24(5), 578–589. https://doi.org/10.1037/met0000209

Lau, R. R., Sigelman, L., & Rovner, I. B. (2007). The effects of negative political campaigns: A meta-analytic reassessment. *Journal of Politics*, 69(4), 1176–1209. https://doi.org/10.1111/j.1468-2508.2007.00618.x

Lee, S., & Feeley, T. H. (2016). The identifiable victim effect: A meta-analytic review. *Social Influence*, 11(3), 199–215. https://doi.org/10.1080/15534510.2016.1216891

Lee, S., Moon, S.-I., & Feeley, T. H. (2019). The 'that's-not-all' compliance-gaining technique: When does it work? *Social Influence*, 14(2), 25–39. https://doi.org/10.1080/15534510.2019.1634146

Levine, T. R., & Weber, R. (2020). Unresolved heterogeneity in meta-analysis: Combined construct invalidity, confounding, and other challenges to understanding mean effect sizes. *Human Communication Research*, 46(2-3), 343–354. https://doi.org/10.1093/hcr/hqz019

Linden, A. H., & Hönekopp, J. (2021). Heterogeneity of research results: A new perspective to assess and promote progress in psychological science. *Perspectives on Psychological Science*, 16(2), 358–376. https://doi.org/10.1177/1745691620964193

Lull, R. B., & Bushman, B. J. (2015). Do sex and violence sell? A meta-analytic review of the effects of sexual and violent media and ad content on memory, attitudes, and buying intentions. *Psychological Bulletin*, 141(5), 1022–1048. https://doi.org/10.1037/bul0000018

Miller, N., Maruyama, G., Beaber, R. J., & Valone, K. (1976). Speed of speech and persuasion. *Journal of Personality and Social Psychology*, 34(4), 615–624. https://doi.org/10.1037/0022-3514.34.4.615

O'Keefe, D. J. (1998). Justification explicitness and persuasive effect: A meta-analytic review of the effects of varying support articulation in persuasive messages. *Argumentation and Advocacy*, 35(2), 61–75. https://doi.org/10.1080/00028533.1998.11951621

O'Keefe, D. J. (1999a). How to handle opposing arguments in persuasive messages: A meta-analytic review of the effects of one-sided and two-sided messages. *Annals of the International Communication Association*, 22(1), 209–249. https://doi.org/10.1080/23808985.1999.11678963

O'Keefe, D. J. (1999b). Variability of persuasive message effects: Meta-analytic evidence and implications. *Document Design*, 1(2), 87–97. https://doi.org/10.1075/dd.1.2.02oke

O'Keefe, D. J. (2002). The persuasive effects of variation in standpoint articulation. In F. H. van Eemeren (Ed.), *Advances in pragma-dialectics* (pp. 65–82). Sic Sat.

O'Keefe, D. J., & Hoeken, H. (2021). Message design choices don't make much difference to persuasiveness and can't be counted on—not even when moderating conditions are specified. *Frontiers in Psychology*, 12, 2533. https://doi.org/10.3389/fpsyg.2021.664160

O'Keefe, D. J., & Jensen, J. D. (2006). The advantages of compliance or the disadvantages of noncompliance? A meta-analytic review of the relative persuasive effectiveness of gain-framed and loss-framed messages. *Annals of the International Communication Association*, 30(1), 1–43. https://doi.org/10.1080/23808985.2006.11679054

Partlett, C., & Riley, R. D. (2017). Random effects meta-analysis: Coverage performance of 95% confidence and prediction intervals following REML estimation. *Statistics in Medicine*, 36, 301–317. https://doi.org/10.1002/sim.7140

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. Springer-Verlag.

Preiss, R. W., Allen, M., Gayle, B., & Kim, S.-Y. (2014, November). *Meta-analysis of the relationship between rate of delivery and message persuasiveness: Linear versus curvilinear tests*. [Paper presentation]. National Communication Association annual meeting, Chicago, IL.

Rains, S. A., Levine, T. R., & Weber, R. (2018). Sixty years of quantitative communication research summarized: Lessons from 149 meta-analyses. *Annals of the International Communication Association*, 42(2), 105–124. https://doi.org/10.1080/23808985.2018.1446350

Rothman, A. J., Bartels, R. D., Wlaschin, J., & Salovey, P. (2006). The strategic use of gain- and loss-framed messages to promote healthy behavior: How theory can inform practice. *Journal of Communication*, 56(S1), S202–S220. https://doi.org/10.1111/j.1460-2466.2006.00290.x

Rücker, G., Schwarzer, G., Carpenter, J. R., & Schumacher, M. (2008). Undue reliance on I^2 in assessing heterogeneity may mislead. *BMC Medical Research Methodology*, 8, 79. https://doi.org/10.1186/1471-2288-8-79

Saucier, C. J., & Walter, N. (2021). Dissecting a frog: A meta-analytic evaluation of humor intensity in persuasion research. *Annals of the International Communication Association*, 45(4), 258–283. https://doi.org/10.1080/23808985.2022.2033634

Seo, K. (2020). Meta-analysis on visual persuasion: Does adding images to texts influence persuasion? *Athens Journal of Mass Media and Communications*, 6(3), 177–190. https://doi.org/10.30958/ajmmc.6-3-3

Seo, K., & Kim, N. Y. (2018, May). *Does adding images to texts influence persuasion? A meta-analysis of visual image effects on persuasive texts* [Paper presentation]. International Communication Association annual meeting, Prague, Czech Republic.

Shen, F., Sheer, V. C., & Li, R. (2015). Impact of narratives on persuasion in health communication: A meta-analysis. *Journal of Advertising*, 44(2), 105–113. https://doi.org/10.1080/00913367.2015.1018467

Stiff, J. B. (1985). *Cognitive processing of persuasive message cues: A meta-analytic review of the effects of supporting information on attitudes* [Doctoral dissertation, Michigan State University]. ProQuest no. 8520567.

Stiff, J. B. (1986). Cognitive processing of persuasive message cues: A meta-analytic review of the effects of supporting information on attitudes. *Communication Monographs*, 53(1), 75–89. https://doi.org/10.1080/03637758609376128

Vann, J. W., Rogers, R. D., & Penrod, J. P. (1987). The cognitive effects of time-compressed advertising. *Journal of Advertising*, 16(2), 10–19. https://doi.org/10.1080/00913367.1987.10673072

Van 't Riet, J., Cox, A. D., Cox, D., Zimet, G. D., de Bruijn, G.-J., Van den Putte, B., de Vries, H., Werrij, M. Q., & Ruiter, R. A. (2016). Does perceived risk influence the effects of message framing? Revisiting the link between prospect theory and message framing. *Health Psychology Review*, 10(4), 447–459. https://doi.org/10.1080/17437199.2016.1176865

Walter, N., Cody, M. J., Xu, L. Z., & Murphy, S. T. (2018). A priest, a rabbi, and a minister walk into a bar: A meta-analysis of humor effects on persuasion. *Human Communication Research*, 44(4), 343–373. https://doi.org/10.1093/hcr/hqy005

Weber, R., & Popova, L. (2012). Testing equivalence in communication research: Theory and application. *Communication Methods and Measures*, 6(3), 190–213. https://doi.org/10.1080/19312458.2012.703834

Wheeless, L. R. (1971). Some effects of time-compressed speech on persuasion. *Journal of Broadcasting*, 15(4), 415–420. https://doi.org/10.1080/08838157109363661

White, B. X., & Albarracín, D. (2018). Investigating belief falsehood: Fear appeals do change behavior in experimental laboratory studies: A commentary on Kok et al. (2018). *Health Psychology Review*, 12(2), 147–150. https://doi.org/10.1080/17437199.2018.1448292

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communication Monographs*, 59(4), 329–349. https://doi.org/10.1080/03637759209376276

Witte, K., & Allen, M. (2000). A meta-analysis of fear appeals: Implications for effective public health programs. *Health Education and Behavior*, 27(5), 591–615. https://doi.org/10.1177/109019810002700506