

THE MESSAGES REPLICATION FACTOR: METHODS TAILORED TO MESSAGES AS OBJECTS OF STUDY

By Sally Jackson, Daniel J. O'Keefe, and Dale E. Brashers



In research on effects of message variables, it is generally necessary to examine responses to actual messages that represent, embody, or instantiate the values of the variable of interest. Researchers have lately become attentive to problems of confounding in the use of individual concrete messages to represent abstract theoretical contrasts, and replicated treatment comparisons are increasingly common in communication research. How to treat the replication factor in the statistical analysis remains controversial. Whether to treat replication factors as fixed or as random hinges on what is assumed about the relationship between abstract treatment contrasts and their concrete material implementations. We argue that reflection on this relationship justifies a general policy of treating replications as random. Two circumstances in which fixed-effects analyses might seem attractive (the case of matched-message designs and the case of experimental manipulations occurring outside of messages) are considered, but it is concluded that these situations also require random-effects analyses.

To do empirical research on effects of message variables, it is generally necessary to examine responses to actual messages that represent, embody, or instantiate the values of the variable of interest. The adequacy of actual concrete messages as instantiations of variables is central to any assessment of the validity of such an experiment. During the long history of experimental message effects research, virtually no attention has been paid to this issue. The seminal studies of message effects conducted by Hovland and associates during and after World War II set a precedent for how to deal with "operationalization" of message variables that has been essentially unchallenged within communication and social psychology.

Only very recently have communication scholars begun to question the validity of the Yale School's basic experimental design choices. The main thrust of this recent critique is that the standard experimental designs built on the Yale School precedent were never adequately tailored to the special properties of messages as objects of study.¹

Our broad objective is to develop an analytic approach that is tailored to the nature of messages and to their role within message effects experiments. This tailoring involves many adjustments to off-the-rack research procedures, of which the most important is the incorporation of "message replications" into the design of many experiments. Adjusting experimental

designs to incorporate multiple message replications, however, raises the issue of whether to consider the messages replication factor (or any similar factor) as “fixed” or “random.”

One line of reasoning would suggest that the messages replication factor should generally be treated as random, since the specific messages used in any experiment are generally only a sample from a large and possibly limitless universe of possible messages.² But it might also be reasoned that the messages replication factor should be treated as fixed, because the specific messages used in an experiment are not in fact randomly selected and because treating messages as random seems to offer low power to detect treatment effects.³

In what follows, we lay out a systematic analysis of different experimental designs for message effects research and discuss the advantages of designs that include replications. We then consider the issue of fixed- vs. random-effects analysis, and argue that replication factors should be treated as random except in very unusual circumstances. We then examine in detail two situations in which fixed-effects analyses might seem attractive, but conclude that these situations do not justify exceptions to a general policy of treating replications as random.

Experimental designs for message effects research can be characterized usefully and clearly by recognizing four basic possibilities arising from two contrasts. The first contrast is between replicated and unreplicated designs: designs in which treatment levels have multiple embodiments versus designs in which each treatment level has a single embodiment. The second contrast is between matched and unmatched designs: designs in which individual messages or other materials arise from manipulation of a controlled template versus designs in which individual messages or other materials are selected “intact” to represent each treatment level.⁴ The four basic experimental designs are thus: (1) the unreplicated unmatched design; (2) the unreplicated matched design; (3) the replicated unmatched design; and (4) the replicated matched design. The relationships among the four designs are summarized in Figure 1.

For simplicity of exposition, we assume throughout this discussion that there is only one treatment variable, with just two levels. What is matched or unmatched between treatment levels will be understood to be experimental messages; it is also possible, of course, to have designs in which sources, situations, confederates, or other sampled materials appear, either matched or unmatched from treatment level to treatment level.

Each of these basic designs can occur either as an independent groups (between-subjects) design or as a repeated measures (within-subjects) design, and of course each can involve more than one treatment factor. These distinctions are well-understood in communication research and need not enter into the present discussion; for the arguments to be made below, it does not matter whether there are additional treatment factors or whether the design includes repeated measures.⁵

The Four Designs. In the simplest case, an unreplicated unmatched messages design involves a comparison between two individual messages, each chosen intact to represent one level of the treatment variable. For example, an investigation of the effectiveness of comparative and noncomparative advertising might select one existing comparative ad and another existing noncomparative ad (even for an entirely different product) as the messages to be used.

Experimental Designs: A Systematic Analysis

FIGURE 1
Four Design Prototypes

	messages sampled intact	messages sampled as templates
M = 1	unreplicated unmatched design	unreplicated matched design
M > 1	replicated unmatched design	replicated matched design

Note: M is the number of distinct message replications.

An unreplicated matched design for the same research question would compare two messages that were versions of the same basic "template message." Although unreplicated unmatched designs are nowadays uncommon, the unreplicated matched design remains standard in message effects research.⁶ A given advertisement would be prepared in two forms (one comparative, one noncomparative), perhaps by the expedient of having a portion of the ad contain either comparative claims or noncomparative claims. For convenience, we will use the term "treatment segments" to refer to the alternative contents inserted into the template to produce the contrasting versions and the term "treatment space" to refer to the slot into which they are inserted. Apart from the treatment segments used to introduce the contrast of interest, the two ads would be identical (hence, matched). Notice that templates can consist of text alone, of text plus context, or of context alone.

A replicated unmatched design would involve more than one message for each treatment level, but these messages would not be matched (paired) across treatments.⁷ For instance, one might collect a large number of existing comparative ads and an equally large number of existing noncomparative ads (thus affording multiple replications within each treatment level).

A replicated matched design for this question would use multiple sets of matched messages; that is, the entire configuration of experimental materials (template and contrasting treatment segments) would be replicated. Rather than having just one template advertisement prepared in both versions, this design would have multiple template advertisements each prepared in both versions.⁸

Why Replicated Designs are Preferable. Replication is desirable in both matched and unmatched designs as a strategy for controlling confounds. Unreplicated designs of both types invite confounding between the abstract theoretical contrasts of interest and the uninteresting content particulars of individual messages.⁹

It should be obvious that in an unreplicated unmatched design, where the abstract contrast is between two message categories, only very rarely will it be reasonable to draw conclusions based on two arbitrary examples chosen to represent the two categories. For example, comparing one corporate "advertorial" with one standard editorial will certainly not suffice to show whether advertorials are as credible or effective as standard editorials; the two pieces may differ in many substantive respects that may have powerful influence on judgments. This problem has been termed "case-category confounding."¹⁰ The confounding of the particular case (the particular

example, the concrete text used) and the abstract treatment category means that properties of the cases examined are not analytically separable from differences among categories.

In an unreplicated matched design, many content features are held constant, and some might suppose for this reason that confounding is not a problem. But even matched designs offer two distinct opportunities for content confounds. First, the contrasting treatment segments can differ in any number of respects other than the abstract contrast of interest—a problem of “surplus variation.”¹¹ Second, the “controlled” template is linguistic material (words, sentences), and there is no guarantee that its meaning will remain unchanged as treatment segments are inserted—a problem of “gestalt effects.”¹²

But controlling for confounds is not the only reason to replicate. Even if materials could be equated perfectly from treatment level to treatment level, it would still be possible that the effect of interest could vary from one instance of the treatment to another.¹³ For example, there is very strong reason to think that the effect of source credibility varies from one message to another, not only in size but also in direction.¹⁴ Whenever it is possible that the effect of interest is itself variable, replication serves the dual purpose of improving the estimate of the effect and permitting a direct estimate of the uniformity or stability of the effect.

Given the advantages of replicated designs, it is unsurprising that these have become more prominent in communication research.¹⁵ But what form of statistical analysis suits such designs remains quite uncertain. The basic data-analytic choice is whether to consider the messages replication factor (or any other similar factor) as “fixed” or as “random.”¹⁶

In our view, the more defensible general policy is to treat replications as random except in very unusual cases. The reasons for including replications in communication experiments have implications for choices made in statistical analysis. The central issue in analysis (whether to treat replications as fixed or as random) depends on the same considerations as the central issue in design (whether to replicate or not). In general, the reasoning that leads to a decision to replicate should also lead to analysis in which replications are treated as a random factor.

There has already been substantial discussion of this question. Some researchers appear to assume that treating replications as random precludes such standard analytic steps as testing for treatment \times replication interactions or conducting followup analysis of such interactions to locate “moderators” of treatment effects.¹⁷ But this assumption is unfounded. The usual mixed model *F* test for the interaction of a fixed and a random factor is the same as the test one would conduct if both factors were considered fixed; and the recognition that messages may be a source of random variability does not in any way preclude the search for systematic sources of variability.¹⁸ The power of random-effects analysis to detect treatment effects has also been questioned.¹⁹ But this potentially serious pitfall has proven on close scrutiny to offer little basis for preferring a treatment of replications as fixed: in the first place, treating replications as fixed tests a null hypothesis other than what is usually wanted, and in the second place, the power of tests with replications as random can be greatly improved with relatively modest increases in the number of replications.²⁰

In some cases – for example, in the unmatched replicated design –

Analyzing Replicated Designs

there is broad consensus in favor of treating replications as random. Here we wish to discuss two circumstances in which it might be thought that treating message replications as a fixed factor is justified: the first is when message replications are matched across treatments, so that all but the treatment components of the message are held constant from one treatment condition to another; the other is when the experimental manipulation lies outside the message proper, so that the entire message is apparently held constant from one treatment condition to another. In each of these circumstances, the effect of messages on outcomes may appear to have been neutralized by holding constant all of the message content exclusive of the treatment itself. But we will argue that even in these circumstances the appropriate data-analytic decision, with very few exceptions, is to treat message replications as random, not fixed. Only in very unusual circumstances will a decision to treat replications as fixed be justified.

Matched-Replication Designs. One circumstance in which one might have doubts about the necessity of a random-effects analysis is when replications have been matched carefully from one treatment condition to another. Specifically, it might be suggested that (a) matching eliminates confounding, and (b) matching assures that the specific messages used will be inconsequential so far as measurement of the treatment effect is concerned. Hence (this reasoning runs) one can safely treat message replications as a fixed factor in the data analysis.²¹

But both of the presumptions behind this reasoning have already been seen to be defective. First, matching does not eliminate confounding, since even though a message template may be held constant, the treatment segments may introduce incidental content variations ("surplus variation," discussed above), and since the insertion of the treatment segments into a treatment space may alter the meaning of the supposedly fixed template ("gestalt effects," discussed above). Second, matching does not assure that the template will be "neutralized." If templates had only "main effects" on the dependent variable, this might be true, but since templates may also interact with treatments, the notion that fixing the template neutralizes its impact on the dependent variable is simply not tenable. To be sure, matching may reduce the magnitude of the variation associated with message replications (just as repeated-measures designs improve on independent-groups designs), but matching does not eliminate the variation and does not obviate the need to take this variation into account in testing the treatment effect (any more than in the parallel case of repeated-measures designs).

In fact, treating matched replications as fixed has undesirable statistical consequences. These consequences have been discussed thoroughly elsewhere, so only a brief summary will be offered here.²² In a replicated matched design, where templates are sampled and subjected to treatment, an analysis treating replications as fixed will lead to a test of the treatment effect that is biased to a degree determined by the size of the treatment x replication interaction effect and by the study size. Specifically, the more variable a treatment effect from one replication to the next (i.e., the bigger the interaction), the more biased the test; the more respondents given a fixed number of replications, the more biased the test; and the fewer replications given a fixed number of respondents, the more biased the test.

This bias translates, when the null hypothesis is true, into an uncontrolled increase in the Type I error rate (alpha inflation). When the net effect of the treatment is nil, but incidental content variations from replication to replication introduce variations both positive and negative around the true

effect, an analysis treating replications as fixed will lead to too many rejections of the null hypothesis. What this means, substantively, is that when a significant treatment effect is reported for a fixed-model test applied to a replicated matched design, there actually is not a good basis for rejecting the null hypothesis, because that test can have a very high probability of achieving significance even if the null hypothesis is true.

Hence whether replications are matched or unmatched, the appropriate general data-analytic policy is consistently to treat replications as random. This policy is justified by the possibility that the specific replications selected will contribute uncontrolled variation to the estimate of the treatment effect.

Extratextual Manipulations. Another circumstance in which one might have doubts about the necessity of a random-effects analysis of message replications is in a special case of matching defined by the use of what we will call "extratextual" experimental manipulations. By extratextual manipulations, we mean manipulations that occur outside the message itself, as when the experimental treatment involves a change in the situation in which a fixed message is delivered or a change in the source or medium of the message. In such cases, where messages appear to "serve merely as the context for the manipulations that operationalize the experimental treatment," one might suppose that a fixed effects analysis would be quite reasonable.²³

In the language introduced earlier for differentiation among various design types, this form of manipulation involves a template consisting of fixed text and some fixed extratextual features, with the treatment space located among the extratextual features. To use an example of Michael Slater's, if we were to "present the same message as being published in the *Washington Post* or the *People's Daily*,"²⁴ the template would consist of the message and setup, the treatment space would be a source identification slot, and the treatment segments would be the contrasting source attributions. The suggestion to be considered in this section is that when such manipulations occur in a replicated design, the replication factor may properly be considered a fixed factor.

The case for treating message replications as fixed in such designs rests on the same two presumptions as before: that when the message is exactly the same from one level of the treatment to the other, no problems of confounding arise, and that if the message serves only as a context for the manipulation, it does not contribute to the effect of the manipulation. Each of these presumptions has already been found questionable. The extratextuality of the treatment space is no security against incidental confounds in the treatment segments, nor does it safeguard the treatment contrast from variability associated with the selection of particular templates.

Indeed, whether a manipulation occurs intratextually or extratextually is entirely irrelevant to the issue of whether replications should be treated as fixed or random, as can be seen by considering a series of related examples. Imagine first a replicated matched design built on the plan described above for the study of credibility effects. A number of texts are selected or created, and *each one* is attributed alternately to two different sources chosen to represent higher and lower credibility. For each independent replication, it may be said that the message serves only as context for the manipulation; for each message replication, the contrasting source attributions embody the treatment variable (credibility). On the proposal that messages be treated as a fixed effect if the manipulation occurs extratextually, this design would be

analyzed as a standard two-way factorial with both factors (credibility and replications) treated as fixed.

Now consider a second design for this study, created through a minor variation in the example. Suppose that instead of locating the credibility manipulation outside the text, the message content itself is altered to include reference to the source of the message's claims. Neither the design nor its ability to support a conclusion has changed, but the rationale for treating messages as fixed has vanished, for the manipulation is no longer extratextual. A variation of this sort *should* make no difference to decisions about analysis, from which it should be apparent that whether a manipulation appears intratextually or extratextually really does not affect the question of how the design should be analyzed. What affects that question is whether there is any opportunity for confounding (and in both cases there is) and whether there is any reason to suppose that the effect of interest varies from one implementation to another (and again, in both cases there is).

To make the same point from a different approach, imagine a third design for this study, one that keeps the credibility manipulation clearly distinct from the messages. Suppose a large number of texts are selected and randomly divided into two interchangeable sets. Half are presented as originating with diverse low credibility sources and the other half as originating with diverse high credibility sources. The design is replicated, but unmatched, so replications are nested within treatment levels. Following the argument that if the manipulation occurs extratextually, the messages may be treated as fixed, the analysis would involve fixed replications nested under fixed treatments, an analysis that is specifically ruled out by the random allocation of messages to treatment conditions. This third design, despite involving an extratextual manipulation of credibility, obviously requires that messages be regarded as random.²⁵

As this series of examples displays, for purposes of statistical analysis it is irrelevant whether a manipulation occurs as variation in text elements (intratextually) or as variation in non-text elements (extratextually); these forms of manipulation are structurally indistinguishable within an abstract representation of a design in terms of relationships among observations within the design. A separation between "the message" and "the manipulation" is a critical component of the view that where extratextual manipulations are used, the message functions only as an irrelevant background "context within which an experimental manipulation takes place."²⁶ Yet this separation, if one chooses to grant it at all, is plainly irrelevant to the question of whether experimental materials such as messages and their sources are contributors to random variation in experimental outcomes – and hence is irrelevant to the question of whether message replications should be treated as a fixed or random factor.

In short, the decision about whether to treat replications as fixed or random cannot be justified by whether the manipulation of an independent variable appears inside or outside a message. For this special case of matching, and for matched designs generally, message replications and other similar factors should almost always be considered random.

Legitimate Exceptions. Only in special and limited circumstances will it be reasonable to treat replication factors as fixed, and these circumstances have nothing to do with whether the replications are matched or unmatched or with whether the manipulations are intratextual or extratextual. Briefly, the legitimate exceptions all involve circumstances in which one's interest is only in learning whether one specific set of concrete stimuli differs from

another. We would say that it is reasonable to treat messages, speakers, situations, or other replication-like entities as fixed when interest centers on the particular cases selected for study, or when the cases selected are *not* substitutable.²⁷

A useful example is provided by an experiment on the effect of medium on perceptions of Reagan and Mondale, in which records of a presidential debate were presented in videotape, audiotape, or transcript form.²⁸ If the goal of the research is to understand how communication channels advantage or disadvantage particular candidates (so as to understand this particular historical event), both medium and candidate should be treated as fixed effects. What makes this experiment exceptional is that the candidates selected for study are not a sample from a large set of equally interesting cases, but are themselves the focus of study. In other words, treatment of messages or other such stimuli as fixed is most clearly justified when the research purpose is idiographic (focussed on a particular event such as a significant speech).

The general topic of the operationalization of message-related variables in communication research has received rather little explicit attention. To encourage some closer consideration of this topic, we wish to critically discuss two faulty assumptions about the operationalization of message variables. The first is the view that messages function within experiments as neutral context for important communicative processes; the second is the view that messages themselves are usefully thought of as operationalizing variables.

What we have said so far suggests that – no matter whether manipulations occur intratextually or extratextually, and no matter whether the design involves matched or unmatched messages – the particular messages to which experimental respondents are exposed can make a difference to the nature and size of the treatment effects of interest. Messages do not serve as a neutral backdrop against which abstract variables such as language intensity or source credibility play out their parts in communication processes. Experimental messages inevitably (if implicitly) represent classes of messages to which a treatment can be applied, and if there is any reason to worry that the treatment will have a varying effect from message to message, the messages actually selected cannot be considered either neutral context or self-sufficient as a context for demonstration of an effect.²⁹

In considering how concrete experimental messages are related to abstract message classes (variables), it can be tempting to think that the relationship is one of operationalization, that is, that concrete messages operationalize abstract message classes. After all, the idea of “operationalizing” a variable involves giving some concrete embodiment to that variable; so, for instance, a matched set of comparative and noncomparative ads might naturally be referred to as “the operationalization” of the comparative/noncomparative contrast.

But “operationalization” can be a misleading way of characterizing the relationship between concrete messages and abstract message variables. It can be misleading precisely because it can encourage one to think of the concrete messages as a neutral or self-sufficient context; it draws attention away from the role that the particular messages (or treatment segments) may play in determining the character of the treatment effects.

Our view is that the relationship of concrete messages to abstract

The Operationalization of Message Variables

message classes seems better captured by saying that *concrete messages do not operationalize, but only instantiate, abstract message variables*.³⁰ Whether messages are selected as examples within distinct categories (as in an unmatched design) or as templates into which a treatment contrast can be inserted (as in a matched design), the resulting concrete messages instantiate, but do not operationalize, the abstract contrasts of interest. The *procedure* (the operation) that produces the concrete messages might usefully be termed the “operationalization” of the treatment variable, but the *product* of that procedure should not. So, for example, in a matched-messages design, it is more accurate and more sensible to think of the message-transformation procedure as the operationalization of the treatment variable, and to think of the relevant treatment segments as instantiating or exemplifying the treatment variable.

Our interest here is not actually centered on the use of the term “operationalization.” Our interest is in how to understand the role of concrete messages in communication research (and, consequently, how to undertake statistical analysis of message data). We urge this way of thinking about concrete messages – as instantiations, not operationalizations, of abstract message variables – precisely so as to underscore the potential contributions of the concrete message materials to the observed treatment effects. Where messages are understood to be instances (examples, cases, exemplars), it is easier to see that the appropriate statistical analysis is one that treats these as random.

In both matched and unmatched designs, and in both intratextual and extratextual manipulations, messages are instances. Because the character of particular messages can influence the observed treatment effect, the messages cannot safely be treated as neutral “contexts” that can be ignored in analysis. In matched designs, treatment segments (be they text or non-text elements) have the same relationship to the treatment contrast as intact messages have to a treatment contrast in an unmatched design, namely, as instantiations (not operationalizations) that are representatives of a larger class, and hence require appropriate (random-effects) analysis. Where extratextual manipulations are used, the replicated message materials are nevertheless instantiations (not operationalizations) of a larger class, and hence require appropriate (random-effects) analysis. By thinking of concrete experimental message materials as instantiations, not operationalizations, one may be reminded of their potential contributions to observed effects – and hence reminded of the importance of statistical analyses that take advantage of the opportunity to estimate those contributions.

Conclusion

Our objective in this essay has been to further develop an analytic approach that is tailored to the nature of messages and their role within message effects experiments. No matter whether message replications are matched across conditions, and no matter whether experimental manipulations occur inside or outside replicated messages, the messages themselves (as parts or wholes) instantiate the variables of interest and are a source of variance in treatment effects.

As a result, unreplicated designs confound the abstract contrasts of interest with concrete messages or message segments, threatening the internal validity of any conclusion drawn about treatment effects. Replicated designs do not eliminate the extra variations due to concrete content, but provide the opportunity to measure them and take them into account

statistically. Whether replicated materials are matched or unmatched, and whether the experimental manipulations occur inside or outside the materials, replications should be regarded as a source of uncontrolled variation in the measure of the treatment differences, and should therefore be treated as a random effect.

NOTES

1. See especially James J. Bradac, "Threats to Generalization in the Use of Elicited, Purloined, and Contrived Messages in Human Communication Research," *Communication Quarterly* 34 (Winter 1986): 55-65; and Sally Jackson, *Message Effects Research: Principles of Design and Analysis* (NY: Guilford, 1992). See also Sally Jackson and Scott Jacobs, "Generalizing about Messages: Suggestions for the Design and Analysis of Experiments," *Human Communication Research* 9 (Winter 1983): 169-191; Daniel J. O'Keefe, *Persuasion: Theory and Research* (Newbury Park, CA: Sage, 1990), 121-129.

2. Jackson and Jacobs, "Generalizing about Messages"; Sally Jackson, Daniel J. O'Keefe, and Scott Jacobs, "The Search for Reliable Generalizations about Messages: A Comparison of Research Strategies," *Human Communication Research* 15 (Fall 1988): 127-141; Sally Jackson, Daniel J. O'Keefe, Scott Jacobs, and Dale Brashers, "Messages as Replications: Toward a Message-Centered Design Strategy," *Communication Monographs* 56 (December 1989): 364-384; Sally Jackson, Dale E. Brashers, and Joseph E. Massey, "Statistical Testing in Treatment by Replication Designs: Three Options Reconsidered," *Communication Quarterly* 40 (Summer 1992): 211-227. See also the large literature on the "language-as-fixed-effect fallacy" in psycholinguistics stimulated by Edmund B. Coleman, "Generalizing to a Language Population," *Psychological Reports* 14 (February 1964): 219-226 and Herbert H. Clark, "The Language-as-Fixed-Effect Fallacy: A Critique of Language Statistics in Psycholinguistics," *Journal of Verbal Learning and Verbal Behavior* 12 (August 1973): 335-359. Although it is often assumed that random sampling is logically required for treatment of a factor as random, much contemporary thinking suggests instead that a factor should be treated as random whenever its levels are chosen arbitrarily from among other equally acceptable levels. For detailed discussion of this issue, see Sally Jackson and Dale E. Brashers, *Random Factors in ANOVA*, Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-098 (Thousand Oaks, CA: Sage, 1994), 1-7, or Sally Jackson and Dale E. Brashers, "M > 1: Analysis of Treatment Replication Designs," *Human Communication Research* 20 (March 1994): 356-389.

3. Donald Dean Morley, "Meta-Analytic Techniques: When Generalizing to Message Populations is Not Possible," *Human Communication Research* 15 (Fall 1988): 112-126.

4. Our use of "unmatched" and "matched" parallels the use of "nested" and "crossed" by Jackson, O'Keefe, Jacobs, and Brashers. We might have used the nested/crossed distinction, but these terms seem appropriate only for replicated designs, and the present discussion encompasses unreplicated designs as well. Jackson (*Message Effects Research*, 27-28) draws a related distinction between "categorical comparisons" and "treatment comparisons." Categorical comparisons are between categories of messages with distinct membership, so experimental messages representing each category are "unmatched" and (in replicated designs) nested under categories. Treat-

ment comparisons are between versions of messages whose basic content has been "treated," so experimental messages are matched from one treatment level to another and (in replicated designs) crossed with treatments. For a more extensive discussion of these design possibilities, including their comparative strengths and weaknesses, see Jackson, *Message Effects Research*.

5. For detailed discussion of many variations of these types, see Jackson, *Message Effects Research*. Note that substantively "unmatched" messages may become "matched" under certain circumstances. For example, if unmatched replications are sorted into sets and each set administered to an independent block of subjects (so that subjects within a block respond to all of the messages within a set), the messages within a set become for all practical purposes matched, since responses to the messages within sets will be nonindependent.

6. For recent examples, see Evelyne J. Dyck and Gary Coldevin, "Using Positive vs. Negative Photographs for Third-World Fund Raising," *Journalism Quarterly* 69 (Autumn 1992): 572-579; Rhonda Gibson and Dolf Zillman, "The Impact of Quotation in News Reports on Issue Perception," *Journalism Quarterly* 70 (Winter 1993): 793-800; Laurence B. Lain and Philip J. Harwood, "Mug Shots and Reader Attitudes Toward People in the News," *Journalism Quarterly* 69 (Summer 1992): 293-300; and Douglas B. Ward, "The Effectiveness of Sidebar Graphics," *Journalism Quarterly* 69 (Summer 1992): 318-328.

7. For recent examples, see Albert C. Gunther and Esther Thorson, "Perceived Persuasive Effects of Product Commercials and Public Service Announcements," *Communication Research* 19 (October 1992): 574-596; and Richard F. Yalch, "Memory in a Jingle Jungle: Music as a Mnemonic Device in Communicating Advertising Slogans," *Journal of Applied Psychology* 76 (April 1991): 268-275.

8. For recent examples, see Duangkamol Chartprasert, "How Bureaucratic Writing Style Affects Source Credibility," *Journalism Quarterly* 70 (Spring 1993): 150-159; Prabu David, "Accuracy of Visual Perception of Quantitative Graphics: An Exploratory Study," *Journalism Quarterly* 69 (Summer 1992): 273-292; Luis Buceta Facorro and Melvin L. DeFleur, "A Cross-Cultural Experiment on How Well Audiences Remember News Stories From Newspaper, Computer, Television, and Radio Sources," *Journalism Quarterly* 70 (Autumn 1993): 585-601; James D. Kelly, "The Effects of Display Format and Data Density on Time Spent Reading Statistics in Text, Tables, and Graphs," *Journalism Quarterly* 70 (Spring 1993): 140-149; Michael A. Shapiro and Robert H. Rieger, "Comparing Positive and Negative Political Advertising on Radio," *Journalism Quarterly* 69 (Spring 1992): 135-145; Michael D. Slater and Donna Rouner, "Confidence in Beliefs about Social Groups as an Outcome of Message Exposure and its Role in Belief Change Persistence," *Communication Research* 19 (October 1992): 597-617; and Esther Thorson and Annie Lang, "The Effects of Television Videographics and Lecture Familiarity on Adult Cardiac Orienting Responses and Memory," *Communication Research* 19 (June 1992): 346-369.

9. For a more thorough discussion, see Jackson, *Message Effects Research*.

10. See Jackson, *Message Effects Research*, 31-36; see also Edwin J. Kay and Martin L. Richter, "The Category-Confound: A Design Error," *Journal of Social Psychology* 103 (October 1977): 57-63.

11. See Jackson, *Message Effects Research*, 41-43.

12. See Jackson, *Message Effects Research*, 43-45.

13. See Jackson, *Message Effects Research*, 45-47.

14. For discussion and further references, see O'Keefe, *Persuasion*, 141-145.

15. According to a recent review, about a third of experiments on message effects now use replicated designs; see Michael D. Slater, "Use of Message Stimuli in Mass Communication Experiments: A Methodological Assessment and Discussion," *Journalism Quarterly* 68 (Autumn 1991): 412-421. The examples cited in notes 7 and 8 illustrate a wide range of different implementations of replicated designs and a wide range of analytic choices.

16. In practice, researchers often drop the replication factor from analysis altogether, handling all observations within a given treatment level as independent even if the observations occur as subgroups defined by common experimental materials. Slater and Rouner, "Confidence in Beliefs about Social Groups," 614-615 n. 5, describe this as "common practice" in message effects research. The consequences of this practice are discussed in Jackson, Brashers, and Massey, "Statistical Testing in Treatment x Replication Designs." Even when "protected" by preliminary tests for replication-related effects, this analytic strategy violates the nonindependence assumption of the *F* test, as shown by Sally Jackson and Dale E. Brashers, "Assuming Independence When Dependence is Not Evident: A Fallacy of Misplaced Presumption" (Paper delivered at the International Communication Association Annual Meeting, Washington, DC, 1993).

17. See, e.g., John E. Hunter, Mark L. Hamilton, and Mike Allen, "The Design and Analysis of Language Experiments," *Communication Monographs* 56 (December 1989): 341-363.

18. For side-by-side comparison of interaction tests with replications fixed and random, see Geoffrey Keppel, *Design and Analysis: A Researcher's Handbook*, 2d ed. (Englewood Cliffs, NJ: Prentice-Hall, 1983), 527. The interpretation of interactions involving random factors is discussed thoroughly in Jackson and Brashers, *Random Factors in ANOVA*. Procedures for identifying moderators in treatment x replication designs are discussed in detail by Sally Jackson, "Meta-Analysis for Primary and Secondary Data Analysis: The Super-Experiment Metaphor," *Communication Monographs* 58 (December 1991): 449-462.

19. Michael Burgoon, John Hall, and Michael Pfau, "A Test of the 'Messages-as-Fixed-Effect Fallacy' Argument: Empirical and Theoretical Implications of Design Choices," *Communication Quarterly* 39 (Winter 1991): 18-34; Hunter, Hamilton, and Allen, "The Design and Analysis of Language Experiments"; and Slater, "Use of Message Stimuli in Mass Communication Experiments."

20. Jackson and Brashers, "M > 1." For detailed differentiation among possible null hypotheses tested in matched replicated designs, see esp. 368-369. For information on the relationship between power and number of replications, see esp. 374-375. More general methods for evaluation of power in replicated designs may be found in Jackson and Brashers, *Random Factors in ANOVA*, esp. 38-41.

21. This position is argued, for example, by Hunter, Hamilton, and Allen, "The Design and Analysis of Language Experiments." See also Slater, "Use of Message Stimuli in Mass Communication Experiments." Slater's discussion invokes a distinction between "between-message" and "within-message" designs, but this distinction is not as transparent as one might like, so it is not entirely clear whether he advocates treating replications as fixed when they are matched or when the treatment occurs outside the message. The latter circumstance is discussed in detail in the next section.

22. For more complete discussion of the consequences of treating matched replications as fixed, see Jackson and Brashers, "M > 1." For parallel

discussions concerning the case of replicated unmatched designs, see K. I. Forster and R. G. Dickinson, "More on the Language-as-Fixed-Effect Fallacy: Monte Carlo Estimates of Error Rates for F1, F2, F', and min F'," *Journal of Verbal Learning and Verbal Behavior* 15 (April 1976): 135-142; John L. Santa, John J. Miller, and Marilyn L. Shaw, "Using Quasi F to Prevent Alpha Inflation Due to Stimulus Variation," *Psychological Bulletin* 86 (January 1979): 37-46; and Thomas D. Wickens and Geoffrey Keppel, "On the Choice of Design and of Test Statistic in the Analysis of Experiments with Sampled Materials," *Journal of Verbal Learning and Verbal Behavior* 22 (June 1983): 296-309.

23. Slater, "Use of Message Stimuli," 419.

24. Slater, "Use of Message Stimuli," 413. Actual examples are common: In Lain and Harwood, "Mug Shots and Reader Attitudes," a fixed text was presented with several different photos of the author; in Ward, "The Effectiveness of Sidebar Graphics," the fixed text was presented with or without graphics; in Dyck and Coldevin, "Using Positive vs. Negative Photographs," the fixed text was accompanied by varied photos of impoverished children.

25. Slater, "Use of Message Stimuli," 417.

26. Slater, "Use of Message Stimuli," 418.

27. For further discussion, see Jackson, *Message Effects Research*, 92-96, or Jackson and Brashers, "M > 1," 383 n. 4.

28. Miles L. Patterson, Mary E. Churchill, Gary K. Burger, and Jack L. Powell, "Verbal and Nonverbal Modality Effects on Impressions of Political Candidates: Analysis from the 1984 Presidential Debates," *Communication Monographs* 59 (September 1992): 231-242.

29. For additional discussion of message-to-message variation in treatment effects, including examples of investigations reporting such variation, see Jackson, O'Keefe, Jacobs, and Brashers.

30. For amplification of this view, see Sally Jackson, "How to do Things to Words: The Experimental Manipulation of Message Variables," *Southern Communication Journal* 58 (Winter 1993): 103-114.