Defining Variables in Relationship to Other Variables: When Interactions Suddenly Turn Out to Be Main Effects

Markus Brauer

CNRS and Université Blaise Pascal, Clermont–Ferrand, France

and

Charles M. Judd

University of Colorado, Boulder

It is frequently the case that independent variables in experimental designs in social psychology are defined in relationship to levels of other independent variables. For instance, an experimental design in social cognition research might examine the effects of two different expectations on memory for information that is either consistent with the first expectation (and inconsistent with the second) or consistent with the second expectation. When the information factor is defined in this way, its interaction with expectation is perfectly confounded with the main effect of the stimulus factor defined in an absolute rather than a relative manner. Because we are trained to interpret interactions as qualifications of main effects, this alternative, and frequently more parsimonious, interpretation may be ignored. We illustrate the general issue and then review literature where interpretational ambiguities have resulted.   © 2000 Academic Press

In many areas of social psychological research, independent variables are defined in relationship to the levels of other independent variables rather than in some absolute way. For example, in a study on person memory, participants may receive an initial description of a hypothetical target person which is either

positively valenced or negatively valenced. They then read a number of positively valenced and negatively valenced behaviors supposedly enacted by this target person. They are then asked to recall the behaviors they have read. There are two ways to define the two independent variables in this study. One way defines both of them absolutely: The initial description is either positive or negative and the behaviors are either positive or negative. The other way is to define the second variable in relationship to the levels of the first: The initial description is either positive or negative and the behaviors are either consistent or inconsistent in their valence with that initial description.

It is well known (Abelson, 1995; Shaffer, 1977) that the analysis following from the first definition is identical to the analysis following from the second, except that one of the main effects and the interaction trade places. Although the two analyses yield identical results, it is our belief that the relational definition often results in theoretical confusion because researchers may fail to realize that the interaction under the relational definition is identical to the main effect under the absolute definition. In this paper, we illustrate the issue and then we consider a variety of research questions in social psychology where we believe interpretational ambiguities have resulted.

ILLUSTRATION OF THE ISSUE

Consider a recent study by Koomen and Dijker (1997), who examined potential moderators of the "stereotype-consistency bias." According to this bias, we preferentially process information that is consistent rather than inconsistent with the stereotype about a target group (for a review see Stangor & McMillan, 1992). Koomen and Dijker argued that the stereotype-consistency bias should depend on whether the information refers to an in-group target or an out-group target. Specifically, they predicted that individuals would mostly recall stereotype-consistent information when dealing with the out-group but that this effect should disappear, or even be reversed, when the information referred to a member of the in-group. To examine this prediction, they presented Dutch participants with behaviors that were either stereotypic or counterstereotypic of either a Dutch target person (in-group) or a Turkish target person (out-group). Behaviors that were stereotypic for one target group were counterstereotypic for the other. Four days later, participants were unexpectedly visited again and were given a recognition memory test. The results of the study are presented in Fig. 1.

An analysis of variance was conducted on the data, analyzing the recognition accuracy as a function of two crossed factors: Target (Dutch target vs Turkish target) × Item Consistency (stereotype-consistent vs stereotype-inconsistent). The data analyses showed no main effect for Item Consistency, which means that in this study, there was no evidence for a generalized stereotype-consistency bias. However, the interaction between the two factors was reliable, $F(1, 167) = 59.60$, $p < .001$. These data seem to point to an asymmetric consistency bias: Participants display the classic stereotype-consistency bias when they deal with
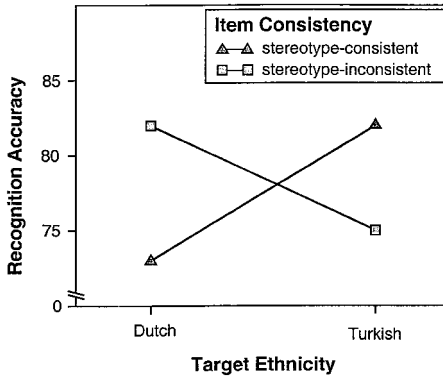
FIG. 1. Results of Koomen and Dijker's (1997) study on the asymmetric consistency bias. Item Consistency is defined in relationship to the target that is being judged. The dependent variable is recognition accuracy for behaviors that had been presented previously. All participants were native Dutch students.

an out-group target but this effect is reversed when the behavioral information refers to an in-group target.

What is problematic about this interpretation? Our argument is that the way in which Koomen and Dijker (1997) have defined their independent variables may cause us to overlook that the interaction, which is theoretically meaningful, may be more parsimoniously interpreted as a relatively uninteresting main effect. The problem originates from the fact that the second independent variable, Item Consistency, has been defined relative to the levels of the first independent variable rather than in an absolute way. For a Dutch target, the stereotypically Dutch behaviors are obviously stereotype-consistent behaviors and the stereotypically Turkish behaviors are stereotype-inconsistent behaviors. For a Turkish target, this is obviously reversed. Thus, the definition of the Item Consistency factor depends on the level of Target, the other independent variable in the design.

An alternative approach to the analysis is to define the two factors in an absolute manner: Target (Dutch vs Turkish) and Item Type (stereotypically Dutch vs stereotypically Turkish). In Table 1, we compare these two ways of defining the independent variables in terms of the various main effect and interaction contrasts that are tested in the alternative analyses of variance. The main effect of the first factor, Target, remains constant across the two definitions. However, the other main effect and the interaction switch places depending on whether the second factor is defined relatively (Item Consistency) or defined absolutely (Item Type). When the second factor is defined relatively, its main effect is equivalent to the interaction in the analysis where the second factor is defined as Item Type. The interaction, when the second factor is defined relatively, is equivalent to the main effect of Item Type under the absolute definition of the second factor.

TABLE 1
Contrast Codes Used in Koomen & Dijker's (1997) Study (Top Panel) and Alternative
Contrast Codes where Variables Are Defined Absolutely (Bottom Panel)

Design 1: defining variables in relationship to other variables

Target	Behavior	Target (Ta) [Dutch: −1, Turkish: 1]	Item Consistency (IC) [cons.: 1, incons.: −1]	Interaction (Ta × IC)
Dutch	Dutch	−1	1	−1
Dutch	Turkish	−1	−1	1
Turkish	Dutch	1	−1	−1
Turkish	Turkish	1	1	1
		ns	*ns*	*p* < .001

Design 2: defining variables absolutely

Target	Behavior	Target (Ta) [Dutch: −1, Turkish: 1]	Item Type (IT) [Dutch: −1, Turkish: 1]	Interaction (Ta × IT)
Dutch	Dutch	−1	−1	1
Dutch	Turkish	−1	1	−1
Turkish	Dutch	1	−1	−1
Turkish	Turkish	1	1	1
		ns	*p* < .001	*ns*

Figure 2 presents the same recognition accuracy scores as those in Fig. 1, but this time the second factor is defined absolutely as Item Type. In this figure, preferential memory for stereotype-consistent items would take the form of an interaction of the two factors. Obviously, the data do not show this effect.
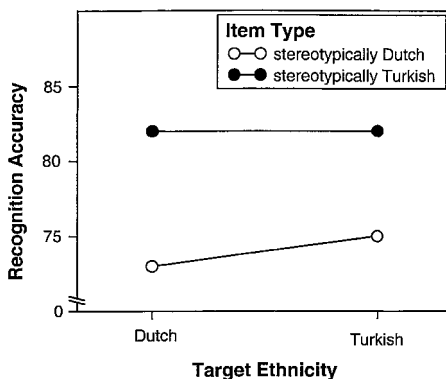

FIG. 2. Alternative representation of Koomen and Dijker's (1997) results. Item Type is defined absolutely; i.e., its levels refer to the target group for which it is stereotypical.

Instead, they show a main effect for Item Type: All participants recognized stereo-typically Turkish behaviors better than stereotypically Dutch behaviors. What looks like an asymmetric consistency bias in Fig. 1 may be nothing more than the finding that the stereotypical Turkish behaviors were more memorable than the stereotypical Dutch behaviors. It should be noted that Koomen and Dijker (1997) are fully aware of this confound. They explicitly point out the alternative interpretation and attempt to argue against it in the discussion of their article (see p. 598). However, not all authors alert the reader to the potential interpretational ambiguities, and some researchers may not even be aware that a confound exists in their data.

The point of this illustration is to suggest that relational definitions of inde-pendent variables can cause interpretational ambiguities when one fails to realize that an interaction can parsimoniously be interpreted as a main effect of an absolutely defined independent variable. In the following section, we point to particular literatures in social psychology where we think this has been problem. Before doing so, however, two points need further clarification.

First, our argument concerns interpretational ambiguities, not analytic alter-natives. Thus, the analyses are correct regardless of how one defines the inde-pendent variables. It is only their meaning that we are debating.

Second, there may well be cases when an interaction between a relatively defined independent variable and another independent variable is best interpreted theoretically as a qualification of the main effect of that relatively defined variable. Such an interpretation may at times be more appropriate than its alternative interpretation as the main effect of an absolutely defined independent variable. We suggest, however, that our training has naturally led most of us to interpret interactions as qualifying main effects and therefore we will tend to forget about the alternative main effect interpretation of the interaction between a relatively defined independent variable and another variable. We will tend to overlook this alternative interpretation, even when it is a more parsimonious one.

HOW WIDESPREAD IS THE PROBLEM?

In social cognition research, prior expectations have been shown to have a variety of effects on memory and judgment. For instance, an expectation may affect what information is attended to in a stimulus array, what information is retrieved at the time of judgment of the stimulus, and the content of the judgment itself. A typical research paradigm in this literature involves two independent variables: Typically expectation varies between participants and then the nature of the stimulus information is manipulated within participants. Attention is then focused on whether information that is consistent with the expectation is treated differently from information that is inconsistent with the expectation.

Consider a study by Borgida and DeBono (1989) examining whether expertise in a domain reduces the tendency to recall information that is consistent with an expectation. There were two samples of participants in this study, an expert sample consisting of professional librarians and a novice one consisting of psychology undergraduates. They read a paragraph describing the behaviors of a

stimulus person. The paragraph contained approximately equal numbers of introverted and of extroverted behaviors. One week later, participants returned and were asked to recall the behaviors of the person. At the time of recall, participants were given one of two expectations: Either they were told that the person was applying for a job as a research librarian or they were told that she was applying for a job as a real estate agent. Thus, expectation was manipulated between participants and type of information varied within participants, with half of the information being consistent with the expectation (introverted behaviors with the librarian expectation and extroverted ones with the real estate agent expectation) and half inconsistent. The researchers hypothesized that expectations should lead to a recall advantage for consistent information, but that this difference should be reduced when participants were relatively expert in a given domain (i.e., when the librarian participants were given the librarian expectation). Consistent with this hypothesis, the data from the librarian participants showed a consistency of information by expectation interaction such that recall was better for consistent than for inconsistent information when the real estate agent expectation was given and there was no difference (actually a slight reversal) when the librarian expectation was given.

Consistency, of course, is defined relatively in this analysis, since extroverted behaviors are consistent with one expectation and introverted ones are consistent with the other. As a result, the consistency by expectation interaction can equivalently, and perhaps more parsimoniously, be interpreted as a main effect for type of behavior: Extroverted behaviors were simply recalled with greater frequency by these librarian participants than were introverted behaviors regardless of the expectation given. The authors would undoubtedly question this alternative explanation based on the parallel data they collected from the sample of undergraduate psychology participants who showed only a main effect for information consistency. In a later section, we discuss how and when such additional participants may help solve the problem.

Social stereotypes can serve as expectations that influence the recall of consistent and inconsistent information, and the magnitude of that recall difference has been hypothesized to depend on whether the stereotype concerns an in-group or an out-group. The above-mentioned study by Koomen and Dijker (1997) is a typical example. Other sorts of stereotyping effects are also susceptible to interpretational problems deriving from defining variables relatively. Consider the well-known result that out-groups are typically seen as being more homogeneous than in-groups (Judd & Park, 1988; Park & Rothbart, 1982). Typically this effect is demonstrated with participants from two different groups who judge the variability of both their own group and the other group. Here, the judged group can be defined either relatively (in-group vs out-group) or absolutely (Target Group 1 vs Target Group 2). Interpretational problems arise when one predicts an asymmetry in out-group homogeneity, such that it is weaker or stronger for one participant group than for another. For instance, Simon and Brown (1987) and Simon and Pettigrew (1990) have argued that minority

participant groups ought to show a relatively smaller out-group homogeneity effect than majority groups. Likewise, Lorenzi-Cioldi (1998) suggested that social status of the groups should moderate the out-group homogeneity effect. And Lorenzi-Cioldi, Eagly, and Stewart (1995) have argued that this status difference means that female participants should show less out-group homogeneity than male participants. When the judged group is defined relatively (i.e., in-group vs out-group), these asymmetry hypotheses predict an interaction between participant group (males vs females, high-status vs low-status, majority versus minority) and whether the judged group is an in-group or an out-group. This interaction, of course, is equivalent to the main effect of the target group, when defined absolutely.

The in-group–out-group domain is a research area where relatively defined independent variables are routinely used and their use may well have led to incorrect theoretical conclusions. There have been a large number of studies that have examined differences between participant groups (based on gender, social status, power, size, etc.) in the magnitude of intergroup biases such as ethnocentrism and out-group homogeneity. Across studies, the patterns of effects that are obtained are confusing and contradictory. In most cases, the researchers conclude that the high-status, high-power groups are more biased than low-status, low-power groups (Brewer, 1979; Devos, Comby, & Deschamps, 1996; Fiske, 1993; Jost & Banaji, 1994; Lee & Ottati, 1993; Lorenzi-Cioldi, 1998) but several studies also point toward the opposite effect (Branthwaite & Jones, 1975; Brown & Smith, 1989; Grant & Brown, 1995; see Brauer, in press, for a more detailed discussion of the contradictions). In nearly all cases, conclusions in individual studies rely on interactions between a two-level participant group factor and the in-group/out-group (relatively defined) factor. It seems likely that the resulting confusion about the nature of these interactions arises from the specific target group effects with which these interactions are confounded.

The sorts of theoretical ambiguities that arise from relatively defined independent variables may be particularly difficult to uncover in more complicated designs involving multiple factors. Consider further work on out-group homogeneity conducted in our own laboratory. (We do not want to be seen as criticizing only the work of others!) One form of out-group homogeneity that we have identified (Park & Judd, 1990) involves the tendency to see a greater prevalence of individuals who confirm the stereotype and a smaller prevalence of counterstereotypic individuals in the out-group than the in-group. So, we ask participants from both groups to indicate the percentage of both their own group and the other group who possess attributes that are either stereotypic of the judged group or counterstereotypic of that group. Even though we define target group in these studies absolutely, we have typically defined attribute stereotypicality relatively. So particular attributes are defined as either stereotypic or counterstereotypic depending on the target group that is being rated (e.g., aggressive is stereotypic for the male target group but counterstereotypic for the female target group). The out-group homogeneity prediction in this research is

TABLE 2

Contrast Codes for Judd, Park, Ryan, Brauer, and Kraus's (1995) Study on Intergroup Perceptions between Ethnic Groups in the United States

Part. group	Targ. group	Attribute	[1] Part. group	[2] Targ. group	[3] Attr. stereo.	[4] Attr. type	[5] $[1] \times [2] \times [3]$	[6] $[1] \times [4]$
AA	AA	Poor	-1	-1	1	-1	1	1
AA	AA	Stuffy	-1	-1	-1	1	-1	-1
AA	WA	Poor	-1	1	-1	-1	1	1
AA	WA	Stuffy	-1	1	1	1	-1	-1
WA	AA	Poor	1	-1	1	-1	-1	-1
WA	AA	Stuffy	1	-1	-1	1	1	1
WA	WA	Poor	1	1	-1	-1	-1	-1
WA	WA	Stuffy	1	1	1	1	1	1

Note. AA = African American; WA = White American.

[1] Participant group: African Americans = -1; White Americans = $+1$.

[2] Target group: African Americans = -1; White Americans = $+1$.

[3] Attribute stereotypicality: counterstereotypic = -1; stereotypic = 1 for the target group to be judged.

[4] Attribute type: stereotypic for African Americans = -1, stereotypic for White Americans = 1.

the triple interaction involving participant group, target group, and attribute stereotypicality, such that the difference in the percentage estimates between stereotypic attributes and counterstereotypic attributes is greater for out-groups than for in-groups (Judd, Park, Ryan, Brauer, & Kraus, 1995; Park & Judd, 1990; Park, Ryan, & Judd, 1992).

The attribute factor in this research could have been defined absolutely rather than relatively. So, rather than defining attributes as stereotypic or counterstereotypic of the target group being judged, we could have defined the factor as attributes stereotypic of one group or of the other (e.g., aggressive is a stereotypically male attribute, nurturant is stereotypically female). Had we defined the attribute factor this way, then it would have been clear that the triple interaction that we interpreted as evidence of out-group homogeneity is equivalent to a participant group by attribute type two-way interaction. That is, we can also interpret it as indicating that attributes that are stereotypic of the out-group are judged to be more prevalent than are attributes that are stereotypic of the in-group, regardless of the specific target group being rated.

To show this equivalence, in Table 2 we present contrast codes for the relevant effects using attributes taken from Judd et al. (1995, Study 4), where out-group homogeneity was found with African American and White American participant and target groups. In this study, some attributes were stereotypic of White Americans (e.g., stuffy) and others were stereotypic of African Americans (e.g., poor). The contrast codes in column 1 of the table define the participant group.

Those in column 2 define the target group. Column 3 presents codes that define attributes in relationship to the target group. That is, stuffy is stereotypic of the White American target group and counterstereotypic of the African American target group; poor is stereotypic of the African American target group and counterstereotypic of the White American target group. In column 4, the alternative codes are given for the attributes, defining them simply by whether they are stereotypic of one group or the other rather than defining them relative to a particular target group. Column 5 presents the product of the codes in the first three columns, thus defining the participant group by target group by attribute stereotypicality triple interaction, which we have interpreted as out-group homogeneity. The exact same set of codes is contained in column 6, but this time these codes derive from multiplying the codes of columns 1 and 4, thus revealing that the out-group homogeneity triple interaction is equivalently a participant group by attribute type two-way interaction.

In other words, in this design, we now have two interpretations for the critical effect: (1) out-group homogeneity or (2) the tendency to see out-group attributes as more prevalent in both target groups than in-group attributes. The second interpretation would imply that White American participants see the trait "poor" as more prevalent in the population at large than African American participants do whereas African American participants see the trait "stuffy" as more prevalent in the population at large. In this case, we believe there is good evidence to suggest that the former interpretation is probably the more correct, since there exists considerable evidence that one actually overestimates the prevalence of one's own attributes and underestimates the prevalence of attributes one does not possess (i.e., the false consensus effect; Ross, Greene, & House, 1977). As Ryan and Judd (1992) have pointed out, out-group homogeneity and false consensus are negatively confounded with each other.

The above examples illustrate that the problem of defining variables in relationship to other variables occurs in a variety of research domains in social psychology. In general, our experience suggests that problems are particularly likely to appear either when one is exploring the role of prior expectations on subsequent judgments or behavior or when one is exploring in-group versus out-group or self versus other judgments.

In the first case, the expectations under consideration may be derived from prior experiences and interactions with particular targets, hypotheses that are to be tested by the participant, priming stimuli that participants are exposed to, or stereotypic expectations about group behaviors. The research typically focuses on the ways in which judgments and behavior depend on whether subsequent stimuli are consistent or inconsistent with those expectations. So researchers want to know whether memory for consistent or inconsistent information is better retained. They want to know whether behaviors that are consistent or inconsistent with a particular person impression are attended to more. They want to know whether interpersonal expectancy effects lead to an increased probability of expectancy-consistent behaviors compared to inconsistent ones. They want to

know whether primes decrease the judgmental latency for semantically consistent compared to inconsistent probes. They want to know whether information that disconfirms a group stereotype is less impactful than confirming information. In general the research approach involves comparing two or more expectations, hypotheses, priming stimuli, or stereotypes. And the second independent variable, typically defined in relation to the different expectations, concerns the consistency or inconsistency of the target stimulus to the expectations. In such cases, whenever one hypothesizes an asymmetric effect of the expectations, such that the consistent versus inconsistent difference in responses is greater for one sort of expectation than for the other, the resulting expectation by consistency interaction that is hypothesized can often be more plausibly interpreted as a relatively uninteresting main effect of the type of target stimulus, defining this second factor absolutely rather than relatively.

In the second case, the case of in-group–out-group designs or designs involving the self and another, it is this factor that is typically defined relationally rather than absolutely. So members of two participant groups make judgments about their own group or the other group. Two individuals who interact make judgments about themselves or their interacting partner. These are both relationally defined variables since the specific target group or person that is judged depends on which participant or participant group is doing the judging. Again, when one hypothesizes an asymmetric effect, such that the in-group versus out-group difference or the self versus other difference is greater for one sort of participant group than for another, the resulting interaction can often be more parsimoniously interpreted as a main effect of the specific target person or group being judged, defining this latter factor absolutely rather than relatively.

HOW CAN THE PROBLEM BE ADDRESSED?

It might seem that the logical implication of the argument presented thus far is that variables should be defined absolutely rather than relatively, in which case the alternative main effect explanation would be less likely to be ignored. However, although the confound is certainly more apparent when variables are defined absolutely, it is not in fact our recommendation to avoid relational definition of variables altogether. In all of the designs that we have considered there are two effects that are confounded: a main effect of an absolutely defined variable and an interaction between that same variable, if defined relatively, and a second independent variable that provides its relative definition. This confound exists regardless of whether the variable is defined absolutely or relatively. While the main effect interpretation may be more visible when the variable is defined absolutely, doing so does not get rid of the confound.

In general, there are three strategies that can be used to address the problem. First, one can appeal to theoretical arguments and prior research to argue against the plausibility of the confounded main effect of the absolutely defined independent variable. Obviously this strategy depends on the existence of compelling reasons for doubting the main effect. It is this strategy that we adopted in

suggesting that the confound in the Judd et al. (1995) research was implausible since it ran in the opposite direction from the well-known false consensus effect.

The second strategy that can be pursued in order to unconfound the predicted effect from the alternative main effect is to modify the research design, generally by collecting ratings from more than one participant group or incorporating appropriate control groups. Consider Borgida and DeBono's (1989) prediction that expectancy-consistent information should be recalled better than inconsistent information, but that this difference was asymmetric in that it should be reduced if one is an expert in the domain under consideration. So, librarian participants showed a stronger expectancy effect in the case of the extroverted real estate agent (where extroverted behaviors were consistent with the expectancy) than in the case of the librarian expectation (where introverted behaviors were consistent). As already discussed, this interaction is equivalently interpreted as a main effect of stimulus behaviors: The librarians simply recalled more extroverted behaviors than introverted ones. To their credit, Borgida and DeBono (1989) added a control group to their design by presenting college student participants with the same stimulus material. The control group displayed the classic consistency bias such that behaviors that were consistent with the initial expectation were memorized better than behaviors that were inconsistent. Furthermore, this bias did not depend on whether the initial expectation was a librarian or a real estate agent. These results help rule out the plausibility of the stimulus effect interpretation: It is evident that the extroverted behaviors were as easy to recall as the introverted behaviors, at least for college students. A similar reasoning can be applied to the study by Koomen and Dijker (1997). Had they collected data not only from Dutch participants but also from Turkish participants, they could argue convincingly that their results are due to an asymmetric consistency bias and not to particular characteristics of the stimulus material.[1]

Very often, the second strategy cannot be employed because the researchers have already included members from both participant groups, and ratings from an

---

[1] Of course, the adding of the control group in the Borgida and DeBono (1989) study does not eliminate another alternative interpretation of the observed interaction: There could be some sort of "self-reference effect" according to which it is easier to recall behaviors that one is unlikely to engage in, so the (presumably introverted) librarians recalled more of the extroverted behaviors because they hardly ever perform these behaviors. It should be noted, however, that this confound is not due to the way the independent variables were defined (relationally or absolute) but rather to the fact that participants had not been assigned randomly to levels of expertise. A similar argument can be made for Koomen and Dijker's (1997) study. Even if ratings from two participant groups had been collected one still could not exclude the already-mentioned self-reference effect: If the Participant Group $\times$ Target $\times$ Item Consistency interaction were significant we would not know if individuals display an asymmetric stereotype-consistency bias or if they simply have better memory for behaviors that are counterstereotypic of the in-group, i.e., behaviors that they are unlikely to engage in. Randomly assigning individuals to the levels of participant group (e.g., using artificial groups in the laboratory) would not address the problem. Rather, the solution would be to construct the stimulus material so that behaviors that are stereotypical and counterstereotypical of one group are nonstereotypical (neither stereotypical nor counterstereotypical) of the other group.

appropriate control group are meaningless. The numerous studies that tested an asymmetric homogeneity bias or an asymmetric ethnocentrism bias are good examples (Fiske, 1993; Jost & Banaji, 1994; Lorenzi-Cioldi, 1998; Lorenzi-Cioldi et al., 1995). For such a case, we would suggest a third strategy, which consists of using multiple participant and target groups in the design. Brauer (in press), for example, tested the effects of social status on out-group homogeneity and recruited participants from two high-status groups (e.g., doctors and lawyers) and from two low-status groups (e.g., waiters and hairdressers). Participants then judged three of the four target groups from which participants were recruited: their own group, the target group with the same social status as their in-group, and one of the two target groups with a different social status than their in-group. Brauer employed an analytic strategy that involved the examination of all possible pairwise comparisons among participant groups (see the original article for more details). By analyzing the effect size of the participant group by target group interactions, he demonstrated that members of low-status groups tend to display more out-group homogeneity than members of high-status groups. The same data were analyzed via a different analytic strategy by McClelland, Judd, and Brauer (1999), who compared homogeneity ratings between in-groups and out-groups holding target group status constant. Thus, for high-status participants, comparisons were made between the variability of their high-status in-group and the perceived variability of the out-group who was also high-status. Similar comparisons were made for low-status participants, but this time focusing only on the low-status target groups. Confirming the results of the original analyses, doctors and lawyers showed a smaller out-group homogeneity effect when judging doctors and lawyers (one in-group and one equal-status out-group) than waiters and barbers when judging waiters and barbers (another in-group and out-group set of equal status).

Note that the conclusion from Brauer's (in press) study is contrary to the dominant finding in the literature. Whereas most studies seem to point to an asymmetric "high-status" out-group homogeneity effect (high-status groups display more out-group homogeneity than low-status groups), Brauer's findings indicate that members of low-status groups are more biased. The present conclusion is that in previous studies, the asymmetric out-group homogeneity effect was confounded with a target group main effect. Therefore, the results of those studies are difficult to interpret. In the Brauer study, however, in which multiple participant and target groups were used, the two effects were unconfounded. It seems to us that this is a good example of how an entire research area has come to incorrect conclusions by failing to realize that the interaction between a relationally defined variable and another variable may be more parsimoniously explained by a main effect.

CONCLUSION

The methodological training of most social psychologists has traditionally included a heavy emphasis on experimental designs, where theoretical factors of

interest are manipulated and the main and interactive effects of those factors are evaluated using analysis of variance procedures. A by-product of this training is the belief that experimental factors are the explanatory constructs of primary interest and that interactions are most appropriately interpreted as qualification or modifications of the effects of those constructs. However, in many domains of research, the levels of one factor are defined in relationship to the levels of another. When this is the case, interactions involving this factor and the other one can be interpreted in multiple ways. The contention of this paper is that such interactions are frequently interpreted as qualifications and that researchers routinely fail to appreciate that these interactions can equivalently, and frequently more parsimoniously, be interpreted as main effects of relatively uninteresting stimulus variation manipulations.

In making our argument, we have illustrated the interpretational problems in a variety of research areas in social psychology and we have identified those domains where relationally defined independent variables are particularly prevalent. Additionally we have argued that at least in one domain, focusing on the moderation of intergroup biases as a function of participant characteristics, the problem has led to considerable interpretational difficulties.

In general, whenever one theoretically predicts an interaction between one factor and a relationally defined second factor, the confound between that interaction and the main effect of the second factor, defined absolutely, is unavoidable. Other than theoretical arguments about the plausibility of that main effect, one solution is to add a control group to the design or to use a full in-group–out-group design. Another solution is to collect data from multiple participant and target groups, so that the effects become unconfounded. Our hope is that this paper will alert researchers to the interpretational ambiguities that we have pointed to and encourage them to undertake appropriate design modifications to deal with the problem.

REFERENCES

Abelson, R. P. (1995). *Statistics as principled argument.* Hillsdale, NJ: Erlbaum.

Borgida, E., & DeBono, K. G. (1989). Social hypothesis testing and the role of expertise. *Personality and Social Psychology Bulletin,* **15,** 212–221.

Branthwaite, A., & Jones, J. E. (1975). Fairness and discrimination: English versus Welsh. *European Journal of Social Psychology,* **5,** 323–338.

Brauer, M. (in press). Intergroup perception in the social context: The effects of social status and group membership on perceived outgroup homogeneity and ethnocentrism. *Journal of Experimental Social Psychology.*

Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin,* **86,** 307–324.

Brown, R., & Smith, A. (1989). Perceptions of and by minority groups: The case of women in academia. *European Journal of Social Psychology,* **19,** 61–74.

Devos, T., Comby, L., & Deschamps, J. C. (1996). Asymmetries in judgments of ingroup and outgroup variability. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 7, pp. 95–144). Chichester, UK: Wiley.

Fiske, S. T. (1993). Controlling other people: The impact of power on stereotyping. *American Psychologist,* **48,** 621–628.

Grant, P. R., & Brown, R. (1995). From ethnocentrism to collective protest: Responses to relative deprivation and threats to social identity. *Social Psychology Quarterly,* **58,** 195–211.

Jost, J. T., & Banaji, M. R. (1994). The role of stereotyping in system-justification and the production of false consciousness. *British Journal of Social Psychology,* **33,** 1–27.

Judd, C. M., & Park, B. (1988). Out-group homogeneity: Judgments of variability at the individual and group levels. *Journal of Personality and Social Psychology,* **54,** 778–788.

Judd, C. M., Park, B., Ryan, C. S., Brauer, M., & Kraus, S. L. (1995). Stereotypes and ethnocentrism: Interethnic perceptions of African American and White American college samples. *Journal of Personality and Social Psychology,* **69,** 460–481.

Koomen, W., & Dijker, A. J. (1997). Ingroup and outgroup stereotypes and selective processing. *European Journal of Social Psychology,* **27,** 589–601.

Kroll, J. F., & Stewart, E. (1994). Category inference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language,* **33,** 149–174.

Lee, Y. T., & Ottati, V. (1993). Determinants of in-group and out-group perceptions of heterogeneity. *Journal of Cross-cultural Psychology,* **24,** 298–318.

Lorenzi-Cioldi, F. (1998). Group status and perceptions of homogeneity. In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 9). Chichester, UK: Wiley.

Lorenzi-Cioldi, F., Eagly, A. H., & Stewart, T. L. (1995). Homogeneity of gender groups in memory. *Journal of Experimental Social Psychology,* **31,** 193–217.

McClelland, G. H., Judd, C. M., & Brauer, M. (1999). *Unfettering contrasts from ANOVA's restrictions: Comment on Abelson and Prentice (1997).* Unpublished manuscript, University of Colorado (currently in revision).

Park, B., & Judd, C. M. (1990). Measures and models of perceived group variability. *Journal of Personality and Social Psychology,* **59,** 173–191.

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology,* **42,** 1051–1068.

Park, B., Ryan, C. S., & Judd, C. M. (1992). The role of meaningful subgroups in explaining differences in perceived variability for in-groups and out-groups. *Journal of Personality and Social Psychology,* **63,** 553–567.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology,* **13,** 279–301.

Ryan, C. S., & Judd, C. M. (1992). False consensus and outgroup homogeneity: A methodological note on their relationship. *British Journal of Social Psychology,* **31,** 269–283.

Shaffer, J. P. (1977). Reorganization of variables in analysis of variance and multidimensional contingency tables. *Psychological Bulletin,* **84,** 220–228.

Simon, B., & Brown, R. (1987). Perceived intragroup homogeneity in minority–majority contexts. *Journal of Personality and Social Psychology,* **53,** 703–711.

Simon, B., & Pettigrew, T. F. (1990). Social identity and perceived group homogeneity. *European Journal of Social Psychology,* **20,** 269–286.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and social-developmental literatures. *Psychological Bulletin,* **111,** 42–61.