Samuel Stouffer and Relative Deprivation

Thomas F. Pettigrew¹

¹University of California, Santa Cruz

Corresponding Author:

Thomas F. Pettigrew

Department of Psychology

University of California, Santa Cruz

Santa Cruz, CA 95064

e-mail: pettigr@ucsc.edu
Abstract

This paper first offers a tribute to Samuel Stouffer (1900–1960), a major contributor to social psychology. He helped to establish probability surveys as a useful method for social science, led three major studies at midcentury, and introduced important new concepts and statistical methods. Thus, both conceptually and methodologically, he shaped modern social psychology. Second, the paper revitalizes Stouffer’s most famous concept——relative deprivation. A new meta-analysis demonstrates that relative deprivation predicts a wide range of important outcomes, so long as it measures resentment with data from individuals and is paired with dependent variables of similar scope. Unfortunately, sociology largely abandoned the concept because it failed to meet the overstated early claims made for it in the collective protest domain. The history of this use and disuse of relative deprivation is summarized and critiqued.

Keywords

Samuel Stouffer, relative deprivation, meta-analysis, collective protests, ecological fallacy
It is a great honor to receive the 2014 Cooley-Mead Award. I feel especially honored when I review the names of those who have received this award in past years. Three are former Harvard teachers and colleagues of mine—Fred Bales, Alex Inkeles, and George Homans; many are old friends whose work I have long admired—Larry Bobo, Bernie Cohen, Gary Fine, Hal Kelley, Jane Piliavin, and Sheldon Stryker.

The first Cooley-Mead Award was given to Muzafer Sherif in 1979. Had the award’s inception been two decades earlier, I am confident that Samuel Stouffer (1900–1960), a mentor and role model for me, would have also been a recipient. He was elected president of the American Association of Public Opinion Research in 1952 and the American Sociological Association in 1953. And he is certainly one of the truly great and influential social psychologists in sociology’s history.

His generation followed that of Charles Horton Cooley (1864–1929), George Herbert Mead (1863–1931), W. I. Thomas (1863–1947), and other founders of American sociology in the early development of social psychology. Together with Paul Lazarsfeld and Rensis Likert, Stouffer developed the probability survey into a refined research instrument for all the social sciences.

He was also the vital leader of three of the major social science projects of the mid-twentieth-century: Gunnar Myrdal’s (1944) *An American Dilemma*, for which Stouffer kept the study going when Myrdal returned to Sweden when his nation was threatened by Germany’s *Wehrmacht*; *The American Soldier* series (Stouffer et al. 1949a; Stouffer et al. 1949b; Stouffer, Guttman, et al. 1950; Stouffer 1962, chapter 2), which Stouffer organized and directed on Army morale throughout World War II; and the survey study of McCarthyism published at the height of that dark period in American political
history. This third book, *Communism, Conformity and Civil Liberties* (1955), although six decades old, remains a model of how to analyze large-scale survey data and how to write about it simply and succinctly.

Stouffer’s political courage was involved in this last work that undercut Senator Joseph McCarthy’s claim that the nation was highly concerned about Communists in government. In 1952, The FBI’s J. Edgar Hoover received notice from a secret Harvard informant who claimed that, of all people, “Professor Talcott Parsons was probably the leader of an inner group” of Communist sympathizers at Harvard! The informant alleged that the new Department of Social Relations had turned into a dangerous left-wing center as a result of “Parsons’ manipulations and machinations.”

Based on this dubious “evidence,” Hoover authorized the Boston FBI to initiate a security investigation of Parsons. Because he knew such suspected “Communists” as Parsons, Stouffer was then denied access to classified documents. Infuriated, Parsons in turn immediately prepared an affidavit in defense of Stouffer. “This allegation is so preposterous,” Parsons wrote, “that I cannot understand how any reasonable person could come to the conclusion that I was a member of the Communist Party or ever had been.” To Stouffer, he wrote, “I will fight for you against this evil with everything there is in me: I am in it with you to the death.” (Wikipedia, 2014)

Stouffer soon responded. The Republican Methodist from small-town Iowa not only wrote *Communism, Conformity and Civil Liberties* soon after this episode, but he also successfully faced off a McCarthyist inquisition held in Boston. He arranged for his former doctoral students who were Roman Catholic nuns and priests, replete in their
habits and collars, to sit ostentatiously behind him at the hearing. That proved too much for the nominally Catholic McCarthy; he abruptly ended the session.

Stouffer’s many key contributions are not remembered today as fully as they deserve. So I devote this paper to recalling his life and achievements and to one of his most lasting social psychological conceptual contributions—relative deprivation.

The Irrepressible Stouffer

Stouffer was born and grew up in the small town of Sac City in western Iowa. In 1900, the year of his birth, Sac City boasted only 2,079 people, and today it has virtually the same population. His father was the publisher of the local newspaper, and he initially planned to join him in journalism. He majored in English at the tiny Methodist college, Morningside, just sixty miles to the west in Sioux City. Upon graduation, he set off for Harvard University where he earned his masters degree in English before returning to his father’s newspaper.

His path to sociology began with a chance meeting that he once described to me in detail. One summer in the 1920s, he took his family to the Wisconsin lake country for a vacation. Restless and sociable as always, Stouffer quickly made acquaintance with other vacationers staying at nearby cabins. One of these neighbors was Edward Ross, the famous University of Wisconsin sociologist who at the time was one of the most cited writers in the discipline (Collier, Minton, and Reynolds 1991).

Stouffer’s curious mind led him to pepper Ross with questions about sociology. Ross responded by giving him a copy of his then-popular social psychology textbook (Ross 1908). Stouffer read it voraciously, became fascinated with social science, and wanted to know how one became a sociologist. Ross encouraged him to apply to a PhD
program in the subject. With characteristic decisiveness, Stouffer soon applied to the
University of Chicago’s sociology department. He was accepted, but to pay for his
training he had to work long hours at the Chicago Tribune as a reporter. Years later, he
would attribute his writing style more to the Tribune than to the formal style of academia.
“I’ll never be thought of as a great sociologist,” he liked to joke, “I’m not a theorist and
everyone can understand every damn word I write!”

Stouffer studied at Chicago with two leading American pioneers in quantitative
techniques for social science——William Ogburn in sociology and Lewis Thurstone in
psychology. After receiving his PhD in 1930, he went to the University of London to
learn more about the rapidly developing field of statistics. There he spent a postdoctoral
year with R. A. Fisher and Karl Pearson——world leaders in statistics. Ogburn and
Fisher in particular greatly influenced him, and he often mentioned them in lectures and
seminar discussions. His rigorous body of work consistently reflects the training he
received from these teachers. His important statistical contributions include the H-
technique for making Guttman cumulative scales practical (Stouffer 1962, chapter 14),
and Stouffer’s Z, a straightforward method for combining probabilities (Mosteller and
Bush 1954).

Stouffer’s remarkable energy and single-minded devotion to his research are
legendary. I could tell dozens of stories of his exploits, but I will restrict myself to a few
revealing anecdotes.

In 1948, national political surveys were put in bad repute by predicting that Harry
Truman would lose the presidential election to Thomas Dewey. Fred Mosteller, the
statistician, loved to tell the story of this survey debacle. He wondered how Stouffer
would react, but he should have known. Early the next morning, he saw Stouffer plowing at top speed across Harvard Yard to Emerson Hall, fired up to begin a major study as to why the surveys had gone wrong.

Stouffer was a heavy cigarette smoker. But, consistent with his intense style, he was extremely careless about his cigarette ashes. He regularly allowed the ashes to drop without noticing them. As might be expected, this habit had unfortunate consequences. At regular intervals, small paper fires would spring up from ashes falling on his desk while he worked preoccupied as ever. Usually, he took these incidents in stride. He would pop out from his office and calmly ask his frightened secretary to hand him the fire extinguisher. But once I witnessed a different reaction. He sprang from his office with uncharacteristic alarm as smoke poured out. What made the difference this time was that it was his annual tax returns that were going up in smoke!

In the fall of 1957, I returned to Harvard as an assistant professor of social psychology in Harvard’s then-titled Social Relations Department. I audited Stouffer’s famous survey analysis graduate seminar. In September 1957, the racial desegregation crisis at Little Rock, Arkansas’s Central High School burst forth on the very same day the seminar held its first meeting.

Stouffer rushed excitedly into the seminar room. He looked around the room at the students waiting for the seminar to begin. “Campbell,” he muttered, “and you, Pettigrew—you are the two Southerners around here. I want you to leave right away for Little Rock and study what is going on down there!” The late Ernest Campbell, later of Vanderbilt University, was a visiting scholar at Harvard for the 1957–1958 academic year. We had never met, but we were the only two native Southerners within Stouffer’s
near reach. He wanted us to leave on a plane that night; but we persuaded him that our wives might need to know where we were going and what we were doing. However, we did leave on a plane two days later.

Wild as it may sound, there was a point to this hurried approach to field research. Like the good newspaperman he started out to be, Stouffer wanted to strike while the iron was hot. He strongly believed in gathering field data while events were unfolding, not months later after sharpening and leveling of people’s memories had taken place. His initial guidance was to start by interviewing a group in Little Rock that was caught in cross-pressures. Such a group, he assured us, would lay bare the underlying dynamics of both sides of the conflict.

Ernie and I decided that the white ministers of modern Protestant denominations represented the ideal target group. Most of these men were racially far more liberal and open to school integration than their congregations. However, a major problem arose from the fact that promotions in the ministerial profession are heavily influenced by how successfully one raises money and enlarges the flock. And these critical resources were under the control of their conservative congregations. As our later publications detailed (Campbell and Pettigrew 1959a, 1959b), these cross-pressures did indeed lead to intense situations in which the ministers often had to choose between their beliefs and their careers.

Stouffer did not want us to have all the fun. He soon joined us for a few days to see how we were doing. When we told him what we had learned thus far, he took a special interest in the fundamentalist Protestant ministers. These men were not at all in conflict. They preached against racial integration at every opportunity to the delight of
their pro-segregationist congregations. They held racial segregation to be ordained by
God. Their popular actions drew members away from some of the modern
denominations, which further heightened the pressures on the liberal ministers.

Given his special interest, we took Stouffer to a Sunday evening service at a small
church led by one of Little Rock’s most fervent segregationist ministers. Stouffer soon
made it obvious that he was far more comfortable in this setting than we were. Drawing
on his small-town Methodist boyhood, he knew the words to many of the hymns and was
thoroughly familiar with the Old Testament quotations that laced the heated sermon. But
the Biblical distortions that flowed from the pulpit upset him. “My God,” he said in a too-
loud voice to both of us, “has this guy got a racket! He takes selected Biblical passages
out of context, twists them around, and draws crazy conclusions from them!”

In the small hall, many present must have heard him, and Ernie and I were
worried that he had blown our cover. But we should have never underestimated Stouffer.
Upon leaving the church at the close of the service, he shook the minister’s hand warmly.
He told him that the sermon had recalled his childhood, and pointedly suggested
integrationist Biblical passages from the New Testament that he then proceeded to cite
verbatim. Even this hardened segregationist melted under Stouffer’s charm to our
surprise and relief.

Stouffer was an inspiring but unorthodox teacher. Instruction from him was
invariably informal and empirical. Intensely engrossed in his work, Stouffer taught by
example. You followed him around from office to computing room and back, absorbing as
best you could his excitement, expertise and “feel” for survey research and analysis. To this
day, I have never lost the sense of excitement and curiosity in analyzing survey data instilled
by these memorable occasions. If a member of his graduate seminar on survey analysis would offer an interesting hypothesis, he would often leap from his chair and exclaim, “Let’s test it!” At which point he would lead the entire class to the machine room and start stuffing the survey data cards into the old IBM 101 counter, sorter, and printer——the leading machine for social science in the early 1950s.

While he was intensely focused on his work, Stouffer often displayed a robust, even puckish, sense of humor. When Gordon Allport lost money on his blueberry crop at his summer retreat in Maine, Stouffer began to call him, to Allport’s amusement, “the blueberry king!”

Stouffer loved his family, his work, baseball, Mickey Spillane mysteries, and Shakespeare. His favorite Shakespearian citation from *King Henry IV* has Glendower asserting that he could “call spirits from the vasty deep.” But, Hotspur retorts, “Why, so can I, or so can any man; but will they come when you do call for them?” When Stouffer called for them from survey data, the answers always came.

**Origins of the Relative Deprivation Concept**

Stouffer resisted sociology's fondness for “grand theory.” Consistent with his emphasis on empirical research, he believed in close-to-the-data reasoning and middle-level concepts. For example, he advanced the concept of “intervening opportunities” to account for migration flows (Stouffer 1940, 1962). Thus, to understand the massive African American migration up the Mississippi River to Chicago in the early twentieth century, one has to take into account not only the distance to Chicago but the intervening opportunities offered by Memphis, St. Louis, and other communities along the migration route.
The most famous illustration of Stouffer’s talent for middle-range concepts comes from his most celebrated work—the monumental World War II *American Soldier* studies (Stouffer et al. 1949a; Stouffer et al. 1949b; Stouffer et al. 1950; Stouffer 1962, chapter 2).

Stouffer devised relative deprivation (RD) as a post hoc explanation for the well-known anomalies from these studies. Two examples became especially famous. He and his wartime colleagues found that the military police were more satisfied with their slow promotions than the air corpsmen were with their rapid promotions. Similarly, African American soldiers in southern camps were more satisfied than those in northern camps despite the fact that the racist South of the 1940s remained tightly segregated by race.

These apparent puzzles assume the wrong referent comparisons. Immediate comparisons, Stouffer reasoned, were the salient referents: the military police compared their promotions with other military police—not air corpsmen whom they rarely encountered. Likewise, black soldiers in the South compared their lot with black civilians in the South—not with black soldiers in the North who were out of view. Satisfaction is relative, he held, to the available comparisons we have. Relative deprivation became a major social science concept, because social judgments are shaped not only by absolute standards but also by standards set by social comparisons (Pettigrew 1967, 1978; Smith et al. 2012; Walker and Smith 2001).

A comparative approach to explain his *American Soldier* puzzles came readily to Stouffer. Comparable ideas had been advanced by both Karl Marx (1847) and Alexis de Tocqueville (1857). Moreover, between the world wars, the theory of general relativity won acceptance in physical science while cultural anthropology firmly established a relativity perspective in social science. Hyman (1942) was the first to reflect this emphasis in social
psychology with his introduction of reference group theory. Following World War II, comparative ideas abounded in American social science. In economics, for example, another Harvard professor introduced a related concept in the same year that Stouffer introduced RD. Duesenberry (1949; Dybvig 1995) fashioned “the ratchet effect” (an intolerance for any decline in the standard of living) to help explain consumer behavior.

Following closely Stouffer’s usage of the concept, we can define relative deprivation as a judgment that one or one’s ingroup is disadvantaged compared to a relevant referent, and that this judgment invokes feelings of anger, resentment and entitlement. In addition to the fundamental feature that the concept refers to individuals and their reference groups, note that there are three critical components of this definition that are frequently overlooked. Individuals undergoing RD experience in turn three psychological processes: 1) they first make cognitive comparisons, 2) then cognitive appraisals that they or their ingroup are disadvantaged, and finally 3) that these disadvantages are seen as unfair and arouse angry resentment. If any one of these three requirements is not met, RD is not operating (Smith et al. 2012).

Defined in this manner, relative deprivation is a social psychological concept par excellence. It postulates a subjective state that shapes emotions, cognitions, and behavior. It links the individual with the interpersonal and intergroup levels of analysis. It melds easily with other social psychological processes to provide more integrative theory—a prime disciplinary need (Pettigrew 1991). Moreover, RD challenges conventional wisdom about the leading importance of absolute deprivation. And it has proven useful in a wide range of areas,—as I shall review.
The concept of relative deprivation has now won wide acceptance in criminology (e.g., Lea and Young 1984), economics (e.g., Duclos and Gregoire 2002), and throughout the social sciences. To be sure, two social scientists have used the concept extensively without referencing any work whatsoever of social psychology (Frank 2007; Varshney 2002). Indeed, it has recently even appeared in a popular book, *David and Goliath*, by the journalist, Malcolm Gladwell (2013)—although, unfortunately, it was grossly misused in this bestseller.

After Stouffer introduced RD, Merton (1957; Merton and Kitt 1950) enlarged the idea within a reference group framework. Building on this framework, Davis (1959) provided a formal mathematical model of RD. This work led me (Pettigrew 1967) to point out that RD was but one of a large family of concepts and theories that employed relative comparisons in both sociological and psychological social psychology. From sociology, this theoretical family embraces Hyman’s (1942, 1960) and Merton’s (1957) reference group theory, Lenski’s (1954) concept of status crystallization, Blau’s (1964) concept of fair exchange, and Homans’ (1961) concept of distributive justice. From psychology, these social evaluation ideas include Walster, Walster, and Bersheid’s (1978) equity theory, Festinger’s (1954) social comparison theory, and Thibaut and Kelley’s (1959) concept of comparison level.

Many theories in social psychology burn hot and then suddenly cool. But RD and related ideas have simmered slowly on a back burner for two-thirds of a century. First, Runciman (1966) broadened the RD construct by his invaluable distinction between egoistic (individual) and fraternal (group) RD. People can believe that they are personally deprived (IRD: individual RD) or that a social group to which they belong and identify is deprived
(GRD: group RD). Feelings of GRD should be associated with group-serving attitudes and behavior such as collective action and outgroup prejudice, whereas IRD should be associated with individual-serving attitudes and behavior such as academic achievement and property crime. This is a crucial point for RD theory and research which I shall call “the fit hypothesis.”

During the 1970s, Suls and Miller (1977) edited a volume that offered a host of interconnections between these comparative concepts and such other disciplinary concerns as causal attributions. Crosby (1976, 1982) provided a detailed analysis of individual RD and applied it to working women. Albert (1977) focused on temporal comparisons in an influential article. The following decade witnessed further progress. Mark and Folger (1984) introduced their referent cognitions model of RD, and a major work edited by Olson, Herman, and Zanna (1986) offered additional connections between social comparison theory and RD. A volume edited by Walker and Smith (2001) continued these advances by linking RD theory with attribution, social identity, self-categorization, and equity theories as well as procedural justice and counterfactual thinking processes.

To be sure, not all work on RD has proven useful. In political science, Gurr (1970) wrote a widely cited book, Why Men Rebel. The volume largely ignores social psychological work on the subject, and the fact that RD is a phenomenon of individuals—not societies. Gurr employed such gross macro-level measures of RD as economic and political indices of whole societies. It is a macro-level study that does not involve the micro-level of Stouffer’s conception of RD. Why Men Rebel uncovered some intriguing
findings, but it is not a RD study. As I shall discuss later, criticism of this work in the social movement field mistakenly cast RD as having little value.

A classic ecological fallacy occurs when macro-level findings are placed within a RD micro-level framework (Robinson 1950; Pettigrew 1996, 2006). That is, micro-level phenomena, such as the RD of individuals, are being erroneously assumed from macro-phenomena. This mistake is often seen in loose statements made about individual voters from aggregate voting results. It is a fallacy because macro-units are too broad to determine individual data, and individuals have unique properties that cannot be inferred from macro data. In his famous volume on *Suicide*, Durkheim ([1893] 1951) fully understood this issue. Although he could not obtain individual data, he strove to obtain data from ever-narrowing reporting districts to reduce the bias.

Indeed, the central thrust of RD theory is that individual responses are often different from that which is expected of the macro-category. For instance, given contrasting comparisons, the rich can be dissatisfied and the poor content—just the opposite from that which their macro-income characteristics would indicate. Complicating the issue further, more advantaged members of disadvantaged groups often engage in protest actions rather than the most disadvantaged (Gurin and Epps 1975; Pettigrew 1964; Sears and McConahay 1970). Although not the most objectively deprived group members, they are the most likely to make subjective social comparisons with members of advantaged groups in part because they are more likely to have contact with them (Pettigrew 1964; Pettigrew and Tropp 2011; Taylor and Moghaddam 1994). As we shall see, the ecological fallacy has seriously hurt the development of RD theory in sociology.
In short, RD makes the claim that absolute levels of deprivation of individuals—much less collective levels of deprivation - only partly determine feelings of dissatisfaction and injustice. Imagined alternatives, past experiences, and comparisons with similar others also strongly influence such feelings (Olson, Herman and Zanna 1986; Masters and Smith 1987; Walker and Smith 2001). Relative deprivation describes these subjective evaluations by individuals.

The RD concept offers social scientists an elegant way to explain numerous paradoxes (Tyler and Smith 1998). For example, RD explains why there is often little relationship between objective standards of living and satisfaction with one’s income (Strumpel 1976). Thus, the objectively disadvantaged are often satisfied with receiving low levels of societal resources, while the objectively advantaged are often dissatisfied with high levels of societal resources (Martin 1986; Pettigrew 1964). RD models suggest that the objectively disadvantaged are frequently comparing themselves to others in the same situation or worse, while the objectively advantaged are frequently comparing themselves to those who enjoy even more advantages than they possess.

Thus, the RD concept has inspired a vast international and cross-disciplinary literature. Yet the concept’s initial promise as an explanation for a wide range of social behavior remains unfulfilled. Some investigations strongly support RD models (e.g., Abrams and Grant 2012; Isaac, Mutran, and Stryker 1980; Runciman 1966; Vanneman and Pettigrew 1972; Walker and Mann 1987). But other studies do not (e.g., Gaskell and Smith 1984; McPhail 1971; Snyder and Tilly 1972; Thompson 1989). In response to these inconsistencies, previous literature reviews have sought to clarify the theoretical antecedents and components of the concept (Crosby 1976; Martin 1986; Walker and Pettigrew 1984), or to dismiss its
value altogether (Finkel and Rule 1987; Gurney and Tierney 1982; McPhail 1971; Snyder and Tilly 1972).

Two Initial Relative Deprivation Problems

Two fundamental problems of the RD literature help to explain these persistent discrepancies in results. First, in *The American Soldier*, Stouffer did not measure RD directly; rather, as noted earlier, he inferred it as a post hoc explanation for a series of surprising results. This failure to initiate a prototype measure has led to literally hundreds of diverse and often conflicting measures that have bedeviled RD research ever since.

Indeed, many of the measures throughout the social scientific research literature purporting to tap relative deprivation do not meet the basic features of the concept just outlined. One especially prevalent example involves the Cantril-Kilpatrick Self-Anchoring Scale (Cantril 1965). This measure has respondents place themselves on a ten-step ladder with the top rung labeled as the best possible life and the bottom rung as the worst possible life. When defined in this way, this scale measures discrepancies between people’s attainments and aspirations. But it does not measure discrepancies between their current status and expectations as to what they want and deserve and how they feel about it (Finkel and Rule 1987; Smith et al. 2012). Thus, this measure emphasizes RD’s cognitive component at the expense of its crucially important affective component.

Second, Stouffer offered a concept, not a testable theory. This may have enhanced its popularity, but it, too, has restrained development of the idea. Only recently have full-fledged theories emerged that allow direct testing and falsification. In the 1980s, Heather Smith and I decided that what was needed was a meta-analysis of the far-
flung research literature that employed the concept (Smith et al. 2012). As with most extensive meta-analyses, it took a while to complete the Herculean task—twenty five years of an off-and-on effort to be exact.

**Meta-analytic tests of Relative Deprivation**

Our first task was to clear the underbrush that had sprung up due to the lack of a precise theoretical and measurement model. Using inclusion criteria that ensured that RD as advanced by Stouffer was indeed being tested, a huge 76 percent drop-off occurred. While we initially secured 860 studies that purported to study some aspect of RD, only 210 met our modest criteria and entered the meta-analysis. ¹ Failing to exclude these marginal studies has been a major problem for previous qualitative reviews that did not employ strict inclusion rules. Consequently, to a great extent the criticism of RD involves this enormous research literature that does not actually involve RD.

Like the Gurr study, much of this work is interesting and uncovers important findings; but these studies are testing other phenomena—not RD. Typically they ask if macro-indices of societal deprivation and frustration lead to collective protest. Such a hypothesis contrasts with that of RD; indeed, it is largely a rival hypothesis. It is important to note that the average effect sizes from this macro-level work are in general as large or larger than those found for RD studies; thus, their exclusion does not artificially inflate our meta-analytic RD effects.²

Our second task was to ascertain the mean effect sizes for the entire RD literature as of January 2010. This long-term effort amassed 210 separate studies, composed of 293 independent samples, 421 non-independent tests and 186,073 respondents. Three different tests provided evidence that our meta-analytic data were not altered by a
publication bias that favored positive results. (See Smith et al. 2012 for details.) The mean effect sizes that emerged were highly statistically significant but small—.106 for studies, .144 for samples, and .134 for tests.

Why are the RD effects so small? We next tested three hypotheses for an explanation. First, our affect hypothesis predicts that those RD measures that explicitly included affect would yield significantly larger effects. That is, when people are clearly angry and resentful over their perceived disadvantage, the full RD effects would emerge. As Martin and Murray (1983) insisted years ago, the affect dimension is crucial. One can detect a personal or group disadvantage but believe that it is fully justified—as system justification research has repeatedly shown (Jost, Banaji, and Nosek 2004). Indeed, experiments in New Zealand show that system-justifying beliefs act as a moderator for both IRD with well-being and GRD with political mobilization. <Query, Author: Remind audience of what these acronyms mean.> As expected, subjects high in these beliefs show smaller RD effects (Osborne and Sibley 2013). Hence, we view feelings of anger, resentment, entitlement and deservingness as basic to the RD formulation. (See Feather 2015, for a detailed analysis of these various reactions to RD.)

Our second proposition involves the fit hypothesis. This idea entails both conceptual and methodological considerations. We predicted that RD effects would be larger when the levels of analysis between RD and the outcome variable are the same (Walker and Pettigrew 1984). Put differently, we contend that RD effects are reduced when IRD is used to predict group-level phenomena and when GRD is used to predict individual-level phenomena. Surprisingly many studies mix the RD and dependent variable levels of analysis.
Our third test is purely methodological. The *research quality hypothesis* holds that the more rigorous studies will yield larger effects. The inventor of meta-analysis, Gene Glass (Glass, McGaw, and Smith 1981), broke from traditional reviewing practice by including poorly conducted studies as well as well-conducted ones in order to detect the effects of research quality. If, for example, the major effects of RD are found among the poorest conducted studies—as with the effects of psychotherapy for adult depression (Cuijpers et al. 2010), we would strongly question RD’s predictive power. But we predicted the opposite—that the most rigorous RD studies would reveal the largest effects. We defined quality in terms of the reliability of both the RD and dependent variables.

---

**FIGURE 1 ABOUT HERE**

*The meta-analytic results.* Figure 1 provides the overall results by showing the percentages of variance accounted for by various subsets of the tests. For bar A, the worst conducted RD tests have none of our three desirable characteristics and yield only an \( r \) of .079. The next bar B in the histogram shows the mean \( r \) of .134 for all 421 tests. Bar C shows a mean \( r \) of .165 for those tests that did tap affect but had neither reliable measures nor a fit between the levels of analysis of RD and the outcome variable. Bar D shows a mean \( r \) of .203 when the tests boast both fit and an affect measure but unreliable measures. Finally, bar E records the results of the optimal tests. It reveals a mean \( r \) of .230 when all three of our concerns are met—reliable measures that tap affect and have the same level of analysis between RD and the dependent variable. In addition, direct
statistical tests of our three hypotheses all provide significance levels at the .05 level. (For full details, see Smith et al. 2012.)

But how are we to interpret this final mean $r$ of .23 across many types of dependent variables? First, this effect size represents a Cohen’s $d$ of .47, which is recognized as a solid medium-sized effect in social science research (Cohen 1988). Second, $r = .23$ is consistent with mean effect sizes of published meta-analyses generally. Thus, an extensive meta-analysis of intergroup contact effects on prejudice yielded a mean $r$ of -.21 (Pettigrew and Tropp 2006, 2011). And a synthesis of more than 120 meta-analyses dealing with psychological assessment produced a mean $r$ of .27 (Meyer et al. 2001). Similarly, a synthesis of 22 meta-analyses concerning ability self-evaluations and actual performance obtained a mean $r$ of .29 (Zell and Krizan 2014). Larger effect sizes are rare in meta-analyses because they typically include a heterogeneous variety of research formats, contexts, and subjects. Accordingly, these results solidly support the importance of RD when it is tested appropriately.

**Universality of the RD phenomena.** Our meta-analysis also addresses a question too seldom raised by social psychologists—the universality of the phenomenon. Positive results were recorded from 30 different nations across the globe with widely contrasting societies and cultures (Smith et al. 2012).

**Relative versus objective deprivation.** One limiting possibility of RD effects is that they may simply reflect absolute deprivation. Relevant research rejects this possibility. Adler and her colleagues (2000) found that the subjective sense of social status was more strongly and consistently related to six health factors than objective social status. Smith and her colleagues (2012) uncovered a similar result. They located 26
studies that allow a direct comparison of the relative and absolute deprivation. All 26 studies used income as the objective measure of deprivation. The mean $r$ for the RD measures proved to be .18 as compared with an $r = .12$ for the objective deprivation measures. In terms of the percentage of variance explained, RD’s mean effects are more than twice that of absolute deprivation. These data supply yet another reason why macro-level measures of collective deprivation cannot be used to gauge the perceived RD of individuals.

The range of dependent variables. The wide-ranging RD research literature includes four broad types of dependent variables (DVs). Two explore the links between IRD and internal states and individual behavior; two others explore the links between GRD and intergroup attitudes and collective behavior. The results from each of these four domains resemble the results for the total sample.\(^3\)

1) The internal states dependent variables include such outcomes as psychological stress, depression, physical health and altered self-evaluations. Without our three moderators, the 188 tests of this domain yields a mean $r$ of .173 ($z = 16.19$). For the subset with two of our critical moderators (affect and fit—in this case, IRD), the mean effect size rises significantly to $r = .271$.

2) The individual behavior dependent variables encompass both normative (e.g., church and academic activities) and non-normative (e.g., bullying, absence from work) actions. Escape behaviors (e.g., smoking, drinking alcohol, drug use) are also included. For this category, the basic mean $r$ for 126 tests without any of our three moderators is .118. When affect and fit (in this case, IRD) are also considered, the mean $r$ becomes .142. Interestingly, we found that when affect is measured and the respondents compare
their present state with their own past—rather than comparing with other individuals or groups, the mean $r$ reaches .310. For this class of dependent variables, then, it appears that inward, personal comparisons over time are especially relevant for RD effects.

3) The intergroup attitude dependent variables consist of not only prejudice measures but variables tapping stereotypes, ingroup identification, and nationalism as well as ingroup bias and identification. For the 299 tests in this domain without any of our three moderators the mean effect is $r = .115$ For the subset boasting two of our critical moderators (affect and fit—in this case, GRD), the effect becomes strikingly larger—a mean $r = .320$.

Using straightforward measures of GRD, I have had repeated success over the years predicting prejudice and related indicators. In the 1960s, Reeve Vanneman and I showed that GRD correlated positively with white voting for the racist presidential candidate, George Wallace, and against black mayoral candidates in Gary, Cleveland, Los Angeles, and Newark (Vanneman and Pettigrew 1972). Years later, with Dutch and German colleagues, we employed three different European surveys to show that GRD significantly enhanced both blatant and subtle forms of prejudice against immigrants in each of four nations—France, Germany, the Netherlands, and the United Kingdom (Pettigrew and Meertens 1995; Pettigrew et al. 2008).

This study also found that IRD related positively to prejudice with its effect fully mediated by GRD. In other words, IRD plays its role in the prejudice process by increasing GRD which in turn increases prejudice against immigrants. In addition, these surveys socially located those with high levels of relative deprivation. For both IRD and GRD, the relatively deprived in Western Europe are poorer, pay less attention to politics,
and are more politically alienated and feel less politically efficacious (Pettigrew et al. 2008).

But does GRD contribute to our understanding and prediction of intergroup prejudice beyond that already provided by other, better-known predictors? A 2002 national probability phone survey of German citizens offers an answer by including sixteen major predictors of prejudice against resident foreigners (Pettigrew, Wagner, and Christ 2007, model 7; Pettigrew and Tropp 2011). The largest predictors, as expected from the vast prejudice research literature, are social dominance orientation, positive contact (a lone negative correlate of prejudice), and authoritarianism. But following these “big three,” GRD ranks together with political inefficacy (a close correlate of RD) as the next most important and highly significant predictors of anti-immigrant prejudice in Germany. Thus, GRD adds significantly to the prediction of prejudice even when fifteen other predictors are included in the regression.

4) The collective behavior dependent variables range from self-reported rioting and intentionally sabotaging job performance to a readiness to approve of violent politics, sign petitions, and join strikes. Note that our range of collective behavior dependent variables is somewhat broader than that employed by some other investigators. Of greatest interest to sociologists and political scientists, this category of RD research has been the most criticized.

The meta-analytic results for collective behavior are similar to those for the entire RD data set as shown in Figure 1. Before applying our three moderators (affect, fit, and reliable measures), the mean $r$ for nine tests is only .065. For the total sample, the mean $r$ is .153. For the 23 tests that involve both affect and fit (in this case, GRD) but less
reliable measures, the mean $r$ rises to .211. For the remaining 18 tests with all three of our critical moderators, the mean effect reaches $r = .234$.

**Criticism of Relative Deprivation Research**

Criticism of RD work has centered on its use as a predictor of collective protest and violence. The 1960s witnessed many studies that claimed RD was a central component of such mass activities. This trend culminated in the 1970 publication of Gurr’s *Why Men Rebel*. Then the 1970s and 1980s witnessed a torrent of publications attacking this use of RD. Indeed, such papers became almost a cottage industry in political science as well as sociology. They attracted considerable attention in part because distinguished scholars—such as Edward Muller (1980) in political science and Charles Tilly (2003), John McCarthy and Mayer Zald (McCarthy and Zald 1977, 2001) in sociology—led the way.

The first focus of this critique involved questioning the then-popular J-curve hypothesis. Davies (1962) held that after a period of improvement, expectations for continued improvement develop. But if improvements slow, a sharp discrepancy between expectations and reality occurs that can ignite protest. He cited Dorr’s rebellion in Rhode Island (1842–1843), the Russian revolution (1917), and the Egyptian revolution (1952) to support his claim. He could have added the French revolution (de Tocqueville 1857). And, indeed, later urban race riots in the United States during the 1960s followed a period of economic and political gains for the African Americans, not a period of rising injustice (Pettigrew 1964).

But is this process an intrinsic component of uprisings? Many critics thought not; they supplied data from other collective disruptions that do not fit the J-curve posited by
Davies. For example, Snyder and Tilly (1972) investigated collective violence in France from 1830 to 1960. Their time series over these 131 years did not support the J-curve model. (For a rejoinder, see Davies 1974.) Similarly, Thompson (1989) developed a time-series of political violence in Northern Ireland from 1922 to 1985 and could find no relationship between unemployment rates and fatal violence.

It should be noted that these authors did not claim to be testing RD theory. Instead, they speak of “deprivation” and “frustration.” And Davies (1962) explicitly regretted that he had to use such “crude indices” as industrial strikes and employment rates rather than “cross-sectional interviews” with individuals. The problem arose when later critics of RD cited these macro-studies as evidence against the validity of RD theory generally.

For example, McPhail (1971) crudely combined the results of ten studies of five urban American race riots and claimed that there was scant empirical support for RD as a factor in riot participation. He deserves credit for attempting to combine study results before the development of meta-analysis. But his conclusion is undermined by his use of studies that use only macro-level indices to gauge RD. This failure to use data from individuals, as Finkel and Rule (1987:58) point out, invalidates McPhail’s dismissal of RD “as a potential determinate of individual behavior.”

Gurney and Tierney (1982) published another hypercritical review of a selected subset of studies that purported to test the role of RD in social movements. These authors rightly challenge the use of macro-measures to infer RD; but then they favorably cite studies, such as McPhail’s (1971), that make use of precisely this ecological fallacy. They also ignore the crucial distinction between individual and group RD. Further, these
critics claim there is little or no relationship between attitudes and behavior—an extreme assertion that had been thoroughly repudiated by social psychologists (e.g., Dillehey 1973; Kelman 1974).

Behind much of the sociological rejection of RD is the belief that it is too psychological, that it draws attention away from the important structural factors that are basic to the sociological vision of collective protest (McCarthy and Zald 1977, 2001). I agree. The sweeping argument of the 1960s that RD could by itself explain protest movements went far beyond Stouffer’s conception and did indeed ignore the fundamental social contexts of mass movements. The problem is that the critics tended to pour the baby out with the bath water by asserting that RD played virtually no role. Important as resource mobilization and other structural theories are, they do not exclude the possibility that RD has a role in the social movement process—certainly not the all-encompassing one initially proposed, but a modest role nevertheless as indicated by our meta-analytic results. Indeed, recent work offers successful statistical models of collective action that include RD together with indicators of social mobilization and social identity (Abrams and Grant 2012; Grant et al. 2015; Van Zomeren, Postmes, and Spears 2008).

After sixty years of research in both the sociological and psychological sides of social psychology, I strongly favor efforts to join the two into a more robust contextual social psychology that combines the insights of both branches. Four decades ago, Robin Williams (1975), a close friend, had a similar view that led him to offer a broad overview of the role of RD in organized protest. He hypothesized that RD’s role in sustained protest would be maximized when there is: 1) a collective sense of RD,
especially in prestige and political power, that 2) occurs suddenly, and 3) the deprived collectivity is large, commands substantial economic and political power recently achieved and is internally cohesive, and 4) the established “control elements of the society have given signals of weakness, indecision, disunity, or actual encouragement of militant dissent.” (Williams 1975:373) Several of these predictions have already been supported (e.g., Dube and Guimond 1986), and the rest deserve careful test. Moreover, RD directly informs current political and policy discussions of income inequality. Observe how all four of Williams’s conditions describe the present-day “tea party” movement in the United States.

**A Final Word**

Actually, the debate about the role of RD in protest movements bears little relevance to Stouffer’s basic conception of RD in individuals. I discuss it because this attack on the role of RD in the social movement theory has had the unfortunate effect of causing sociologists to discard RD as a viable predictor for the great range of dependent variables that our meta-analysis has uncovered.

Brush (1996) notes that psychologists have continued to employ RD theory while sociologists have largely abandoned it; this is easily explained. Psychologists use a variety of dependent variables rather than concentrating on collective protest, and they generally employ Stouffer’s original conception of RD together with appropriate measurement.

In sum, this paper has two purposes. The first is to remind us of Samuel Stouffer’s major contributions to social psychology. These contributions, both conceptual and methodological, importantly shape the field today. The second aim was to revitalize his
invaluable concept of relative deprivation. Tested rigorously with reliable measures that tap angry resentment against dependent variables of similar scope, the meta-analytic test demonstrates that RD can be a useful theory in a wide variety of domains of central interest to social psychology.

REFERENCES


The inclusion criteria were: RD is the predictor; individuals report RD directly; RD is the product of a disadvantageous comparison; the relationship between participant and comparison target is clear; the measures are not difference scores created by the researcher; and the impact of the comparison is measured. For further details, see Smith et al. 2012.

2 This is to be expected as macro-level measures typically involve less error and routinely provide larger effects than micro-level measures.

3 For internal states, individual behaviors, and intergroup attitudes, the third moderator—reliable measures—was not included because of small sample sizes in these subsets of tests.
Figure 1
Mean Effects of Test Subsets by
% of Variance Explained

A. $r = .079$

B. $r = .134$

C. $r = .165$

D. $r = .201$

E. $r = .230$

Adapted from Smith et al. 2012.