

Deception in Experiments: Revisiting the Arguments in Its Defense

Ralph Hertwig
University of Basel

Andreas Ortmann
Charles University and Academy of Sciences of the Czech Republic

In psychology, deception is commonly used to increase experimental control. Yet, its use has provoked concerns that it raises participants' suspicions, prompts second-guessing of experimenters' true intentions, and ultimately distorts behavior and endangers the control it is meant to achieve. Over time, these concerns regarding the methodological costs of the use of deception have been subjected to empirical analysis. We review the evidence stemming from these studies.

Keywords: deception, research ethics, experimental control, suspicion

The use of deception [in experiments] has become more and more extensive. ... It is easy to view this problem with alarm, but it is much more difficult to formulate an unambiguous position on the problem. ... I am too well aware of the fact that there are good reasons for using deception in many experiments. There are many significant problems that probably cannot be investigated without the use of deception, at least not at the present level of development of our experimental methodology. (Kelman, 1967, p. 2)

In his well-known article "Human Use of Human Subjects: The Problem of Deception in Social Psychological Experiments," Herbert Kelman (1967) described his dilemma as a social scientist as that of being caught between the Scylla of the use of deception to study important social behaviors and the Charybdis of ethical

and methodological considerations (p. 2). He wrote this article in the wake of a public exchange between Baumrind (1964) and Milgram (1964) and in response to the substantial increase in the use of deception during the 1960s. Whereas the exchange between Baumrind and Milgram had focused on the ethical implications of Milgram's research on obedience, Kelman stressed the long-term consequences of deceptive practices on participants' expectations and behavior. The essence of his concern was this:

As we continue to carry out research of this kind ... our potential subjects become increasingly distrustful of us, and our future relations with them are likely to be undermined. Thus, we are confronted with the anomalous circumstance that the more research we do, the more difficult and questionable it becomes. (p. 7)

We reiterated Kelman's concern that the use of deception may contaminate the participant pool (Ortmann & Hertwig, 1997, 1998) and, to put psychology's research practices into perspective, we noted that researchers in a neighboring discipline—experimental economics—had effectively prohibited the use of deception in their experiments (Hertwig & Ortmann, 2001, 2002; Ortmann & Hertwig, 2002). Our comments prompted responses from several researchers: Bröder (1998), Kimmel (1998), Korn (1998), and Weiss (2001). Four arguments featured prominently in their defense of deception: (a) the use of deception has, after an increase in the 1960s and 1970s, dropped (e.g., Korn); (b) “the preponderance of evidence suggests that deceived participants do not become resentful about having been fooled by researchers” (Kimmel, p. 804); (c) the effects of suspiciousness on research performance “appear to be negligible” (Kimmel, p. 804); and (d) deception is an indispensable tool for achieving experimental control—at least in some socially significant areas of research (all four researchers).

Our original contributions (Ortmann & Hertwig, 1997, 1998), as well as those by our critics, were based more on assertions than on empirical evidence. It is thus the main goal of this article to present empirical evidence that bears on the first three arguments (we turn to the fourth argument in the General Discussion). Although our emphasis is on the possible methodological side effects of deception, we address possible ethical implications of our findings in the General Discussion. We do not claim to have unearthed all available evidence, but by specifying clear and transparent search criteria we believe to have made progress toward a comprehensive assessment of the available evidence.

The following review of the available evidence is a companion piece to Ortmann and Hertwig (2002), building on, complementing and extending their review of psychological research on the methodological costs of deception. Ortmann and Hertwig's review was written for economists, and it focused on the question of whether experimental economics should lift its *de facto* prohibition of deception. The present article, in contrast, addresses an audience of psychologists and discusses, among other issues, whether psychology's regulatory regime regarding the use of deception is effective.

Before we turn to the three arguments in the defense of deception, we first turn to a definition of deception and describe the reasons for its use.

WHAT IS DECEPTION?

Deception is not easily defined. Yet there seems to be considerable agreement about what definitely ought to count as deception. Such agreement is, for instance, manifest among the group of researchers who have studied the prevalence of deception as a research method in (mainly social) psychology. To find such review studies, we conducted a PsycINFO/PsycLIT literature search, using the term *deception* in combination with *frequency*. In addition, we consulted the bibliography of the obtained articles. We found 15 studies that analyzed the frequency of use of deception in various journals: Adair, Dushenko, and Lindsay (1985); Carlson (1971); Epley and Huff (1998); Gross and Fleming (1982); Kimmel (2001); Krupat (1977); Levenson, Gray, and Ingram (1976); McNamara and Woods (1977); Menges (1973); Nicks, Korn, and Mainieri (1997); Seeman (1969); Sieber, Iannuzzo, and Rodriguez (1995); Stricker (1967); Toy, Olsen, and Wright (1989); and Vitelli (1988). Adair et al., for instance, defined deception as “the provision of information that actively misled subjects regarding some aspect of the study” (p. 62). Nicks et al. defined deception as an “explicit misstatement of fact” (p. 70), and Menges described deception as instances where “the subject is given misleading or erroneous information” (p. 1032). These and other definitions reveal that intentional and explicit misrepresentation—that is, lying about, for instance, the purpose of the investigation and the identity of researcher and confederate—is unanimously considered to be deception. This consensus is also shared across disciplinary borders. In the words of economist Hey (1998), “there is a world of difference between not telling subjects things and telling them the wrong things. *The latter is deception, the former is not*” (p. 397).

Hey’s (1998) assertion, furthermore, indicates what seems to be widespread agreement among researchers: Withholding information does not necessarily constitute deception (for a different view on the role of *passive* deception, see Kimmel, 2001). That is, not acquainting participants in advance with all aspects of the research being conducted, such as the hypotheses explored and the full range of experimental conditions, is often not considered deception. In their review of deception studies, Adair et al. (1985), for instance, decided that “the simple failure to disclose the true purpose of the study was not counted as deception” (p. 63). Although Baumrind (1979) suggested that “full disclosure of everything that could affect a given subject’s decision to participate is a worthy idea” (p. 1), this strict critic of deception also conceded that “absence of full disclosure does not constitute intentional deception” (Baumrind, 1985, p. 165). Similarly, experimental economists McDaniel and Starmer (1998) described some forms of “economy with the truth” as “perfectly legitimate” (p. 406), and Hey (1998, p. 397) pointed

out that “ill-defined experiments” (i.e., when the experimenter does not inform participants about all features of the experiment) are an important tool (see Lawson, 2001, for a thorough discussion of the distinction between providing false information and withholding information).

The distinction between deception and nondeception blurs, however, when participants’ default assumptions come into play. One default assumption a participant is likely to have is that experiments start only after an experimenter has clearly indicated its beginning. As a consequence, the participant might assume that his or her initial interactions with the experimenter (upon entering the laboratory) are not an object of investigation. Should violations of such expectations be counted as deception? Some of the researchers who assessed the prevalence of deception did not appear to include such violations (Adair et al., 1985; Gross & Fleming, 1982; Nicks et al., 1997), but others did. Sieber et al. (1995) and Gross and Fleming (1982), for instance, considered participants to be deceived if they were unaware of being research participants at all or were unaware that the study had begun at the time of the manipulation. The fact that some researchers included violations of default assumptions in their definition of deception and others did not might reflect conceptual disagreement. Alternatively it could reflect a pragmatic decision on the part of researchers who struggle to quantify the prevalence of deception—violations of default assumptions are much more difficult to identify than provisions of misinformation.

In sum, a consensus has emerged across disciplinary borders that intentional provision of misinformation is deception and that withholding information about research hypotheses, the range of experimental manipulations, or the like ought not to count as deception. Common ground has not (yet) been established with respect to the violation of participants’ default assumptions. Perhaps the study of default assumption necessitates a completely different approach to the definition of deception. Although deception is commonly defined on the basis of the experimenter’s behavior (e.g., intentionally providing false information), one could define it alternatively on the basis of how participants *perceive* the experimenter’s behavior. According to such a definition, deception would have occurred if participants, after being completely debriefed, had perceived themselves as being misled. Such an approach defines deception empirically and post hoc rather than on the basis of norms devoid of context. We do not know of any attempt to realize such an “inductive” approach.

REASONS FOR DECEPTION AND TWO MECHANISMS OF CONTAMINATION

Why deceive? Deception is often justified with two arguments. The first is that deception allows the researcher to create situations of interest that are not likely to arise naturally. A good illustration of this potential is found in studies of helping behavior in emergency situations, in which researchers stage emergencies (e.g.,

someone experiences a seizure), manipulate situational factors (e.g., absence and presence of others), and then determine the impact of these factors on bystanders' willingness to help (e.g., Darley & Latané, 1968). Because emergency situations occur infrequently, it is difficult to study them experimentally, unless, so the argument goes, one fabricates them.

The second rationale for deception is that certain socially relevant aspects of behavior can only be studied if people are caught off guard (e.g., Cooper, 1976; Weber & Cook, 1972; Weiss, 2001). If they suspected or knew that some socially undesirable aspects of behavior are being observed (e.g., conformity, prejudices, antisocial behavior), then they would alter their "natural" behavior to look as good as possible to the social observers (i.e., experimenter or other participants). Consider conformity behavior as an example. If participants knew that an experiment explores the extent to which they easily give in to social pressure, then they would be less likely to show conformity behavior. Therefore, so the argument goes, studies of conformity behavior need to camouflage the purpose of the experiment to achieve experimental control. If not, then the "psychologist runs the risk of distorting the reactions of his or her subjects and ultimately limiting the applicability of the research findings" (Kimmel, 1996, p. 68).

Challenging the latter rationale, critics of deception have argued that it is the very use of deception that impairs, and eventually even destroys, experimental control, thus threatening the validity of research findings. Kelman (1967) is not the only one to have advanced this argument. Other researchers in the social sciences (i.e., psychologists, sociologists, anthropologists, and economists) have also worried that deception contaminates the participant pool. Whereas in sociology it was suggested that a likely outcome of deceptive practices is participants' future resistance to other research efforts (e.g., Erikson, 1995), psychologists and economists have expressed concern that the expectation of being deceived produces suspicion and second-guessing and that these reactions—rather than the experimenter's scenario and instructions—guide and ultimately distort experimental behavior.

The concern that deception breeds suspicion, and that suspicion, in turn, impairs experimental control comes in two variants. For some researchers, suspicion and second-guessing require *firsthand* experience with deception (i.e., participating and being debriefed in deception experiments; see Seeman, 1969); others assume that *secondhand* experience with deception—for instance, stemming from undergraduate psychology classes, campus scuttlebutt, media coverage of psychological research, and the profession's reputation more generally—suffices to engender in participants the expectation that they will be deceived (e.g., Adelson, 1969; Davis & Holt, 1993; Ledyard, 1995; Orne, 1962; Ring, 1967). The latter assumption is particularly common in experimental economics where deception is effectively prohibited (e.g., see Davis & Holt, 1993; Hey, 1991).

The distinction between firsthand and secondhand experiences is relevant because these different experiences with deception imply potentially different degrees of contamination of the participant pool. If secondhand experience sufficed

to induce suspicion and second-guessing, then the potential side effects of deception would likely be widespread and extend beyond participants with firsthand experience. In contrast, if firsthand experience were necessary to induce suspicion and second-guessing, then contamination would be more contained. In addition, the argument advanced to defend the practice of deception—that its use has declined since its peak in the 1970s—would then gain additional weight.

IS THE USE OF DECEPTION IN DECLINE?

One argument in the defense of deception is that, after an increase in the 1960s and 1970s, the use of deception has dropped. As mentioned earlier, we found 15 studies that analyzed the frequency of use of deception across a wide range of journals, including fields other than social psychology (see the aforementioned search strategy). Owing to these studies, there is now ample evidence that deception is by no means confined to social psychology. In consumer research, for instance, the American Psychological Association (APA; 2002) ethics code is the primary code of conduct for the most common research methods in the field (N. C. Smith, Klein, & Kimmel, 2002). Based on an analysis of the *Journal of Marketing Research (JMR)* and *Journal of Consumer Research (JCR)*, Kimmel (2001) observed a rise in published deception studies in marketing research, from 43% in 1975/1976 to 56% in 1996/1997.

What are the trends in social psychology journals? To answer this question, we turn to the two leading outlets of social psychological research, namely, the *Journal of Personality and Social Psychology (JPSP)* and the *Journal of Experimental Social Psychology (JESP)*. Focusing on these journals has the additional benefit of the most comprehensive data available. Both journals were founded in 1965 (*JPSP* emanated from the *Journal of Abnormal and Social Psychology*). As Figure 1 (top panel) shows, in *JPSP*'s first year of existence about half of its articles involved deception. In the second half of the 1970s, the use of deception peaked, with 73% of all articles involving deception in 1978. In the 1980s, the trend reversed and deception became a markedly less frequently used tool, with about 30% of deception studies in 1989. This decline appears to have come to a halt in the 1990s; deception now appears to be used in about one third to two fifths of all publications.

In *JESP*, trends in the use of deception were quite similar to those observed for *JPSP*, except that the proportion of deception studies was consistently higher (see Figure 1, bottom panel). This trend was already evident in the first year of publication: in 1965, 85% of articles published in *JESP* involved deception. Until the late 1970s, the figures fluctuated at around 70%. Similar to *JPSP*, the change in the propensity to use deception came in the second half of the 1980s. Specifically, in 1987 the proportion of deception studies dropped to 43%. Subsequently, the figures fluctuated, with 66% in 1989 and 50% in 1994. To get a sense of whether a trend toward more or less use of deception prevailed, we analyzed the frequency of de-

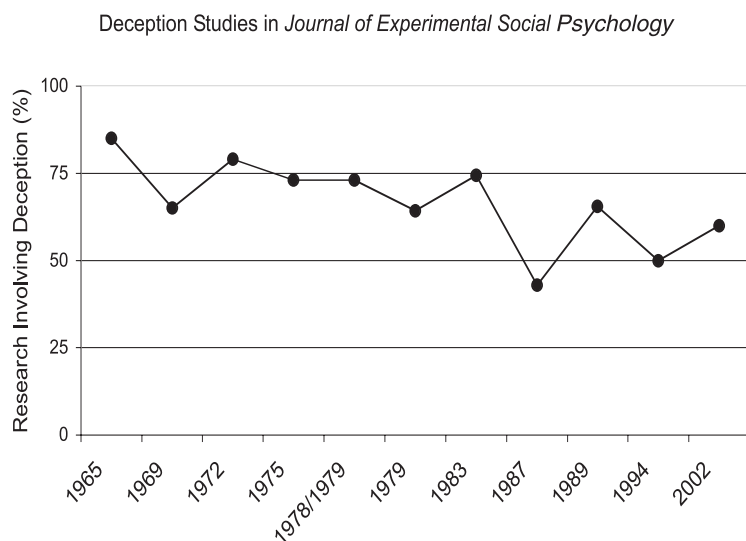
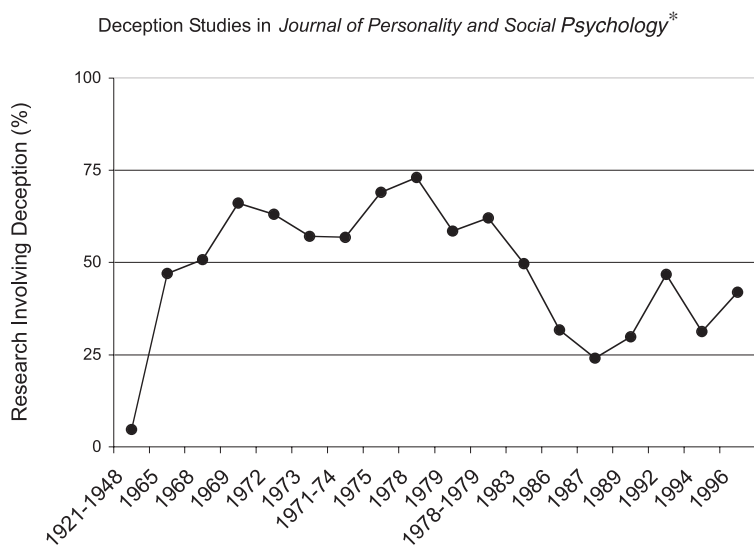


FIGURE 1 The proportion of articles employing deception in the *Journal of Personality and Social Psychology* (top panel), and the *Journal of Experimental Social Psychology* (bottom panel). (Source of data for *JPSP*: Adair et al., 1985; Epley & Huff, 1998; Gross & Fleming, 1982; McNamara & Woods, 1977; Menges, 1973; Nicks et al., 1997; Sieber et al., 1995. Source of data for *JESP*: Adair et al., 1985; Gross & Fleming, 1982; Nicks et al., 1997; the data for 2002 stem from our own analysis, see text.)

*The data for 1921–1948 represent data for *Journal of Abnormal and Social Psychology*.

ception in 2002. Specifically, we coded each study (of each article published in *JESP* in 2002) according to whether deception was used. When deception was used, we also recorded the aspect of the study about which participants were deceived. To this end, we used the taxonomy of methods of deception in Sieber et al. (1995). Table 1 lists their eight methods of deception and the percentage of deception studies that employed each of the eight methods in our sample of studies.

In 2002, *JESP* published 27 articles and 32 reports (the latter are subject to a length limitation), which encompassed 117 studies. Of all 117 studies, 63 (53%) used deception, and 36 (61%) publications reported at least 1 study that used deception. As these figures show, the use of deception in published studies in *JESP* has not become a marginal phenomenon: More than half of the studies drew on at least one method of deception, and, on average, deception studies made use of two to three different methods of deception. The results in Table 1 confirm trends previously observed by others (e.g., Sieber et al., 1995). Specific methods of deception such as using confederates or keeping people unaware of their participation in

TABLE 1
Percentage (and Number) of Studies Published in the *Journal of Experimental Social Psychology* in 2002 Using a Classification of Methods of Deception Proposed by Sieber et al. (1995)

<i>Method of Deception</i>	<i>How Many of the Deception Studies Use a Given Method</i>
<i>False purpose.</i> Participants are given, or be caused to hold, false information about the main purpose of the study	87% (55)
<i>Bogus device.</i> Participants are given false information concerning stimulus material ^a	62% (39)
<i>Role deception.</i> Participants interact with participants about whose identify they have been given false information	24% (15)
<i>False feedback regarding self.</i> Participants are given false feedback about themselves	30% (19)
<i>False feedback regarding others.</i> Participants are given false feedback about another person	24% (15)
<i>Two related studies.</i> Two related studies are presented as unrelated	9.5% (6)
<i>Unaware of measure.</i> Participants are kept unaware that a study is in progress at the time of manipulation or measurement, or unaware of being measured (e.g., videotaped)	3% (2)
<i>Unaware of participation.</i> Participants are kept unaware of being subjects in research	0%

Note. There were 117 studies, of which 63 used deception.

^aAs a *bogus device* we coded every instance in which experimenters made false statements about key aspects of the stimulus material. For illustration, instances in which participants received a photocopy of a bogus sign-up sheet that had ostensibly been filled out by students, were given bogus answers of a potential dating partner, and were falsely told that two questionnaires were written by two different researchers were coded as cases of *bogus device*.

a study (both characteristic of some classic deception studies) are rarely used or not used at all. Other methods such as the use of bogus devices (which may at least partly driven by the ubiquitous use of computers in the laboratory) have almost become a default tool among deception studies.

In drawing conclusions from the results in Figure 1, one needs to be careful in interpreting each and every increase and decline, respectively. Clearly not all of the reported changes mirror evolving methodological preferences, ethical standards, or federal regulations of research. Some of the fluctuations may simply stem from different views of what constitutes deception. (Compare, for instance, the criteria employed by Sieber et al., 1995, with those used by Nicks et al., 1997). Notwithstanding this qualification, however, the following picture emerges: Compared to the zenith of deception in the late 1960s, 1970s, and early 1980s, the use of deception in social psychology has declined in the last 20 years. In both *JESP* and *JPSP*, the turning point occurred in the second half of the 1980s. Since then deception has no longer been the pervasive methodology that it was in the heyday of deception (but see the different trend in consumer psychology).

This change supports the defenders' argument for the use of deception in contemporary studies: Deception is simply not as frequent a phenomenon as it used to be. Figure 1, however, also demonstrates that regardless of the decline in the use of deception its prevalence is not trivial. A conservative estimate is that every third study published in *JPSP* in the 1990s employed deception, and half of all studies published in *JESP* in the 1990s and 2002 still drew on deception. Returning to the distinction of firsthand and secondhand experiences with deception, these figures also suggest that it is not an unlikely event, even today, for students to experience deception personally. Thus, in our view, the argument that frequency has dropped cannot easily allay concerns about possible methodological side effects of deception. Can the next argument in defense of deception dispel them?

DOES DECEPTION BREED NO RESENTMENT?

Based on his review of research about the effect of deception, Christensen (1988) concluded, "This review of the literature, which has attempted to document the impact of deception on research participants, has consistently revealed that research participants do not perceive that they are harmed and do not seem to mind being misled" (p. 668). Curiously, in his review of the evidence Christensen did not include his own study conducted a decade earlier. There, he concluded, "the primary conclusion that can be drawn from the present two studies is that subjects who have knowingly participated in a manipulative experiment will attempt to resist such a manipulative intent in future manipulative experiments" (1977, pp. 399–400).

More recently, Kimmel (1998) concluded that the "preponderance of evidence suggests that deceived participants do not become resentful about having been

fooled by researchers" (p. 804). According to *Merriam-Webster's Collegiate Dictionary*, resentment is "a feeling of indignant displeasure or persistent ill will at something regarded as a wrong, insult, or injury" (p. 1059). Does the empirical evidence indeed suggest that participants do not harbor such feelings? To avoid opportunistic sampling of evidence (see Christensen's, 1988, oversight), we referenced here all published journal articles that Kimmel (1998, pp. 104–107) in his recent review cited in support of his conjecture that the negative effects of deception appear to be minimal. We also included other articles in his support that we encountered outside of his review. Finally, we attempted to unearth further studies that gauged students' feeling of resentment (or lack thereof) about the use of deception in psychology experiments. Using the term *deception* in combination with either *feelings* or *resentment*, we conducted a literature search using PsycINFO/PsycLIT. These searches did not turn up any further hits. Before we turn to the articles, one clarification is in order: We summarize the results in a qualitative rather than quantitative way. What stands in the way of a more meta-analytical treatment of the studies is that they can hardly be compared on just a few key dimensions: Deception varies widely in type and degree, and the dimensions on which the effects are measured vary enormously.

Kimmel's (1998) conclusion seems to rest on the following five observations: First, participants in general do not seem to express negative feelings (i.e., regret having participated) about their experience in deception experiments (e.g., Milgram, 1964; Pihl, Zacchia, & Zeichner, 1981; Ring, Wallston, & Corey, 1970; C. P. Smith, 1981). Second, participants endorse the scientific utility of deception experiments (Clark & Word, 1974; Gerdes, 1979) and seem to be prepared to tolerate deception in the interest of research (Aitkenhead & Dordoy, 1985). Third, participants in deception experiments report having enjoyed the experience more, having felt less bored, and having perceived more educational benefit from their participation than participants in nondeception experiments (e.g., Finney, 1987; S. S. Smith & Richardson, 1983). Fourth, most college students are generally accepting of ethically sensitive research practices such as deception and invasion of privacy (e.g., Collins, Kuhn, & King, 1979; Epstein, Suedfeld, & Silverstein, 1973; Farr & Seaver, 1975) and are less critical of those practices than members of Human Subjects Committees, psychologists, graduate students, and faculty (e.g., Korn, 1987; C. P. Smith & Berard, 1982; Sullivan & Deiker, 1973). Fifth, according to a questionnaire study by Sharpe, Adair, and Roese (1992), the continued use of deception did not evoke an increase in negative attitudes toward psychological research among the participant population.

Based on these observations, Kimmel (1996) concluded that "the negative effects of deception appear to be minimal" (p. 104). A different series of observations, however, provides less reason for such optimism. Fisher and Fyrberg (1994), for instance, reported that the majority of their students believed that participants

in various published deception studies must have felt embarrassed, sad, or uncomfortable. In one experiment, D. F. Allen (1983) found that only participants who had been deceived during the session “rated the experiment as worthless, were annoyed with the experiment, and would not recommend the experiment to a friend” (p. 899; see also Straits, Wuebben, & Majka, 1972). Moreover, Cook et al. (1970, p. 189) found that participants with a history of deception studies considered experiments to be less scientific and less valuable and reported caring less about understanding and following experimental instructions. In addition, Epstein et al. (1973) reported that, next to danger to the participant, deception is the most frequently mentioned reason for withdrawing from an experiment. Oliansky (1991) observed that deception—in this particular case the wrong impression that one has control over another participant, who was in reality a confederate—might trigger severe negative emotions in (some) participants.

We can think of two reasons as to why the evidence regarding people’s feelings is so mixed. First, deception as used in Aitkenhead and Dordoy (1985) is not deception as used in Finney (1987), which is not deception as used in Oliansky (1991). In other words, whether being fooled lies within a participant’s “comfort zone” (Gerdes, 1979) is probably a function of the nature and severity of the deception. Second, participants react on different levels, and negative feelings may not automatically translate into behavior: For instance, in a replication of Asch’s (1956) line-judgment task, Finney observed that deceived participants were more depressed, hostile, and anxious than nondeceived participants; yet their uneasiness did not cause them to avoid future psychological research or to question the study’s scientific value.

Whatever the reasons for the mixed results may be, it seems fair to conclude that the issue of how pervasively deception raises resentment is not yet decided. We propose that one way to further elucidate this issue is to consult related research on the consequences of deception in social interactions beyond those of experimenter and participant. The results of a still small set of negotiation and strategic interactions (i.e., games) studies suggest that being deceived in social interactions has the potential to evoke a wide range of responses, ranging from diminishing desire for future interactions to attribution of untrustworthiness (Boles, Croson, & Murnighan, 2000), to a substantial taste for retribution and for punishment by the deceived players—a taste for which they are even willing to sacrifice money (e.g., Boles et al., 2000; Brandts & Charness, 2002; Croson, Boles, & Murnighan, 2003; see also Schultz, 1969). This research also shows that if people expect lies and deception, then they might not necessarily respond negatively once their expectations are met (Lewicki & Stark, 1996), thus raising the possibility that those students who do not resent being deceived may be the ones who expect deception as part of the game. Such an expectation can, of course, also jeopardize experimental control.

ARE THE EFFECTS OF SUSPICION NEGLIGIBLE?

Psychological experiments may provoke a dynamic that Riecken (1962) described as follows:

The fact that the experimenter controls the information available to the subject and that he never reveals completely what he is trying to discover and how he will judge what he observes—this feature gives the experiment much of its character as a game or contest. It leads to a set of inferential and interpretive activities on the part of the subject in an effort to penetrate the experimenter's inscrutability (p. 31).

Are participants who suspect the experimenter to be lying even more eager to undo this information asymmetry? If so, one may expect the behavior in experiments of participants who suspect foul play to differ from those who do not. Based on his review of the literature, Kimmel (1998), however, arrived at a different conclusion. In his view, "the effects of suspiciousness on research performance, although somewhat inconsistent, appear to be negligible, leading some to conclude that, in general, there are not major differences between the data of suspicious and reportedly naïve participants" (p. 804).

Are the effects "negligible," as has been argued in the defense of deception? Some observational data suggest this may not be so. Take, for example, the following incident. In the middle of a mock jury study, one of the six jurors experienced a genuine epileptic seizure reminiscent of the feigned seizure that served as a manipulation in a classic study by Darley and Latané (1968). The experimenters, MacCoun and Kerr (1987), reported that "three of the five subjects spontaneously reported that they had questioned the authenticity of the attack" (p. 199) and that "there were indications that prior knowledge of psychological research, derived primarily from course work, was related to suspicion" (p. 199). The only person who promptly came to the victim's aid had no prior psychology coursework, whereas "two of the other bystanders reported looking for hidden observers even as they joined in administering aid" (p. 199). Had MacCoun and Kerr's study been concerned with altruistic behavior, then the participants' behavior, that is, withholding help because they were suspicious of deception and expected to be framed, would have been mistaken as evidence for the "bystander effect" (Darley & Latané, 1968).

Is this just a singular incident in which suspicion compromises experimental data and conclusions? We address this question by analyzing three sets of studies that render possible three independent tests of the conjecture that the effects of suspicion are negligible. The first set of studies compares conformity behavior of participants who were identified post-experimentally as being either suspicious or unsuspecting of deception. The second set consists of studies that intentionally provoked the expectation of deception at the outset and then examined experimental behavior as a function of it. In the third set of studies, participants' experimental history (e.g., pre-

vious participation in deception studies) was either recorded or systematically manipulated and their experimental behavior studied as a function of it.

To avoid the risk of opportunistic sampling of studies, we performed a systematic electronic literature search (using the keywords listed below). We searched for specific keywords in titles and abstracts of articles listed in the PsycINFO/PsycLIT database, which covers the academic literature in psychology and related disciplines, including sociology and education, in the period between 1887 and June 2006. We also included all studies cited in a recent review by Bonetti (1998a), who concluded that “deception does not appear to ‘jeopardize future experiments’ or ‘contaminate a subject pool’” (p. 389). Finally, we looked up the studies cited in the articles found using the first two methods and included them if they could be classified into one of the three sets that we used to examine the effects of suspicion.

Test 1: Are the Effects of Self-Reported Suspicion on Conformity Behavior Negligible?

To find studies that examined the effects of postexperimentally identified suspicion, we searched for *deception* in combination with *suspicion* (and its variants, such as *suspicious*, *suspiciousness*, *suspicious*). Our search uncovered two systematic reviews of the social psychology literature that examined the prevalence of suspicion among participants. The studies reviewed by Stricker (1967) excluded, with one exception, suspicious participants, and thus his review does not allow us to examine how suspicion affected experimental behavior. In his review of the literature on social conformity, Stang (1976) found 21 studies that reported the percentage of “suspicious” participants. Out of the 21 studies, Stang (1976, p. 363) cited 9 that systematically compared the behavior of suspicious and unsuspecting participants. Typically, this classification was performed on the basis of post-experimental interviews in which participants responded to questions such as “Do you feel this experiment was deceptive (involved lying) in any way?” (Geller & Endler, 1973, p. 49). In addition to those 9 studies referenced by Stang, our search turned up 5 studies that examined behavior in experiments as a function of suspicion, all of which were also concerned with conformity behavior. It is probably no coincidence that researchers studying conformity have been particularly concerned with the possible repercussions of suspicion. According to Gross and Fleming (1982), researchers in this area have relied heavily on deception, with 96.7% of studies in the area of compliance and conformity having used deception.

As shown in Table 2, in 10 of 14 studies identified by Stang and our additional search, suspicious participants showed less conformity behavior—the target variable in which experimenters were interested—than unsuspecting participants. In 4 studies (Chipman, 1966; Endler, Wiesenthal, & Geller, 1972; Wiesenthal, Endler,

TABLE 2
Are the Effects of Suspicion on Conformity Behavior Negligible?

<i>Authors</i>	<i>Proportion of Suspicious Participants</i>	<i>Experimental Performance of Suspicious Participants (Effect Size^a)</i>
V. L. Allen (1966)	30 of 120 (25%)	Less conformity: On a maximum score of 100% conformity, unsuspecting participants scored on average 26% and suspicious participants 12%.
Stricker, Messick, and Jackson (1967)	38.6% (averaged across sex and suspicion about various aspects of the experiment; see their Table 1).	Less conformity: $r = .49$ (their Table 4), $r = .33$ (their Table 5); averaged across sex and measures of conformity.
Glinski, Glinski, and Slatin (1970)	Sessions 1 and 2: 42 of 55 (76%)	Less conformity: $r = .89$ (Session 1), $r = .86$ (Session 2)
Eitinger, Marino, Endler, Geller, and Natziuk (1971)	15 of 40 (38%)	Less conformity: $eta = .33$
Endler, Wiesenthal, and Geller (1972)	No data	No difference in conformity (no "significant" main effect)
Endler and Hartley (1973)	14 of 40 (35%)	Less conformity: $eta = .31$
Geller and Endler (1973)	28 of 54 (52%)	"Once subjects become suspicious, their conformity sharply decreases" (p. 52): $eta = .6$
Geller, Endler, and Wiesenthal (1973)	21 of 61 (34%)	Less conformity: $eta = .33$
Wiesenthal, Endler, and Geller (1973)	96 of 116 (83%)	No difference in conformity (nonsignificant t test)
Chipman (1966) ^b	19 of 68 (28%)	No significant difference in conformity
Willis and Willis (1970) ^b	54.2%	Little to no effect
Rubin and Moore (1971) ^b	95 of 142 (67%) were either medium or highly suspicious	Less conformity: $r = -.42$
Adair (1972) ^b	38 of 86 (44%)	Less conformity: $eta = .21$
Stang (1976) ^b	13 of 65 (20%)	Less conformity: $eta = .3$; "significant" treatment effects on conformity only when suspicious [participants] were removed from the analyses" (p. 353)

^aEffect sizes calculated (eta, biserial correlation r) when sufficient information was available. ^bObtained from our literature search (search words *deception* and *suspicion* and its variants); articles with no index stem from Stang's (1976) review.

& Geller, 1973; Willis & Willis, 1970) suspicion did not significantly change the amount of conformity behavior, and no study reported that suspicion produced greater conformity. For 9 of the 10 studies in which suspicion triggered less conformity and in which the necessary information was given we calculated an effect size measure (*eta*, or *r*). *Eta* is defined as the square root of the proportion of variance accounted for (Rosenthal & Rosnow, 1991) and is identical to the Pearson product-moment correlation coefficient when $df = 1$, as is the case when two conditions are compared (as in most cases where we calculated *eta*). According to Cohen's (1988) classification of effect sizes, a value of *eta* of .1, .3, or .5 constitutes a small, medium, or large effect size, respectively. In terms of these measures, the reduction in conformity as a function of suspicion was of medium to large effect size.

To conclude, in research on conformity behavior the data of participants who are suspicious of deception and those of naïve participants are different. Those who suspected being tricked were less likely to bend to social pressure than those who trusted the experimental scenario. If one assumes that not all participants reveal their suspicions truthfully (see Altemeyer, 1971; Taylor & Shepperd, 1996), then the true differences between the groups may be even larger. In research on conformity behavior suspicion appears to increase the probability beta of wrongly rejecting the alternative hypothesis (Type II error) rather than increasing the probability alpha of wrongly rejecting the null hypothesis (Type I error; for an example of Type I error due to deception, see the weapon effect below). Taking the risk of increasing the probability of beta (Type II) error is not a negligible threat in a discipline in which the power of experimental tests ($1 - \beta$) in major psychology journals continues to be as low as 50% (assuming a medium-size effect; Cohen, 1992; Gigerenzer, Krauss, & Vitouch, 2004; Sedlmeier & Gigerenzer, 1989).

Test 2: Are the Effects of Suspicion Negligible: Studies That Experimentally Induced Suspicion

To circumvent the problem of participants not admitting to being suspicious, experimenters can systematically "plant" participants' suspicion from the outset and then study their experimental performance as a function of it. To find such studies, we used the search term *deception* in combination with *prebriefing*, or *forewarning*. We found a total of eight studies. The issue with which we are concerned here, namely, the effect of experimentally induced suspicion, was not the explicit focus in all eight studies. Participants' knowledge and thus suspicion of deception ranged from relatively neutral forewarning about experimental procedures in general (e.g., D. F. Allen, 1983, p. 901: "In a few experiments it is necessary for experimenters to deceive subjects concerning some elements of the experiment") to concrete tip-offs by a confederate (e.g., Levy, 1967), to disclosure that deception would occur during the experiment (e.g., Finney, 1987).

Table 3 summarizes how participants' foreknowledge of deception affected behavior. The results are mixed, with some studies finding no effect and others large

TABLE 3
The Effects of Anticipation of Deception on Experimental Performance

<i>Authors</i>	<i>Research Topic</i>	<i>Manipulation</i>	<i>Behavioral Effects (Effect Size^d)</i>
Levy (1967)	Verbal conditioning	Two groups of participants: fully informed (tipped off by a confederate) and uninformed.	Groups differed in the level of performance ($\eta^2 = .41$) but there were no significant differences in the shape of the acquisition curve.
Golding and Lichtenstein (1970)	Valins effect (effect of bogus heart rate feedback on preferences)	Three groups of participants: naïve, suspicious (by being told in a conversation with a confederate that they would be tricked), and completely informed about the deception by a confederate.	No "significant" differences in the Valins effect as a function of prior knowledge. However, participants who admitted awareness of experimental manipulation in a postexperimental questionnaire did not show the Valins effect, whereas those who either were not aware or did not admit their awareness showed a substantial effect ($r = -.48$).
Gallo, Smith, and Mumford (1973)	Conformity behavior	Three groups of participants: complete, partial, or no information about the purpose of the experiment (the information did not reveal that deception was used).	No significant effect ($\eta^2 = .13$).
Turner and Simons (1974)	Aggression (weapons effect)	Three groups of participants: no information, informed that some deception might be involved (by a confederate tip-off), or informed that "the weapons were probably part of the procedure to influence their behavior" (p. 342).	"Increased levels of ... subject sophistication led to decreased numbers of shocks administered by subjects to their frustrators" (p. 341; $\eta^2 = .43$).

Spinner, Adair, and Barnes (1977)	Incidental learning	At the end of the first part of an experiment, designed to arouse suspicion, participants were told: "Sometimes experiments require that a subject be deceived initially" (p. 546). Based on an "awareness questionnaire" administered at the end of the second part of the experiment, participants were classified into three groups as a function of their suspicion and anticipation of other tasks. Two groups of participants: "neutral forewarning" (i.e., "in a few experiments it is necessary for experimenters to deceive subjects concerning some elements of the experiment") vs. no forewarning. Three groups of participants either were instructed that they "may be deceived," or "will be deceived," or did not receive any consent information.	Those who were suspicious and intended to prepare for some other task scored higher than those who did not prepare and/or were not suspicious ($\eta^2 = .46$).
D. F. Allen (1983)	Cooperativeness in a Prisoner's Dilemma game		No significant effect.
Finney (1987)	Conformity behavior		The number of conformity judgments in the "will be deceived" group (4.1) was significantly higher than in the "no consent" group (1.9), but the results in the latter group did not differ from those in the "may be deceived" group (2.3).
Wiener and Erker (1986)	Attribution of responsibility and evaluation of culpability	Two groups: standard informed consent group and prebriefing group (i.e., participants were alerted to the possibility that they might be intentionally misinformed).	No significant effects for sentencing judgment, verdicts, and attribution judgments; significant differences in the attribution process.

^aEffect sizes calculated when sufficient information was available (search words: *deception* and *prebriefings*, *forewarning*, or *informed consent*).

effects. Nevertheless, a trend is discernable. The more concrete the foreknowledge, the more it affects participants' behavior: When participants received detailed tip-offs about the true purpose of the experiment (e.g., Levy, 1967; Turner & Simons, 1974), or were explicitly told that they would be deceived (Finney, 1987), or explicitly acknowledged awareness of experimental manipulation (Golding & Lichtenstein, 1970), suspicion altered experimental performance (albeit not necessarily on all dependent measures). In contrast, when participants were merely informed that some kind of deception might happen (D. F. Allen, 1983; Finney, 1987; Wiener & Erker, 1986) or were told the purpose of the study (without indicating the possibility of deception; Gallo, Smith, & Mumford, 1973), their performance did not differ from that of control participants who were not given this information (but see Spinner, Adair, & Barnes, 1977).

Test 3: Does Previous Experience of Deception Evoke Suspicion and Are Its Effects Negligible?

Yet another way to explore the effects of suspicion is to study how participants' experimental history affects experimental performance. To find studies that adopted this approach, we used the search term *deception* in combination with *experimental history* and found nine studies. Table 4 summarizes a complex series of findings. In brief, the results suggest that firsthand experience with deception or manipulation affects performance, whereas mere disclosure of the possibility of deception in psychological experiments does not (Cook & Perrin, 1971; see also Christensen, 1977, Experiments 1 and 2). Second, Silverman, Shulman, and Wiesenthal (1970) observed that the experience with deception appears to make people more apprehensive of evaluation. Third, the studies by Fillenbaum (1966) and Fillenbaum and Frey (1970) suggest that not all suspicious participants act on their suspicion. Fourth, different dependent variables seem to be differentially affected by the experience with deception. In Cook and Perrin's research, incidental-learning data differed as a function of experimental history, but attitude data did not (but see Experiment 2 in Cook et al., 1970). Finally, the extent to which previous deception experience transfers to other experiments may depend on the similarity between the past and present experimental situation (Brock & Becker, 1966; Cook et al., 1970).

Page and Scheidt (1971) reported a dramatic example involving the "weapons effect," which illustrates how past experience with laboratory deception can distort behavior so extremely that it elicits a phenomenon that "cannot be generalized to nonlaboratory situations" (p. 304). The "weapons effect" (originally reported by Berkowitz & LePage, 1967) suggests that weapons might stimulate aggression by classical conditioning processes resulting from learned associations between aggressive acts and weapons. Page and Scheidt were able to replicate the weapons effect in only one of three of their experiments, and only in a group of participants

TABLE 4
The Effects of Experimental History on Participants' Performance

<i>Authors</i>	<i>Research Topic</i>	<i>Manipulation</i>	<i>Behavioral Effects (Effect Size^d)</i>
Brock and Becker (1966)	Compliance behavior	Students participated in two consecutive experiments, the debriefed experiment and the three groups: no debriefing, partial debriefing, and complete debriefing. For half the participants, the test experiment included an element from the debriefing experience; for the other half the common element was omitted.	Complete debriefing reduced compliance behavior in the test experiment (10%) but only when the debriefing situation and the test experiment were explicitly similar; no reduction in the no and partial debriefing conditions (50% and 50%).
Fillenbaum (1966)	Incidental learning	<i>Experiment 1.</i> Performance on an incidental-learning task after an earlier task that did or did not involve deception. <i>Experiment 2.</i> Same procedure as in Experiment 1 with minor changes.	<i>Experiment 1.</i> Although participants who experienced deception did somewhat better on the incidental-learning task, the difference was "not very large and far from significant" (p. 534, $r = .1$). Difference was larger if one compared participants who reported themselves to be suspicious to those who did not. <i>Experiment 2.</i> Participants with deception experience did better on the incidental-learning task ($r = .27$). As in Experiment 1, difference was larger if one compared participants who reported themselves to be suspicious to those who did not.

(continued)

TABLE 4 (Continued)

Authors	Research Topic	Manipulation	Behavioral Effects (Effect Size ^a)
Cook et al. (1970)	Attitude-change experiments	<p><i>Experiment 1.</i> Experimentally naïve participants took part in one of five attitude-change experiments.</p> <p><i>Experiment 2.</i> Participants were assigned to one of three groups in Experiment 1, which was or was not linked to Experiment 2 (by a common cue). The three groups were no deception, experience of deception, knowledge of deception.</p>	<p><i>Experiment 1.</i> Attitude data did not significantly differ as a function of experimental history. Experimental history, however, affected global attitudes: Participants with deception experiences believed the experimenter less, considered experiments to be less scientific and less valuable, and reported caring less about understanding and following instructions.</p> <p><i>Experiment 2.</i> Attitude was affected by the deception variable and the presence of the cues ($\eta^2 = .34$). Without a cue, experience of deception biased the data (compared to knowledge of deception). With a cue, learning about deception but not experiencing it biased the data.^b</p>
Fillenbaum and Frey (1970)	Incidental learning	Students were given the critical incidental-learning task immediately after a prior and revealed deception on another task. Students were categorized as “trustful” or “suspicious” participants.	Suspicious participants scored higher on the incidental-learning task than trustful participants ($\eta^2 = .31$).
Silverman, Shulman, and Wiesenthal (1970)	Various dependent variables	<p>Experiment 1 involved either deception and debriefing or a memory study without deception. In Experiment 2, all participants were given tests measuring compliance of demands, persuasibility, sentence completion, and a personality test.</p>	<p>“Significant differences between deception and nondeception conditions were observed with all four of the tests used” (p. 209). η^2 equaled .25, .26, and .29 for the compliance of demands, persuasibility, and sentence completion test, respectively. Overall, “the deception experience sensitized subjects to possible ulterior purposes of experiments, increasing evaluation apprehension” (p. 209).</p>

Cook and Perrin (1971)	Attitude change, incidental learning	<p><i>Experiment 1.</i> Participants were assigned to one of three deception conditions: no deception, experience of deception, and knowledge of deception.</p> <p><i>Experiment 2.</i> Attitude-change and incidental-learning measures were obtained (participants did or did not learn that this experiment also involved subsequent deception; we ignore this manipulation here).</p>	<p>The attitude data (unlike in Cook et al., 1970) did not discriminate between conditions. The incidental-learning measure showed that prior experience but not prior learning of deception produced greater incidental learning ($r = .3$), and "experiencing deception produced the strongest evidence of absolute bias" (p. 215). A measure of general suspiciousness ("how truthful are psychology experimenters") but not of particular suspiciousness (concerning the relationship of both experiments) showed a main effect on incidental learning ($\epsilon/\eta = .29$).</p> <p>The weapon effect was obtained for the sophisticated but not for the naive participants ($\epsilon/\eta = .32$). "What appeared to be aggressive behavior to the original investigators seems to have been a sham or an artifact" (p. 315).</p> <p><i>Experiment 1.</i> Conditioning did not occur for the group that experienced prior manipulation and deception, but it did occur for the group that was only told that experiments may involve active manipulation of their behavior.</p> <p><i>Experiment 2.</i> Unlike in the control and nonmanipulation groups, "subjects given a manipulative experimental experience do not exhibit verbal conditioning" (p. 397).</p> <p>Participants who had been debriefed improved more in the parallel form than those who had not been debriefed ($\epsilon/\eta = .3$).</p>
Page and Scheidt (1971)	Aggressiveness (weapons effect)	<p><i>Experiment 3.</i> Two groups of participants: naive participants who took part in a psychological experiment for the first time and sophisticated participants who took part in a deception experiment within the last month.</p> <p><i>Experiment 1.</i> Four experimental groups, including one group in which an active attempt was made to manipulate their behavior. Then they were debriefed and went through the verbal conditioning procedure.</p>	<p>Participants who took part in a psychological experiment for the first time and sophisticated participants who took part in a deception experiment within the last month.</p> <p><i>Experiment 1.</i> Four experimental groups, including one group in which an active attempt was made to manipulate their behavior. Then they were debriefed and went through the verbal conditioning procedure.</p>
Christensen (1977)	Verbal conditioning	<p><i>Experiment 2.</i> Three experimental groups, including one prior manipulation group and one nonmanipulation group.</p>	<p><i>Experiment 2.</i> Three experimental groups, including one prior manipulation group and one nonmanipulation group.</p>
Grueter, Stumpfhauser, and Weyer (1977)	Performance on an intelligence test	<p>Participants received randomly determined feedback about their performance on an intelligence test.</p> <p>Half of them were debriefed about this deception whereas the other half were not. Then they worked on a parallel form of the test a week later.</p>	<p>Participants who had been debriefed improved more in the parallel form than those who had not been debriefed ($\epsilon/\eta = .3$).</p>

^aEffect sizes calculated when sufficient information was available. ^bBias defined as the difference to the no-deception group.

who had taken part in a deception experiment within the previous month; participants unfamiliar with psychological experimentation did not exhibit the effect. Turner and Simons (1974; see also Simons & Turner, 1976) challenged Page and Scheidt's results, and Turner, Simons, Berkowitz, and Frodi (1977) even suggested, "Perhaps the failures to replicate the weapons effect occurred because the researchers used subjects who were not naïve about deception studies or who were very apprehensive about the impression they might create" (p. 369). Although Page and Scheidt and Turner et al. disagreed over the issue of how experience with deception alters experimental performance, they agreed that it does have this potential. Turner and Simons concluded, "Apparently, unless subjects are naïve, the effects of important independent variables may be obscured" (p. 347).

Again: Are the Effects of Suspicion Negligible?

According to a key argument made in the defense of the use of deception, the differences in the data of naïve and suspicious participants are negligible. We tested this conjecture against three sets of empirical studies that systematically explored the effects of suspicion on behavioral data. It seems fair to conclude that Tests 1 through 3 show that the effects of suspicion are not invariably strong. But the analysis also demonstrates that the consequences of suspicion can be substantial. In the set of studies that constituted Test 1, we found that in two thirds of the conformity studies in question, researchers reported evidence—of medium to strong effect size (Table 2)—that naïve and suspicious participants' behavior differs markedly. In the set of studies that constituted Test 2, we found that the concrete but not general foreknowledge of deception (e.g., being forewarned or prebriefed) appears to systematically alter experimental performance (Table 3).

In the set of studies that constituted Test 3, the effects of previous firsthand experience with deception point to the boundary conditions of the effects of suspicion: First, dependent variables appear to differ in the extent to which they provide room for biasing effects of suspicion to occur (e.g., attitude vs. incidental learning). Second, the extent to which future experiments may elicit suspicion may depend on the similarity (or lack thereof) between the previous experiments in which participants experienced deception and the future ones. Third, even if participants are suspicious, not everybody might act on their suspicion. Fourth, the study by Silverman et al. (1970) suggests that the effects of suspicion can also be indirect by affecting participants' motivations. Such motivations include predilections to enact the good participant role, the obedient participant role, the evaluation-apprehensive role, and the negative participant role (see Rosenthal & Rosnow, 1991, chap. 6), respectively. Suspicion could amplify some of these motivations (such as evaluation-apprehension) while crowding out others. Although it may be difficult to discover these indirect effects of suspicion, their consequences need not be negligible as the case study of research on the weapons effect shows.

GENERAL DISCUSSION

We evaluated empirically three arguments that are often advanced to justify and defend the use of deception. Consistent with the first argument, there has been a drop in the use of deception in social psychology (but not in marketing research) in comparison with the heyday of deception in the 1960s and 1970s. This decline, however, has not turned deception into an endangered species. Marketing researchers and social psychologists used it routinely in the 1990s and at the beginning of the new millennium. In some social psychology and marketing journals the deception rates (in laboratory studies) are as high as 50% and higher. In *JPSP*, every third study published in the 1990s employed deception. Second, we found mixed evidence regarding the thesis that deceived participants do not become resentful about having been fooled by researchers. Defenders and critics of deception can point to studies consistent with their point of view. Third, in contradiction to the argument that the effects of suspicion are negligible, we found evidence that suspicion has the potential to adversely impact research outcomes, both in the experiment at hand and in subsequent studies. Undoubtedly, the available empirical evidence does not allow us to finally settle the methodological debate on deception, and there is room for honest differences in evaluating the ultimate impact of deception. In what follows, we discuss what we consider the key implications of the findings.

OLD EVIDENCE AND PRIVATE OBSERVATIONS

In our search for studies that examined the methodological consequences of deception, we discovered that most available studies date back to the decade between the mid-1960s and the mid-1970s. This is no coincidence. Silverman (1978, p. 405) referred to this period as the “most self-critical decade” of psychology, during which much research was devoted to investigating the “threats to validity that reside in ... the interaction between the experimenter and the subject” (Rosenthal & Rosnow, 1991, p. 110). Are the results of this research obsolete today? For several reasons, we do not think so. For one, those who defend the use of deception in psychology (e.g., Kimmel, 1998), or make the case for its use in experimental economics (Bonetti, 1998a, 1998b) typically justify their arguments with reference to this evidence.

Second, the few available recent studies also highlight the potential of firsthand experience to affect behavior in future experiments. Epley and Huff (1998) and Krupat and Garonzik (1994) asked participants to report what their concrete expectations would be if they participated in future research (e.g., “You will be misled or deceived in some way during the course of the study”) and analyzed these expectations as a function of prior experience with deception. Participants’ re-

sponses suggested that with previous exposure to deception, participants were more likely to expect to be misled and deceived in future experiments and to be more suspicious of information presented by the experimenter. Epley and Huff's and Krupat and Garonzik's findings are also consistent with still another category of contemporary evidence that only rarely makes it to the public domain: researchers' unprompted, informal observations.

One example of such an unprompted observation is MacCoun and Kerr's (1987) report described earlier. Is theirs just a rare exception or the tip of an iceberg? We do not know. We are, however, surprised by how many of our colleagues have related unprompted observations to us, ranging from comments on participants' distrust about the promised performance-contingent payment to their distrust of crucial parameters in gambles to their conviction that some coincidental noise outside of the laboratory room is systematically related to the current experiment. These informal observations suggest that there are myriad ways in which suspicion can seep into our labs and studies. To avoid painting a completely lopsided picture, however, we should also emphasize that a number of colleagues have related to us instances of participants being solicitous about the experiment or the experimenter and far removed from seeing the experimenter, in the words of one commentator, as "a liar or an ogre."

HOW PSYCHOLOGISTS MAY CURTAIL NEGATIVE CONSEQUENCES OF SUSPICION

In his essay, Kelman (1967) predicted that as we continue to use deception "our potential subjects become increasingly distrustful of us," and therefore the "more research we do, the more difficult and questionable it becomes" (p. 7). It seems fair to say that this did not happen. Why? One explanation is that Kelman did not anticipate what is possibly an institutional solution to the problem of a contaminated participant pool—psychology's strategy of constantly replenishing the pool, thus reducing the risk of relying on suspicious participants. In response to the possible side effects of deception, Silverman et al. (1970) recommended that "that the practice of using the same subjects repeatedly be curtailed, and whenever administratively possible, subjects who have been deceived and debriefed be excluded from further participation" (p. 211).

Indeed, it seems psychologists have taken this advice to heart. More than in the past undergraduates have become a major source of experimental research data, and typically, undergraduate participation is enforced through the use of a participant pool. Participant pools are replenished by requiring undergraduate students—notably students from introductory classes—to participate in research projects as part of their course requirements. In a survey of 242 U.S. psychology departments (with participant pools and graduate programs), Sieber and Saks (1989) found that

93.4% of departments recruited from introductory courses. This does not mean that 93% of their participants are from introductory courses, as 35% of the responding departments also recruit from other lower division courses (p. 1057).

Participant pools, however, have not always relied so heavily on students from introductory classes. In his analysis of the participant selection in studies published in the period 1966–1967 in the two largest journals of the APA, Schultz (1969) found that 41% (*Journal of Experimental Psychology*) and 34% (*Journal of Personality and Social Psychology*) of studies relied on students from introductory psychology courses as participants. What has prompted this drastic change in the composition of psychology's participant pools? One reader of Schultz's article suggested a myriad of reasons including reliance on participant pools in which students are obligated to participate in experiments in exchange for experimental credit guarantees participant availability, makes recruitment less effortful, and reduces costs by minimizing the need for monetary compensation. Although this is speculation, we suggest that another contributory factor to the current practice of recruiting participants mostly from introductory courses has been the need for minimizing the contaminating effects of deception and suspicion on the participant pool. By replenishing pools with ever-new and naïve participants and using them as the prime source of data, a recruiting mechanism has evolved in psychology that promises to curtail the possibly distorting side effects of firsthand and secondhand experience with deception.

Psychologists also appear to take individual precautions to curtail the negative consequences of participants' suspicion. For instance, a prominent social psychologist told us that at his laboratory, in which deception is used and in which students are eligible to participate in multiple experiments, experimenters routinely probe for suspicion at the end of the studies. In addition, they ask the participants to list the previous studies in which they have participated. If experimenters need naïve participants, they can discount all data from participants who have previously participated in a deception study. Or they might choose to analyze those data separately and estimate the effects of experience with prior deceptions. To the best of our knowledge, this practice is not institutionalized throughout psychological laboratories but is left to the discretion of the individual researcher. Therefore, such arrangements may have the unfortunate, and paradoxical, consequence that researchers who do not use deception are more likely to become victim of its potentially distorting side effects, because they are likely to be less inclined to probe their participants for suspicion and thus be less able to control for the effect of prior experience with deception.

DOES THE RULE OF CONDUCT RULE OUR CONDUCT?

Deception has been defended as an indispensable strategy of last resort for the study of those facets of behavior that are of great social importance and for which

alternative research methods either are unavailable or would produce invalid data (e.g., Bröder, 1998; Kimmel, 1998; Korn, 1998; Weiss, 2001). By this argument, the costs of *not* conducting such research (e.g., on conformity, obedience, racial stereotypes, bystander effect, and aggression) outweigh the costs of using deception (e.g., Trice, 1986). This argument is also explicitly endorsed by the APA rules of conduct. According to those rules, “psychologists do not conduct a study involving deception unless they have determined that the use of deceptive techniques is justified by the study’s prospective scientific, educational, or applied value and that effective nondeceptive alternative procedures are not feasible” (APA, 2002, Standard 8.07).

Does this rule of conduct rule our conduct? One way of answering the question of whether deception is used as a last-resort method is by simply looking at the numbers. Despite APA’s exhortation to treat deception as a method of last resort deception is frequently used in journals such as *JPSP*, *JESP*, *JMR*, and *JCR*. In our own analysis of published articles in *JESP* in 2002, we found that about every second study employed some method of deception, and Kimmel (2001) reported similarly high numbers for *JMR* and *JCR*. Can a method so frequently drawn on (in these journals) count as the rarely used tool that one expects a method of last resort method to be, one that is to be employed only in those cases in which the study’s prospective utility justified the use of deception and in which equally effective alternative procedures were not feasible? Rather than looking at the numbers, one can also browse through recent issues of leading journals in social and experimental psychology and look for topics studied both in psychology and economics, such as strategic games, negotiations, choices between gambles, and so on. In psychology, participants are routinely misled to believe that their decisions in games and gambles will determine their final payoffs, that assignment of roles in an experiment will be determined by chance, that they will be paired up with another person, that the feedback they receive will be veridical, that information they provide will be made public, and so on. Experimental economic studies that address similar and sometimes the exact same questions—and that do so without deception—raise serious doubts about whether the false claims in the corresponding studies in psychology were indeed indispensable.

Why are the APA rules of conduct not more effective in enforcing deception as a strategy of last resort? One problem is, perhaps, the halfhearted way we teach the rules. In the late 1960s, Kelman worried that psychologists use deception without question, and he felt that “we are training a generation of students who do not know that there is any other way of doing experiments in our field—who feel that deception is as much de rigueur as significance at the .05 level” (1967, p. 3). Since then, some things have changed. Today, we certainly do not teach students that deception is de rigueur. Yet we doubt that they typically learn—for instance, by way of published studies—that deception is meant to be a strategy of last resort. Rather, implicitly or explicitly we signal to students that deception is a commonly

accepted practice that needs, however, to be justified to what is often perceived as a capricious and overly cautious ethical review by institutional review boards (IRBs).

Another pragmatic problem with the APA rule is that that the decision of whether deception is justified by its anticipated utility is left to those who stand to benefit from its use (Ortmann & Hertwig, 1997, 1998). Notwithstanding the mediating role of IRBs (which tend to focus on the ethical rather than the methodological consequences of deception), this practice leaves the assessment of private benefits (e.g., deception is often less expensive and more convenient than alternative procedures, thus promising quicker publications; see Adelson, 1969) and public costs (e.g., possible contamination of the participant pool) to the interested party (the experimenter)—a classic moral-hazard problem whose solution is likely to sacrifice the public interest.

How can one enforce the APA rules without necessarily expanding the somewhat daunting role of IRBs, which doubtlessly have complicated the business of experimentation? According to Pittenger (2002), the APA should provide more specific standards regarding the permissibility of deception and its appropriate use. Beyond clearer standards, we believe that the most promising solution to this dilemma is to implement mechanisms in which the individual researcher has an incentive to forgo the routine tool deception and to search for and implement alternative procedures. Such incentives could, for instance, be put in place via the editorial process. Specifically, the APA could lobby the editors of the relevant journals to change their editorial practices and to impose more transparent and stricter rules on the submission of studies incorporating deception. If editors required researchers to justify their use of deception and defined stringent and clear criteria for the justification of deception, then researchers would be spurred on to consider more thoroughly the use of nondeceptive experimental methods and to develop alternatives.

One key criterion against which editors can judge the use of decision is past research practices. Specifically, authors could be asked to briefly characterize previous research practices regarding the topic they investigate (either in a cover sheet or in the body of the manuscript) *and* explain why they cannot adopt the nondeceptive designs that others have used before. Would not such a requirement perpetuate the use of deception, simply because in many areas few nondeceptive studies have ever been conducted? No. Even in research traditions in which deception has been considered to be indispensable, alternative research techniques have often been available. This follows logically from Gross and Fleming's (1982) review of 1,188 journal articles in leading social psychology journals (between 1959 and 1979). The authors analyzed the prevalence of deception in 24 research areas—including conformity, altruism, impression formation, and attitude change—in social psychology and observed a wide variation in how often deception was used in different areas. Researchers in about half of the areas used deception in less than

half of all studies. That is, in areas in which deception has often been advocated as indispensable—facets of human behavior that are of great social importance—studies devoid of deception have been conducted. Even in research areas such as conformity alternatives are available. Stricker, Messick, and Jackson (1969), for instance, observed that 20% of all conformity studies published in 1964 in four leading social psychology journals did not use deception.

ETHICAL IMPLICATIONS OF METHODOLOGICAL ISSUES

Although deception is still frequently used, ethical guidelines for research have become substantially stricter (for a short history of the “ten commandments of the APA,” see Rosnow & Rosenthal, 1997). As a consequence, the profession has succeeded in reducing the severity of deceptive methods used. Rosnow and Rosenthal, for instance, concluded that many of the seminal studies that were conducted then and that raised daunting ethical issues (e.g., Baumrind, 1964, 1971, 1985; see also Aguinis & Handelsman, 1997; Herrera, 1996; Kimmel & Smith, 2001) “would be impossible today” (p. 114).

Rather than on ethical implications of deception, our review focused on possible methodological implications. They are, however, not divorced. To appreciate this, let us turn to the rationale behind economists’ ban of deception. Davis and Holt (1993) in their authoritative textbook *Experimental Economics* described the rationale as follows:

The researcher should ... be careful to avoid deceiving participants. Most economists are very concerned about developing and maintaining a reputation among the student population for honesty in order to ensure that subject actions are motivated by the induced monetary rewards rather than by psychological reactions to suspected manipulation. Subjects may suspect deception if it is present. Moreover, even if subjects fail to detect deception within a session, it may jeopardize future experiments if the subjects ever find out that they were deceived and report this information to their friends. (pp. 23–24)

In the parlance of economists, participants’ expectation that they will not be deceived (i.e., honesty on the part of the experimenter) is a common good of sorts (such as clean water or the arctic wildlife) that would be depleted quickly even if only a few experimenters practiced deception. On theoretical and empirical grounds, economists also do not trust experimenters to make an unbiased analysis of the (private) benefits of deception and its (public) costs. The “moral hazard” of reaping the private benefits of deception (e.g., in terms of a more convenient experimental design or swifter publication) is perceived to be simply too great. In other

words, even though conservation of the common good (trustworthiness brought about by honesty) may be in all experimenters' joint interest, any given individual experimenter has an incentive to take a free ride off the contributions of others. If everyone follows such private incentives, any good will be produced at an inefficient, low level, or not at all.

Whether or not one agrees with this view, it highlights that the ethical problems of deception are not restricted to the experimenter–participant dyad. On economists' views, other experimenters who do without deception end up paying the public potential costs (e.g., participants' reactions to suspected deception) of others' use of deception, thus violating an implicit "social contract" between experimenters.

CONCLUSION

Undoubtedly, the empirical evidence is not as clear-cut as one might hope—and as either the proponents or the critics of the use of deception sometimes imply. Rather than ending with the clichéd call for more empirical work, let us spell out three conclusions that we draw. First, we believe that there is room for honest differences in interpreting the evidence reviewed here. Second, in light of the still wobbly empirical basis, one may decide to reserve deception for clearly specified circumstances (thus containing a potential methodological damage whose likelihood is unclear). Third, although this is exactly the self-imposed policy of the APA, psychology's rules of conduct and its research reality are two different animals. Unless we mean those rules to be diluted, the discipline ought to address the gap between them and reality (see also Hertwig & Ortmann, 2008).

ACKNOWLEDGMENTS

We thank Valerie M. Chase, William D. Crano, Gerd Gigerenzer, Anita Todd, Peter M. Todd, and David Weiss for many constructive comments, and Laura Wiles for improving the readability of our article.

REFERENCES

- Adair, J. G. (1972). Demand characteristics or conformity? Suspiciousness of deception and experimenter bias in conformity research. *Canadian Journal of Behavioral Science*, *4*, 238–248.
- Adair, J. G., Dushenko, T. W., & Lindsay, R. C. L. (1985). Ethical regulations and their impact on research practice. *American Psychologist*, *40*, 59–72.
- Adelson, J. (1969). Personality. *Annual Review of Psychology*, *20*, 217–252.

- Aguinis, H., & Handelsman, M. M. (1997). Ethical issues in the use of the bogus pipeline. *Journal of Applied Social Psychology, 27*, 557–573.
- Aitkenhead, M., & Dordoy, J. (1985). What the subjects have to say. *British Journal of Social Psychology, 24*, 293–305.
- Allen, D. F. (1983). Follow-up analysis of use of forewarning and deception in psychological experiments. *Psychological Reports, 52*, 899–906.
- Allen, V. L. (1966). Effect of knowledge of deception on conformity. *Journal of Social Psychology, 69*, 101–106.
- Altemeyer, R. A. (1971). Subject pool pollution and the postexperimental interview. *Journal of Experimental Research in Personality, 5*, 79–84.
- American Psychological Association. (2002). *Ethical principles of psychologists and code of conduct*. Washington, DC: Author. Retrieved from <http://www.apa.org/ethics/>
- Asch, S. E. (1956). Studies of independence and conformity: I. A minority of one against a unanimous majority. *Psychological Monographs, 70*(No. 416).
- Baumrind, D. (1964). Some thoughts on ethics of research. After reading Milgram's "Behavioral study of obedience." *American Psychologist, 19*, 421–423.
- Baumrind, D. (1971). Principles of ethical conduct in the treatment of subjects: Reaction to the draft report of the Committee on Ethical Standards in Psychological Research. *American Psychologist, 26*, 887–896.
- Baumrind, D. (1979). IRBs and social science research: The costs of deception. *IRB: A Review of Human Subjects Research, 1*, 1–4.
- Baumrind, D. (1985). Research using intentional deception: Ethical issues revisited. *American Psychologist, 40*, 165–174.
- Berkowitz, L., & LePage, A. (1967). Weapons as aggression-eliciting stimuli. *Journal of Personality and Social Psychology, 7*, 202–207.
- Boles, T. L., Croson, R. T. A., & Murnighan, J. K. (2000). Deception and retribution in repeated ultimatum bargaining. *Organizational Behavior and Human Decision Processes, 83*, 235–259.
- Bonetti, S. (1998a). Experimental economics and deception. *Journal of Economic Psychology, 19*, 377–395.
- Bonetti, S. (1998b). Reply to Hey and Starmer & Mc Daniel. *Journal of Economic Psychology, 19*, 411–414.
- Brandts, J., & Charness, G. (2003). Truth or consequences: An experiment. *Management Science, 49*, 116–130.
- Brock, T. C., & Becker, L. A. (1966). "Debriefing" and susceptibility to subsequent experimental manipulations. *Journal of Experimental Social Psychology, 2*, 314–323.
- Bröder, A. (1998). Deception can be acceptable. *American Psychologist, 53*, 805–806.
- Carlson, R. (1971). Where is the person in personality research? *Psychological Bulletin, 75*, 203–219.
- Chipman, A. (1966). Conformity as a differential function of social pressure and judgment difficulty. *Journal of Social Psychology, 70*, 299–311.
- Christensen, L. (1977). The negative subject: Myth, reality, or a prior experimental experience effect? *Journal of Personality and Social Psychology, 35*, 392–400.
- Christensen, L. (1988). Deception in psychological research: When is its use justified? *Personality and Social Psychology Bulletin, 14*, 664–675.
- Clark, R. D., & Word, L. E. (1974). Where is the apathetic bystander? Situational characteristics of the emergency. *Journal of Personality and Social Psychology, 29*, 279–287.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.
- Collins, F. L., Jr., Kuhn, I. F., Jr., & King, G. D. (1979). Variables affecting subjects' ethical ratings of proposed experiments. *Psychological Reports, 44*, 155–164.

- Cook, T. D., Bean, J. R., Calder, B. J., Frey, R., Krovetz, M. L., & Reisman, S. R. (1970). Demand characteristics and three conceptions of the frequently deceived subject. *Journal of Personality and Social Psychology, 14*, 185–194.
- Cook, T. D., & Perrin, B. F. (1971). The effects of suspiciousness of deception and the perceived legitimacy of deception on task performance in an attitude change experiment. *Journal of Personality, 39*, 204–224.
- Cooper, J. (1976). Deception and role-playing: On telling the good guys from the bad guys. *American Psychologist, 31*, 605–610.
- Croson, R., Boles, T., & Murnighan, J. K. (2003). Cheap talk in bargaining experiments: lying and threats in ultimatum games. *Journal of Economic Behavior and Organization, 51*, 143–159.
- Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies. Diffusion of responsibility. *Journal of Personality and Social Psychology, 8*, 377–383.
- Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton, NJ: Princeton University Press.
- Endler, N. S., & Hartley, S. (1973). Relative competence, reinforcement and conformity. *European Journal of Social Psychology, 3*, 63–72.
- Endler, N. S., Wiesenenthal, D. L., & Geller, S. H. (1972). The generalization of the effects of agreement and correctness on relative competence mediating conformity. *Canadian Journal of Behavioral Science, 4*, 322–329.
- Epley, N., & Huff, C. (1998). Suspicion, affective response, and educational benefit as a result of deception in psychology research. *Personality and Social Psychology Bulletin, 24*, 759–768.
- Epstein, Y. M., Suedfeld, P., & Silverstein, S. J. (1973). The experimental contract: Subjects' expectations of and reactions to some behaviors of experimenters. *American Psychologist, 28*, 212–221.
- Erikson, K. (1995). Commentary. *American Sociologist, 26*, 4–11.
- Ettinger, R. F., Marino, C. J., Endler, N. S., Geller, S. H., & Natziuk, T. (1971). The effects of agreement and correctness on relative competence and conformity. *Journal of Personality and Social Psychology, 19*, 204–212.
- Farr, J. L., & Seaver, W. B. (1975). Stress and discomfort in psychological research: Subject perception of experimental procedures. *American Psychologist, 30*, 770–773.
- Fillenbaum, S. (1966). Prior deception and subsequent experimental performance: The “faithful” subject. *Journal of Personality and Social Psychology, 4*, 532–537.
- Fillenbaum, S., & Frey, R. (1970). More on the “faithful” behavior of suspicious subjects. *Journal of Personality, 38*, 43–51.
- Finney, P. D. (1987). When consent information refers to risk and deception: Implications for social research. *Journal of Social Behavior and Personality, 2*, 37–48.
- Fisher, C. B., & Fyrberg, D. (1994). Participant partners: College students weigh the costs and benefits of deceptive research. *American Psychologist, 49*, 417–427.
- Gallo, P. S., Jr., Smith, S., & Mumford, S. (1973). Effects of deceiving subjects upon experimental results. *Journal of Social Psychology, 89*, 99–107.
- Geller, S. H., & Endler, N. S. (1973). The effects of subject roles, demand characteristics, and suspicion on conformity. *Canadian Journal of Behavioral Science, 5*, 46–54.
- Geller, S. H., Endler, N. S., & Wiesenenthal, D. L. (1973). Conformity as a function of task generalization and relative competence. *European Journal of Social Psychology, 3*, 53–62.
- Gerdes, E. P. (1979). College students' reactions to social psychological experiments involving deception. *Journal of Social Psychology, 107*, 99–110.
- Gigerenzer, G., Krauss, S., & Vitouch, O. (2004). The null ritual: What you always wanted to know about significance testing but were afraid to ask. In D. Kaplan (Ed.), *The Sage handbook of quantitative methodology for the social sciences* (pp. 391–408). Thousand Oaks, CA: Sage.
- Glinski, R. J., Glinski, B. C., & Slatin, G. T. (1970). Nonnaivety contamination in conformity experiments: Sources, effects, and implications for control. *Journal of Personality and Social Psychology, 16*, 478–485.

- Golding, S. L., & Lichtenstein, E. (1970). Confession of awareness and prior knowledge of deception as a function of interview set and approval motivation. *Journal of Personality and Social Psychology*, *14*, 213–223.
- Gross, A. E., & Fleming, I. (1982). Twenty years of deception in social psychology. *Personality and Social Psychology Bulletin*, *8*, 402–408.
- Gruder, C. L., Stumpfhauser, & Wyer, R. S., Jr. (1977). Improvement in experimental performance as a result of debriefing about deception. *Personality and Social Psychology Bulletin*, *3*, 434–437.
- Herrera, C. D. (1996). An ethical argument against leaving psychologists to their statistical devices. *Journal of Psychology*, *130*, 125–130.
- Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: A methodological challenge for psychologists? *Behavioral and Brain Sciences*, *24*, 383–451.
- Hertwig, R., & Ortmann, A. (2002). Economists' and psychologists' experimental practices: How they differ, why they differ and how they could converge. In I. Brocas & J. D. Carillo (Eds.), *The psychology of economic decisions* (pp. 253–272). New York: Oxford University Press.
- Hertwig, R., & Ortmann, A. (2008). Deception in psychological experiments: Two misconceptions and a research agenda. *Social Psychology Quarterly*.
- Hey, J. D. (1991). *Experiments in economics*. Oxford, England: Basil Blackwell.
- Hey, J. D. (1998). Experimental economics and deception. A comment. *Journal of Economic Psychology*, *19*, 397–401.
- Kelman, H. C. (1967). Human use of human subjects: The problem of deception in social psychological experiments. *Psychological Bulletin*, *67*, 1–11.
- Kimmel, A. J. (1996). *Ethical issues in behavioral research: A survey*. Cambridge, MA: Blackwell.
- Kimmel, A. J. (1998). In defense of deception. *American Psychologist*, *53*, 803–805.
- Kimmel, A. J. (2001). Ethical trends in marketing and psychological research. *Ethics and Behavior*, *11*, 131–149.
- Kimmel, A. J., & Smith, N. C. (2001). Deception in marketing research: Ethical, methodological, and disciplinary implications. *Psychology and Marketing*, *18*, 663–689.
- Korn, J. H. (1987). Judgments of acceptability of deception in psychological research. *Journal of General Psychology*, *114*, 205–216.
- Korn, J. H. (1998). The reality of deception. *American Psychologist*, *53*, 805.
- Krupat, E. (1977). A re-assessment of role playing as a technique in social psychology. *Personality & Social Psychology Bulletin*, *3*, 498–504.
- Krupat, E., & Garonzik, R. (1994). Subjects' expectations and the search for alternatives to deception in social psychology. *British Journal of Social Psychology*, *33*, 211–222.
- Lawson, E. (2001). Informational and relational meanings of deception: Implications for deception methods in research. *Ethics and Behavior*, *11*, 115–130.
- Ledyard, J. O. (1995). Public goods: A survey of experimental research. In J. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics* (pp. 111–194). Princeton, NJ: Princeton University Press.
- Levenson, H., Gray, M. J., & Ingram, A. (1976). Research methods in personality five years after Carlson's survey. *Personality and Social Psychology Bulletin*, *2*, 158–161.
- Levy, L. (1967). Awareness, learning and the beneficent subject as expert witness. *Journal of Personality and Social Psychology*, *6*, 363–370.
- Lewicki, R. J., & Stark, N. (1996). What is ethically appropriate in negotiations: An empirical examination of bargaining tactics. *Social Justice Research*, *9*, 69–95.
- MacCoun, R. J., & Kerr, N. L. (1987). Suspicion in the psychological laboratory: Kelman's prophecy revisited. *American Psychologist*, *42*, 199.
- McDaniel, T., & Starmer, C. (1998). Experimental economics and deception: A comment. *Journal of Economic Psychology*, *19*, 403–409.
- McNamara, J. R., & Woods, K. M. (1977). Ethical considerations in psychological research: A comparative review. *Behavior Therapy*, *8*, 703–708.

- Menges, R. J. (1973). Openness and honesty v. coercion and deception in psychological research. *American Psychologist*, 28, 1030–1034.
- Merriam-Webster's Collegiate Dictionary* (11th ed.). Springfield, MA: Merriam-Webster.
- Milgram, S. (1964). Issues in the study of obedience: A reply to Baumrind. *American Psychologist*, 19, 848–852.
- Nicks, S. D., Korn, J. H., & Mainieri, T. (1997). The rise and fall of deception in social psychology and personality research, 1921 to 1994. *Ethics and Behavior*, 7, 69–77.
- Oliansky, A. (1991). A confederate's perspective on deception. *Ethics and Behavior*, 1, 253–258.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17, 776–783.
- Ortman, A., & Hertwig, R. (1997). Is deception acceptable? *American Psychologist*, 52, 746–747.
- Ortman, A., & Hertwig, R. (1998). The question remains: Is deception acceptable? *American Psychologist*, 53, 806–807.
- Ortman, A., & Hertwig, R. (2002). The empirical costs of deception: Evidence from psychology. *Experimental Economics*, 5, 111–131.
- Page, M. M., & Scheidt, R. H. (1971). The elusive weapons effect: Demand awareness, evaluation apprehension, and slightly sophisticated subjects. *Journal of Personality and Social Psychology*, 20, 304–318.
- Pihl, R. O., Zaccchia, C., & Zeichner, A. (1981). Follow-up analysis of the use of deception and aversive contingencies in psychological experiments. *Psychological Reports*, 48, 927–930.
- Pittenger, D. J. (2002). Deception in research: Distinctions and solutions from the perspective of utilitarianism. *Ethics and Behavior*, 12, 117–142.
- Riecken, H. W. (1962). A program for research on experiments in social psychology. In N. F. Washburne (Ed.), *Decisions, values, and groups* (Vol. 2) (pp. 25–41). New York: Pergamon.
- Ring, K. (1967). Experimental social psychology: Some sober questions about some frivolous values. *Journal of Experimental Social Psychology*, 3, 113–123.
- Ring, K., Wallston, K., & Corey, M. (1970). Mode of debriefing as a factor affecting subjective reaction to a Milgram-type obedience experiment: An ethical inquiry. *Representative Research in Social Psychology*, 1, 67–88.
- Rosenthal, R., & Rosnow, R. L. (1991). *Essentials of behavioral research: Methods and data analysis* (2nd ed.). Boston: McGraw-Hill.
- Rosnow, R. L., & Rosenthal, R. (1997). *People studying people: Artifacts and ethics in behavioral research*. New York: Freeman.
- Rubin, Z., & Moore, J. C., Jr. (1971). Assessment of subjects' suspicions. *Journal of Personality and Social Psychology*, 17, 163–170.
- Schultz, D. P. (1969). The human subject in psychological research. *Psychological Bulletin*, 72, 214–228.
- Sedlmeier, P., & Gigerenzer, G. (1989). Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin*, 105, 309–316.
- Seeman, J. (1969). Deception in psychological research. *American Psychologist*, 24, 1025–1028.
- Sharpe, D., Adair, J. G., & Roese, N. J. (1992). Twenty years of deception research: A decline in subjects' trust? *Personality and Social Psychology Bulletin*, 18, 585–590.
- Sieber, J. E., Iannuzzo, R., & Rodriguez, B. (1995). Deception methods in psychology: Have they changed in 23 years? *Ethics and Behavior*, 5, 67–85.
- Sieber, J. E., & Saks, M. J. (1989). A census of subject pool characteristics and policies. *American Psychologist*, 44, 1053–1061.
- Silverman, I. (1978). Expectancy effects revisited. *Behavioral and Brain Sciences*, 1, 404–405.
- Silverman, I., Shulman, A. D., & Wiesenthal, D. L. (1970). Effects of deceiving and debriefing psychological subjects on performance in later experiments. *Journal of Personality and Social Psychology*, 14, 203–212.

- Simons, L. S., & Turner, C. W. (1976). Evaluation apprehension, hypothesis awareness, and the weapons effect. *Aggressive Behavior*, 2, 77–87.
- Smith, C. P. (1981). How (un)acceptable is research involving deception? *IRB: A Review of Human Subjects Research*, 3, 1–4.
- Smith, C. P., & Berard, S. P. (1982). Why are human subjects less concerned about ethically problematic research than human subject committees? *Journal of Applied Social Psychology*, 12, 209–222.
- Smith, N. C., Klein, J. G., & Kimmel, A. J. (2002). *The ethics of deception in consumer research*. Unpublished manuscript, London Business School, London.
- Smith, S. S., & Richardson, D. (1983). Amelioration of deception and harm in psychological research: The important role of debriefing. *Journal of Personality and Social Psychology*, 44, 1075–1082.
- Spinner, B., Adair, J. G., & Barnes, G. E. (1977). A reexamination of the faithful subject role. *Journal of Experimental Social Psychology*, 13, 543–551.
- Stang, D. J. (1976). Ineffective deception in conformity research: Some causes and consequences. *European Journal of Social Psychology*, 6, 353–367.
- Straits, B. C., Wuebben, P. L., & Majka, T. J. (1972). Influences on subjects' perceptions of experimental research situations. *Sociometry*, 35, 499–518.
- Stricker, L. (1967). The true deceiver. *Psychological Bulletin*, 68, 13–20.
- Stricker, L. J., Messick, S., & Jackson, D. N. (1967). Suspicion of deception: Implication for conformity research. *Journal of Personality and Social Psychology*, 5, 379–389.
- Stricker, L. J., Messick, S., & Jackson, D. N. (1969). Evaluating deception in psychological research. *Psychological Bulletin*, 71, 343–351.
- Sullivan, D. S., & Deiker, T. A. (1973). Subject-experimenter perceptions of ethical issues in human research. *American Psychologist*, 28, 587–591.
- Taylor, K. M., & Shepperd, J. A. (1996). Probing suspicion among participants in deception research. *American Psychologist*, 51, 886–887.
- Toy, D., Olsen, J., & Wright, L. (1989). Effects of debriefing in marketing research involving "mild" deceptions. *Psychology and Marketing*, 6, 69–85.
- Trice, A. D. (1986). Ethical variables? *American Psychologist*, 41, 482–483.
- Turner, C. W., & Simons, L. S. (1974). Effects of subject sophistication and evaluation apprehension on aggressive responses to weapons. *Journal of Personality and Social Psychology*, 30, 341–348.
- Turner, C. W., Simons, L. S., Berkowitz, L., & Frodi, A. (1977). The stimulating and inhibiting effects of weapons on aggressive behavior. *Aggressive Behavior*, 3, 355–378.
- Vitelli, R. (1988). The crisis issue assessed: An empirical analysis. *Basic and Applied Social Psychology*, 9, 301–309.
- Weber, S. J., & Cook, T. D. (1972). Subjects effects in laboratory research: An examination of subject roles, demand characteristics, and valid inferences. *Psychological Bulletin*, 77, 273–295.
- Weiss, D. J. (2001). Deception by researchers is necessary and not necessarily evil. *Behavioral and Brain Sciences*, 24, 431–432.
- Wiener, R. L., & Erker, P. V. (1986). The effects of prebriefing misinformed research participants on their attributions of responsibility. *Journal of Psychology*, 120, 397–410.
- Wiesenthal, D. L., Endler, N. S., & Geller, S. H. (1973). Effects of prior group agreement and task correctness on relative competence mediating conformity. *European Journal of Social Psychology*, 3, 193–203.
- Willis, R., & Willis, Y. (1970). Role playing versus deception: An experimental comparison. *Journal of Personality and Social Psychology*, 16, 472–477.