

Whether to Use Deception in Experiments? Proposal of a Cost-Benefit Approach

Abstract

This article argues that in some research contexts, deception can undermine the validity of research results. A case-by-case approach should be favoured over a universal resolution of whether deception is indeed more useful than hurtful for the scientific validity of the experiment. This article proposes a cost-benefit approach rather than a systematic use of deception in marketing experiments. We argue that the prevalence of demand artefacts is relatively low, and its consequences, tacitly overvalued. We expose the ethical and methodological weaknesses of deception at a micro and macro level and propose less costly methods for control of demand artefacts.

Key words: Deception; Cost-benefit approach; Experiments; Methodology; Demand artefacts

“Seldom, very seldom, does complete truth belong to any human disclosure; seldom can it happen that something is not a little disguised or a little mistaken.”

—Jane Austen, Emma

> Damien Hallegatte

Associate professor of marketing, Member of LaboNFC
Université du Québec à Chicoutimi, Chicoutimi, Canada
Email : Damien_Hallegatte@uqac.ca

> Myriam Ertz

Assistant professor of marketing, Head of LaboNFC
Université du Québec à Chicoutimi, Chicoutimi, Canada
Email: Myriam_Ertz@uqac.ca

1. Introduction

The editorial of *Journal of Consumer Research* entitled “Death to dichotomizing” (Fitzsimons, 2008) reveals the importance of rigour in marketing research, which could be compromised by dichotomising of continuous independent variables in order to test their combined effect with manipulated variables in experiments. The very title of this editorial includes an unusually strong normative injunction, which has had the expected outcome, namely a significant reduction in this popularly accepted practice (Pham, 2015). However, the idea that this analytical procedure has disadvantages such that it should be systematically excluded, has recently been questioned (Iacobucci, Posavac, Kardes, Schneider & Popovich, 2015b). Beyond the question of merits and demerits of the dichotomisation and competing methods, this controversy highlights two different stances relating to the methodology of marketing research. The first is to maximise the methodological rigour, which implies, for example, establishing a prescriptive requirement that applies to all cases of analysis of the effect of continuous independent variables in experimentations, the use of regression analysis (McClelland, Lynch, Irwin, Spiller, & Fitzsimons, 2015; Rucker, McShane, & Preacher, 2015). The second approach could be described as cost-benefit, where the best procedure is applied according to the case (Iacobucci, Posavac, Kardes, Schneider & Popovich, 2015a, Iacobucci et al., 2015b), thus is less drastic than the outright rejection of a method which has indisputable merits, knowing that the alternative methods presented as superior have their own weaknesses.

We propose to apply a critical view, similar to that of Iacobucci et al. (2015a, 2015b), on the systematic use of deception. This critical approach to deception is also a contribution to the scientific nature of our marketing discipline, and could even be of relevance to the fields of psychology, sociology, anthropology, or economics. However, unlike Iacobucci et al. (2015a, 2015b), it is not to rehabilitate a procedure that has recently been shunned from the discipline, but to propose an alternative to current systematic use of deception in experimentation; specifically, a cost-benefit approach to decide whether or not to use deception in experimental research. Although the cost-benefit approach has the potential of improving validity of experimental methodology, it never seemed to find favour with most researchers, at least in the recent history. To this end, we summarize all the arguments in favour of a cost-benefit approach and highlight the weakness of the arguments in favour of the systematic application of these two procedures. We demonstrate that the use of deception in experiments is not always necessary, and that the disadvantages of this procedure may surpass its advantages. In particular, we maintain that the prevalence of demand artefacts is relatively low and that its consequences are tacitly overvalued. Moreover, we show the weaknesses and disadvantages of deception, from an ethical and methodological point of view, at a micro and macro level, and propose less costly methods for controlling of demand artefacts.

2. Controlling Demand Artefacts

Demand artefacts are an important issue in terms of the validity of experiments, since they are likely to influence the results (Orne, 1962). The subject of the experiment may assume a role according to what he believes to be the hypotheses of the research and thus provide biased answers (Orne, 1969). At least three types of roles possibly adopted by the subject are likely to introduce bias (see Sawyer, 1975). If he adopts the “positive role”, the subject is motivated to confirm what he believes to be the researcher’s assumptions. If he adopts the “negative role”, the subject will try to disprove them, behaving in a contradictory, random or neutral manner. Finally, if he is an apprehensive subject, his primary motivation is to look good in the eyes of the researcher. In each of these three cases, the subject does not adopt the role of “integral subject”, and this poses a threat both to the internal and external validity of the results (Sawyer, 1975). Thus, since the seminal article by Sawyer (1975), demand artefacts is a major concern for researchers in marketing.

However, according to Shimp, Hyatt and Snyder (1991), the prevalence and consequences of demand artefacts could be tacitly overvalued by researchers. Indeed, it is generally accepted that, on the one hand, guessing a hypothesis induces demand artefacts and, on the other hand, that demand artefacts constitute a source of systematic error (rather than random) (Shimp, Hyatt & Snyder, 1991). These authors propose three conditions that must be met in order for a demand artefact to occur: the subject must encode a demand index, then that he discerns the hypothesis, and finally that he acts according to this hypothesis. These conditions are related to the receptivity and motivation of the Rosenthal and Rosnow model (see also Rosnow & Aiken, 1973), both of which are considered mediating variables of demand artefacts (Allen, 2004). To our knowledge, if the implications of these conditions have been criticised (see Darley & Lim, 1993), the conditions themselves have not been questioned.

Thus, the probability that the results of an experiment are biased due to demand artefacts depends on the least of probabilities of the three necessary but not sufficient conditions individually stated by Shimp, Hyatt and Snyder (1991). In other words, it is not because a subject doubts that the researcher has a specific objective that he will guess a hypothesis, and it is not because the subject has guessed a hypothesis that he will amend his behaviour accordingly. Even if several subjects had altered their behaviour, this would more likely result in a random rather than a systematic error (Shimp et al., 1991). Indeed, for such an error to be systematic, it would be necessary that a significant number of subjects guess

the same hypothesis and consequently adopt the same behaviour to influence the results in the same direction (Shimp et al., 1991). It is therefore reasonable to assume that such an error, if it exists, has a good chance of being random and thus does not constitute a bias in the results, but only a decrease in the magnitude of the effects, which appears much less serious. Then, guessing a hypothesis does not necessarily imply a demand artefact, and a demand artefact does not necessarily imply a systematic error. The prevalence of demand artefacts would be less important, and the potential consequences less damaging, than what most researchers think.

Notwithstanding the prevalence and consequences of demand artefacts, the precautionary principle would require that procedures be implemented to ensure that the problem is controlled, albeit minimally. To do this, a method used in research is to perform a post-experimental survey, which eliminates those who apparently discovered experimental hypotheses (Shimp et al., 1991). Such an approach is problematic for several reasons. First, as discussed previously, discerning a hypothesis does not necessarily imply a demand artefact, since the subject must modify his behaviour accordingly. This procedure therefore leads to the elimination of subjects whose answers were valid. Secondly, measuring the discernment of the hypotheses by the subjects is delicate and possibly invalid, in particular because such a procedure aimed at eliminating the demand artefacts is itself likely to be affected (Gorn, Jacobs & Mana, 1987). Finally, these post-experimental procedures can introduce a bias, which is systematic (Shimp et al., 1991). Indeed, subjects who correctly guessed the hypotheses could be systematically different from others in terms of intelligence or need for cognition, for example (Shimp et al., 1991). Moreover, such a procedure leads to the elimination of subjects in the experimental groups and not in the control groups, as the case may be.

Thus, the post-experimental survey may threaten both the external and internal validity of the experiment (Shimp et al., 1991). It is probably for this reason that such procedures are rarely used in some disciplines such as marketing. In a content analysis of experimental research published in 259 articles in five journals of marketing, namely the *Journal of Marketing*, the *Journal of Marketing Research*, the *Journal of Consumer Research*, the *Journal of the Academy of Marketing Science*, and the *International Journal of Research in Marketing*.

Of course, such a point of view, minimising the occurrence of the demand artefact issue, the magnitude of its consequences and invalidating the most used method of coping with it, has been criticised. Darley and Lim (1993) argue that the prevalence is higher and the consequences more important than Shimp, Hyatt and Snyder (1991) suggest. On the one hand, the prevalence of discernment of hypotheses is higher than what the empirical data indicate, as it is likely that some subjects discern the hypotheses without indicating it. In other words, the subjects are able to guess the purpose of the research and adjust their behavior in any way that befits them, accordingly. On the other hand, the consequences of such discernment may be more damaging than suggested by Shimp, Hyatt and Snyder (1991), since the adoption of a role by a subject may be unconscious (Darley & Lim, 1993). Darley and Lim (1993) therefore suggest seeking to systematically control demand artefacts in experiments. With respect to the design of the experiment, they recommend verifying the realism of the experiment and the involvement of the participants in a pilot study and in the main study. In addition, they promote deception to minimise the risk of discernment of hypotheses by the participants. With respect to the validation of the experiment, Darley and Lim (1993) suggest that non-experimentation and hetero method (see Sawyer, 1975) be used systematically and jointly. The first consists of a replication of the experimental procedure, but without treatment, with only the description of the treatments. The second consists of replication with other ways of operationalising the independent variables as well as variations in procedures.

The recommendations of Darley and Lim (1993) deserve several remarks. Regarding the design of the experimentation, they appear reasonable. In particular, as far as the use of deception is concerned, it has been a common practice since at least the early 1980s (Kimmel, 2011, Perrien, 1997). As far as the validation of experimentation is concerned, the systematic and joint use of non-experimentation and hetero method appears as an "extreme" recommendation (Shimp, Hyatt & Snyder, 1993). They involve a complete replication for each experimentation and therefore, multiplication of the number of subjects accordingly. Moreover, among the 259 articles analysed by Perrien (1997), only one reported on the use of non-experimentation and none on the use of the hetero method. But it is even more striking to note that the post-experimental survey is recommended by Darley and Lim (1993)—those advocates of methodological orthodoxy—only for the pilot studies, that is to say, on the plan of the design and not of the validation of the experimentation. The fact of advocating a post-experimental procedure before the main experimentation is, if not an oxymoron, a clear recognition of these weaknesses, even a disavowal. Moreover, Darley and Lim (1993) do not attempt to argue against the criticisms of Shimp, Hyatt, and Snyder (1991) on post-experimental surveys, and thus tacitly endorse them. In conclusion, however, they indicate that it would be more harmful not to delete the subjects likely to have guessed the hypotheses than to retain them. Nevertheless, this is only a matter of faith, as it is in complete contradiction with their own suggestion to confine the post-experimental survey to pilot studies. Table 1 summarises the cases in which the demand artefact risks are more or less high.

3. Deception, a common but not always justified practice

If post-experimental surveys are not a common practice to reduce demand artefacts (Khan, 2011; Perrien, 1997), deception is. If a thorough discussion of this point is not necessary (see Hertwig & Ortmann, 2008a for an excellent discussion), there are some key aspects worth discussing. In the Perrien (1997) survey, 42% of articles explicitly used deception in their experimental design, while only 29.4% referred to demand artefacts. Thus, more articles have used deception than dealing specifically with the issue of demand artefacts. For Perrien (1997), “the explanation is obvious: deception becomes part of the experimental process, although it should depend on a high probability of responses with a demand bias”. Other studies in marketing, which were focusing on two journals, namely the Journal of Consumer Research and the Journal of Marketing Research, but were taking into account all empirical articles and not only experimentations, show a progression of 43% in 1975–1976 to 56% in 1996–1997 (Kimmel, 2001), then 66% in 2001–2002 (Kimmel, 2004), and finally 77% in 2006–2007 (N. C. Smith, Kimmel & Klein, 2009)2 of articles that used some form of deception. Conversely in social psychology, the use of deception has decreased. In the Journal of Personality and Social Psychology, the use of deception reached a peak of 73% of the articles published in 1978, subsequently declining before stabilising between one third and two fifths in the 1990s (Hertwig & Ortmann, 2008a). Therefore, we can conclude that deception is not a systematic practice in social psychology, whereas it is almost the case in marketing, and especially top marketing research outlets. The reason for this may relate to the influx of many scholars from the field of psychology in general – not only social psychology, in particular - for whom deception-based treatments of subjects remains an elementary form of methodology. Yet, it is clear that deception is a debatable practice not only from an ethical point of view, but also from a methodological point of view. From an ethical perspective, deception can have negative effects on participants, ranging from discomfort to loss of self-esteem (for a review of literature, see Smith et al., 2009). The codes of ethics of the American Psychological Association and the American Sociological Association advocate a cost-benefit approach, whereas the American Anthropological Association apparently prohibits deception (Smith et al., 2009). In addition, it is explicitly prohibited in experimental economics (see Jamison, Karlan, & Schechter, 2008). Apart from ethical considerations, deception is criticised from a methodological point of view. Indeed, the effective use of deception depends on four elements: a high level of naivety on the subjects’ part; an experimental procedure which does not provide clues to the subjects that deception is used; a valid verification of the effectiveness of deception; a suspicion of deception on the part of the subjects which do not modify the experimental effects (Golding & Lichtenstein, 1970). Thus, the use of deception to eliminate demand artefacts may increase them if the latter is not effective (Sawyer, 1975; Smith et al., 2009). Specifically, rather than reduce demand artefacts, deception may, if detected during experimentation by participants, alter their behaviour and thus introduce bias (Jamison et al., 2008).

The methodological problems of deception largely exceed the framework of the validity of an individual study. Using deception could increase suspicion among potential participants and thus decrease the number of naïve participants (Kimmel, 2011). As a matter of fact, a participant who has been cheated once can expect to be cheated again the next time. More generally, deception can tarnish the image of the discipline (Kimmel, 2011). Non-suspicious participants constitute a “public good” which must be protected (Jamison et al., 2008). Certainly, since Milgram’s famous experience on obedience in the 1960s, it seems that deception has not had the deleterious effects apprehended. However, this does not detract from the possibility that such negative effects will occur in the future, particularly because of information technologies that allow rapid dissemination of information on ethically questionable research (Smith et al., 2009). In addition, commonly used post-experimental debriefing can mitigate the short-term negative consequences on participants, but probably has the effect of increasing suspicion, since it reveals deception.

Table 1

Evaluating the Incidence of Deception

Incidence of Deception	Lowest	Highest
Criteria		
Level of naivety on the part of participants.	Low	High
Risk that the experimental procedure provides clues to the participants that deception is used.	High	Low
Risk that suspicion of deception on the part of the participants modifies the experimental effects.	High	Low
Validity of the proposed procedure for the effectiveness of deception.	Low	High

Source: adapted from Golding & Lichtenstein (1970).

4. A Cost-Benefit Approach

No advocate of the maximum methodological rigour (e.g., Darley & Lim, 1993) would deny the disadvantages of control methods of demand artefacts, either on the design of the experiment or on its validation. Similarly, no advocate of methodological simplification (e.g., Shimp, Hyatt & Snyder, 1991, 1993) would deny that demand artefacts can compromise the validity of the results of an experiment or other empirical research.

The use of deception should therefore always be subject to a cost-benefit type of calculation (Pascual-Leone, Singh & Scoboria, 2010), as advocated by the codes of ethics of the American Psychological Association and the American Sociological Association (Smith et al., 2009). Such an approach has also been advanced in the field of economics on the ground that “there are important research questions for which deception is truly unavoidable” (Hertwig & Ortmann, 2008b, p. 225). For Shimp, Hyatt and Snyder (1993), the costs of rigorous efforts, such as proposed by Darley and Lim (1993), greatly exceed the benefits, and this, for the discipline in its entirety. The study of Perrien (1997)—indicating that, on 259 articles, only one has mentioned using non-experimentation, and none of the hetero method, although prior to the article of Darley and Lim (1993)—seems supportive of Shimp, Hyatt and Snyder (1993). In fact, deception should be used only if the probability of demand artefacts is high, and not be a systematic procedure (Perrien, 1997), that is to say that it should be used as a last resort (Kimmel, 2011; Smith et al., 2009) and researchers should consider the use of a procedure without deception (Sawyer, 1975).

The challenge for experimental researchers is therefore to assess the risk of demand artefacts and choose one or several proportionate measures to monitor them. Indeed, “a general awareness and a commitment are more important than any specific methodology” (Sawyer, 1975). It entails assessing the risks of occurrence, which are relatively low in experimental designs of the type “between subjects” less likely to be affected by demand artefacts than designs of the type “within subjects” (Sawyer, 1975). The risk is even lower when the participants are not from a sample of students, which are a source of demand artefacts, since they are more likely to adopt a positive role, namely to confirm the assumed hypotheses (Orne, 1962). It is also interesting to note that, if only 7.7 per cent of marketing studies analysed by Perrien (1997) have used a “within subject” design, 74.3% have used samples of students. Finally, the dependent variables relating to behaviour are less likely to be affected by demand artefacts than other types of variables (Sawyer, 1975). Thus, the type of design, the nature of the sample and the type of dependent variable attenuate the risk of demand artefacts.

With regard to the choice of methods to control the possible demand artefacts, researchers may proceed by elimination. On the validation side of the experiment, non-experimentation and hetero method are very expensive methods, almost absent from research in marketing, for example, as seen previously. With regard to post-experimental surveys, it is not the implementation cost, but the associated risks that act against it in the majority of marketing cases, as seen earlier. On the design of experimentation side, deception is a method commonly used in marketing, but less frequently in social psychology. One of the reasons is that marketing has seen a steady rise in the consumer behavior domain which studies individuals rather than organizations and strategy, while on the other hand social psychology seeks to understand cognition and behavior as they naturally occur in a social context. As such, deception has been challenged on methodological grounds but also on ethical ones, and other research strategies have been favoured instead, such as role-playing, for example (Klein, 2014). Deception is thus not only ethically objectionable, but methodologically delicate, as discussed previously. Deception is all the less justified from an ethical point of view since the methodological gains are questionable (Smith et al., 2009). For example, in the case of recruiting participants on specialised discussion forums, where they are asked questions on a topic of particular interest to them, the risks of suspicion appear high and the consequences thereof are potentially very damaging. Indeed, some participants will seek to know the true motivations of the researcher. As they are on a discussion forum which, by definition, is a place promoting discussions, they will not only be able to discuss their assumptions but, in addition, possibly deter other participants from answering, or even encourage them to sabotage. Moreover, in such a case a possible debriefing procedure appears difficult to implement, since the data collection takes place over time. A discussion forum therefore appears as a particularly dangerous context for the use of deception. If, at the beginning of the introduction of deception in marketing, it was already clear that the latter was a source of problems (Sawyer, 1975), the context of recruitment of participants on discussion forums existing today exacerbates them. Thus, in the case of experimentations taking place on the Internet, the costs (or rather risks) of such a procedure for the control of demand artefacts can all the more exceed its profits. Table 2 summarises the different criteria for determining the effectiveness of deception in a given situation according to a cost-benefit approach.

Table 2

Assessment of the Risks of Demand Artefacts

Risks of demand artefact	Lower	Higher
Criteria		
Type of experimental design	Between subjects	Within subjects
Type of sample	Non-student	Student
Type of Dependent Variable	Dependent behavioural Variables	Dependent cognitive variables

Sources: Sawyer, 1975, Orne, 1962.

5. Alternative of the Scenario Method

Waiving deception may imply that respondents are instead presented openly hypothetical scenarios. In marketing, such procedures are called scenario method or role-playing method, since the latter implies that the respondents imagine being in a situation in which they are not in reality. Bolton, Warlop and Alba (2003) even use the term "paradigm of scenarios". Traditionally used in the literature on service encounters (e.g., Bitner, 1990; A. K. Smith, Bolton & Wagner, 1999; Surprenant & Solomon, 1987), the scenario method is currently used in several other areas, such as consumer loyalty (Wagner, Hennig-Thurau, & Rudolph, 2009), relationship marketing (Melancon, Noble & Noble, 2011), price management (Drèze & Nunes, 2004), distribution (Rungtusanatham, Wallin, & Eckerd, 2011), and retailing (Mukhopadhyay & Johar, 2009). Although generally associated with quantitative approaches, the scenario method is also used in qualitative studies (e.g., Grønhoj & Bech-Larsen, 2010). This method has also been used in social psychology for several decades (see, e.g., Freedman, 1969; Kwon & Weingart, 2004; Robinson & Close, 2001; Stolte, 1994), in which it is also called the "vignette research" (Stolte, 1994) or "vignette-based methodology" (Robinson & Close, 2001), where "vignette" is a written scenario. In the literature using the scenario method, the two main concerns are to ensure, on the one hand, an acceptable level of realism in the scenarios, and on the other hand, a certain involvement of respondents. Indeed, although at least two studies, one in marketing and the other in psychology, have shown that the scenario method may yield similar results to a real situation (Robinson & Close, 2001; Surprenant & Churchill Jr, 1984), verifying the realism of the scenarios and the involvement of respondents allows for reducing the risks of demand artefacts (Darley & Lim, 1993). In order to simultaneously increase realism and the involvement of the respondents, but also their ability to respond, Smith, Bolton and Wagner (1999) have asked them to appoint a service recently used and have then presented a scenario where they had to imagine going back to the hotel or the restaurant itself and experiment a hypothetical service encounter. In the same vein, Mukhopadhyay and Johar (2009) have imposed a purchase decision on respondents, and then asked them to give reasons for a decision that they had not really taken. The reasons per se had no interest to researchers in marketing, but were to be used to increase the involvement and the realism of a decision that the respondents had not taken. Another possible approach, but a costly one, is media production of video recordings of service encounters (e.g., Sparks, Bradley, & Callan, 1997). The realism of the scenarios and the involvement of the respondents can also be increased by compatibility between a context and a sample. For example, Dabholkar (1996) chose post-secondary students for a study on self-service technology.

The realism of the scenarios can also be validated by groups of discussion, which play the role of pilot study at this point. The aim is to establish different realistic scenarios in the eyes of the participants and discussion groups and, particularly, that none of them gives rise to massive rejection. It is a verification of realism initially more thorough than what is usually performed in pilot studies in marketing, as suggested by Darley and Lim (1993). As to the realism of the task itself, either to evaluate hypothetical scenarios, we can ensure a proximity between the task requested and what the respondents do in real life, not only by the choice of our sample, like Smith, Bolton and Wagner (1999) and Dabholkar (1996), but also by the choice of data collection method, in this case by discussion forums. As a matter of fact, a discussion forum, by definition, consists of people who discussed a particular topic of concern to them. Developing hypothetical scenarios is probably, for the majority of the members of these forums, a normal and pleasant activity. If the scenarios are realistic, and the task of assessing a scenario appears as a "natural" activity for the members of the forum in question, waiving deception appears as a methodologically reasonable choice.

Furthermore, regarding the involvement of the respondents, the latter is fostered if there is compatibility between the context of the research and the study population, like Smith, Bolton and Wagner (1999) and Dabholkar (1996). If there is a good compatibility between the subject of the study and the study population, the use of the scenario method is less prone to bias, according to the criteria provided by Sawyer (1977). First, the subjects play their own role as a fervent supporter

of a particular trademark, and not, for example, the role of a car salesman. Secondly, there is little chance that the subject contains socially undesirable elements. Thirdly, the context is familiar to respondents, where they have “a basis for their projection”. Fourthly, the scenarios have good chances to appear as simple to the eyes of these connoisseurs of the subject. In addition, the context in which individuals react to the scenario has an effect on the quality of the answers. Responding to scenarios in a quiet and isolated environment minimises the respondents’ tendency to simply seek satisfaction, the latter involving that respondents treat the information in a way less effective than in the reality (Stolte, 1994). However, it is reasonable to think that the majority of members of the discussion forums, who are of interest to the majority of researchers in marketing, frequent these environments in moments of rest or relaxation.

In addition to the methodological precautions discussed above to ensure the realism of the scenarios and the involvement of the respondents, it is possible to take other measures to minimise the risk that the respondents shall adopt an inadequate role. Among the roles that may be adopted by participants is that of the apprehensive subject, for whom projecting a good self-image will be a motivation likely to bias his answers. To minimise this apprehension, Sawyer (1975) suggests emphasising the status of the researcher and to guarantee anonymity. However, attempting to cheat the respondents on this point, for example by presenting the person responsible for the study as a student while he is a professor must be avoided, since a simple query on a common search engine would reveal the true status. A member of the discussion forum used for a study could conclude that the identity of the researcher is a lie, share this information with other members, and thus seriously jeopardise the data collection process. In addition, if it is necessary to ensure anonymity, Sawyer (1975) cautions on the fact of not encouraging too low an involvement by insisting on anonymity. Table 3 summarises the criteria of relevance of the scenario method as an alternative to deception.

Table 3

Evaluation of the Relevance of the Scenario Method

Relevance of the scenario method	Highest	Lowest
Criteria		
Role played by the participants	Own role (e.g.: fervent supporter of a trademark in particular)	Fictional role (e.g.: seller)
Type of topic addressed	No elements socially undesirable	Contains elements socially undesirable
Context of the study	Familiar to the participants	Little familiar to the participants
Perception of the scenarios by the participants	Simple	Complex

Source: adapted from Sawyer, 1977.

6. Conclusions

This article emphasised the idea that the systematic use of deception in experimentations is unjustified. When almost systematically using this procedure when feasible, this paper proposes the systematic recourse to a cost-benefit analysis, in arguing that the prevalence of the demand artefacts is relatively low, and that its consequences are tacitly overvalued. We have highlighted the weaknesses and disadvantages of deception, methodological procedure commonly used both from the ethical and methodological point of view, at the micro and macro level, and proposed less costly methods of control of demand artefacts. Hence, the force of these arguments in favour of a cost-benefit approach makes it plausible that the systematic application of these two procedures in scientific literature is less a question of strength of scientific arguments than political victory of the proponents of utmost methodological rigour, facilitated by a general willingness to promote the scientific nature of a particular discipline. Although the good faith of these researchers is beyond doubt, the rigour seems to have a disproportionate importance when addressing the relevance and clarity in the research. If strengthening the scientific status of a discipline is an important objective, the methodological sophistication has costs, including the decline in the relative importance of the subject studied and of the substantive contributions of the research (Lehmann, McAlister, & Staelin, 2011). In research, rigour must not become an end in itself and methodological elegance must have precedence on the socio-political relevance of the research, so that our research continues to have an impact on society, our ultimate objective.

Conflicts of interest: none

Endnotes

1. This figure corresponds to articles that have “considered demand artifacts”, except for those that merely indicated it as a limitation in the research. Thus, this figure may include any a priori or a posteriori procedure. Hence, this evaluation is considered to be a very moderate assessment of the occurrence of post-experimental surveys.
2. These authors used the same methodology as Kimmel (2001, 2004) so that results are comparable.

References

- Allen, C. T. (2004). A Theory-Based Approach for Improving Demand Artefact Assessment in Advertising Experiments. *Journal of Advertising*, 33(2), 63–73.
- Armstrong, J. S. (2003). Discovery and communication of important marketing findings: Evidence and proposals. *Journal of Business Research* 56(1), 69–84.
- Bitner, M. J. (1990). Evaluating Service Encounters: The Effects of Physical Surroundings and Employee Responses. *Journal of Marketing*, 54(2), 69–82.
- Bolton, L. E., Warlop, L., & Alba, J. W. (2003). Consumer Perceptions of Price (Un) Fairness. *Journal of Consumer Research*, 29(4), 474–491.
- Dabholkar, P. A. (1996). Consumer evaluations of new technology-based self-service options: An investigation of alternative models of service quality. *International Journal of Research in Marketing*, 13(1), 29–51.
- Darley, W. K., & Lim, J. S. (1993). Assessing Demand Artefacts in Consumer Research: An Alternative Perspective. *Journal of Consumer Research*, 20(3), 489–495.
- Drèze X., & Nunes, J. C. (2004). Using Combined-Currency Prices to Lower Consumers' Perceived Cost. *Journal of Marketing Research*, 41(1), 59–72.
- Festinger, L. (1953). Laboratory experiments. In L. Festinger, & D. Katz (Eds.) *Research methods in the behavioral sciences* (pp. 136-172). New York: Holt, Rinehart and Winston.
- Fitzsimons, G. J. (2008). Death to Dichotomizing. *Journal of Consumer Research*, 35(1), 9–11.
- Freedman, J. L. (1969). Role playing: Psychology by consensus. *Journal of Personality and Social Psychology*, 13(2), 107–114. doi: 10.1037/h0028067
- 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(3), 213–223. doi: 10.1037/h0028853
- Gorn, G. J., Jacobs W. J., & Mana, M. J. (1987). Observations on Awareness and Conditioning *Advances in Consumer Research*, 14(1), 415–416.
- Grønhoj, A., & Bech-Larsen, T. (2010). Using vignettes to study family consumption processes. *Psychology & Marketing*, 27(5), 445–464.
- Hertwig, R., & Ortmann, A. (2008a). Deception in Experiments: Revisiting the Arguments in its Defense. *Ethics and Behavior*, 18(1), 59–92.
- Hertwig, R., & Ortmann, A. (2008b). Deception in social psychology experiments: Two misconceptions and a research agenda. *Social Psychology Quarterly*, 71(3), 222–227.
- Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015a). The median split: Robust, refined, and revived. *Journal of Consumer Psychology*, 25(4), 690-704. doi: 10.1016/j.jcps.2015.06.014
- Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015b). Toward a more nuanced understanding of the statistical properties of a median split. *Journal of Consumer Psychology*, 25(4), 652–665. doi: 10.1016/j.jcps.2014.12.002
- Jamison, J., Karlan, D., & Schechter, L. (2008). To deceive or not to deceive: The effect of deception on behavior in future laboratory experiments. *Journal of Economic Behavior & Organization*, 68(3–4), 477–488. doi: 10.1016/j.jebo.2008.09.002
- Khan, J. (2011). Validation in marketing experiments revisited. *Journal of Business Research*, 64(7), 687–692. doi: 10.1016/j.jbusres.2010.08.009
- Kidd, R. F. (1976). Manipulation checks: Advantage or disadvantage? *Representative Research in Social Psychology*, 7(2), 160–165.
- Kimmel, A. J. (2001). Ethical trends in marketing and psychological research. *Ethics & Behavior*, 11(1), 131–149.
- Kimmel, A. J. (2004). Ethical trends in marketing and psychological research 112th annual APA convention. Honolulu.
- Kimmel, A. J. (2011). Deception in psychological research—a necessary evil? *The Psychologist*, 24(8), 580–585.
- Klein, J. (2014). *Working with Groups (Psychology Revivals): The Social Psychology of Discussion and Decision*. Routledge.
- Kwon, S., & Weingart, L. R. (2004). Unilateral Concessions From the Other Party: Concession Behavior, Attributions, and Negotiation Judgments. *Journal of Applied Psychology*, 89(2), 263–278.

- Lehmann, D. R., McAlister L., & Staelin, R. (2011). Sophistication in Research in Marketing. *Journal of Marketing*, 75(4), 155–165. doi: 10.1509/jmkg.75.4.155
- McClelland, G. H. Jr, Lynch, J. G., Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, Type II errors, and false—positive consumer psychology: Don't fight the power. *Journal of Consumer Psychology*, 25(4), 679–689. doi: 10.1016/j.jcps.2015.05.006
- Melancon, J. P., Noble, S. M., & Noble, C. H. (2011). Managing rewards to enhance relational worth. *Journal of the Academy of Marketing Science*, 39(3), 341–362. doi: 10.1007/s11747010-0206-5
- Mukhopadhyay, A., & Johar, G. V. (2009). Indulgence as self-reward for prior shopping restraint: A justification-based mechanism. *Journal of Consumer Psychology*, 19(3), 334–345. doi: 10.1016/j.jcps.2009.02.016
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(9), 776–783.
- Orne, M. T. (1969). Demand Characteristics and the Concept of Quasi-Controls. In R. Rosenthal, & R. L. Rosnow (Eds.) *Artifact in Behavioral Research* (pp. 143-179). New York: Academic Press.
- Pascual-Leone, A., Singh, T., & Scoboria, A. (2010). Using deception ethically: Practical research guidelines for researchers and reviewers. *Canadian Psychology*, 51(4), 241–248. doi: 10.1037/a0021119
- Perdue, B. C., & Summers, J. O. (1986). Checking the Success of Manipulations in Marketing Experiments. *Journal of Marketing Research*, 23(4), 317–326.
- Perrien, J. (1997). Repositioning Demand Artifacts in Consumer Research. *Advances in Consumer Research*, 24(1), 267–271.
- Pham, M. T. (2015). Is it OK to dichotomize? A research dialogue. *Journal of Consumer Psychology*, 25(4), 650–651. doi: 10.1016/j.jcps.2015.07.003
- Robinson, M. D., and Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin*, 27(11), 1520–1532. doi: 10.1177/01461672012711012
- Rosenthal, R., & Rosnow, R. L. (1969). *Artifact in Behavioral Research*. New York: Academic Press.
- Rosnow, R. L., & Aiken, L. S. (1973). Mediation of artifacts in behavioral research. *Journal of Experimental Social Psychology*, 9(3), 181–201. doi: 10.1016/0022–1031 (73)90009-7
- Rucker, D. D., McShane, B. B., & Preacher, K. J. (2015). A researcher's guide to regression, discretization, and median splits of continuous variables. *Journal of Consumer Psychology*, 25(4), 666–678. doi: 10.1016/j.jcps.2015.04.004.
- Rungtusanatham, M., Wallin, C., & Eckerd, S. (2011). The Vignette in a Scenario-Based RolePlaying Experiment. *Journal of Supply Chain Management*, 47(3), 9–16. doi: 10.1111/j.1745-493X.2011.03232.x
- Sawyer, A. G. (1975). Demand Artifacts in Laboratory Experiments in Consumer Research. *Journal of Consumer Research*, 1(4), 20–30.
- Sawyer, A. G. (1977). The Role of Role Playing in Experiments about Consumer Behavior. In B. A. Greenberg, & D. N. Bellenger (Eds.), *Contemporary Marketing Thought* (pp. 191-194). Chicago: American Marketing Association.
- Sawyer, A. G., Lynch, J. G. Jr, & Brinberg, D. L. (1995). A Bayesian Analysis of the Information Value of Manipulation and Confounding Checks in Theory Tests. *Journal of Consumer Research*, 21(4), 581–595.
- Shimp, T. A., Hyatt, E. M., & Snyder, D. J. (1991). A Critical Appraisal of Demand Artifacts in Consumer Research. *Journal of Consumer Research*, 18(3), 273–283.
- Shimp, T. A., Hyatt, E. M., & Snyder, D. J. (1993). A Critique of Darley and Lim's "Alternative Perspective". *Journal of Consumer Research*, 20(3), 496–501.
- Simmons, C. J., Bickart, B. A., & Lynch, J. G. Jr (1993). Capturing and Creating Public Opinion in Survey Research. *Journal of Consumer Research*, 20(2), 316–329.
- Smith, A. K., Bolton, R. N., & Wagner, J. (1999). A Model of Customer Satisfaction with Service Encounters Involving Failure and Recovery. *Journal of Marketing Research*, 36(3), 356-372.
- Smith, N. C., Kimmel, A. J., & Klein, J. G. (2009). Social contract theory and the ethics of deception in consumer research. *Journal of Consumer Psychology*, 19(1), 486–496.
- Sparks, B. A., Bradley, G. L., & Callan, V. J. (1997). The Impact of Staff Empowerment and Communication Style on Customer Evaluations: The Special Case of Service Failure. *Psychology & Marketing*, 14(5), 475–493.
- Sternthal, B., Tybout, A. M., & Calder, B. J. (1987). Confirmatory Versus Comparative Approaches to Judging Theory Tests. *Journal of Consumer Research*, 14(1), 114–125.
- Stolte, J. F. (1994). The context of satisficing in vignette research. *The Journal of Social Psychology*, 134(6), 727–733.
- Surprenant, C. F., & Churchill, G. A. Jr (1984). Can Role Playing be Substituted for Actual Consumption? *Advances in Consumer Research*, 11(1), 122–126
- Surprenant, C. F., & Solomon, M. R. (1987). Predictability and Personalization in the Service Encounter. *Journal of Marketing*, 51(2), 86–96.
- Wagner, T., Hennig-Thurau, T., & Rudolph, T. (2009). Does Customer Demotion Jeopardize Loyalty? *Journal of Marketing*, 73(3), 69–85. doi: 10.1509/jmkg.73.3.69
- Wetzel, C. G. (1977). Manipulation checks: A Reply to Kidd. *Representative Research in Social Psychology*, 8(2), 88–93.