

Chapter 10

Challenges to Validity

from the Standpoint of Methodological Pluralism: The Case of Survey Research in Economics



2
3
4
5
6

María Caamaño-Alegre

7

Abstract By focusing on the case of survey research in economics, the paper shows how methodological pluralism emerges as a natural consequence from a very common dynamics of feedback between problems and solutions taking place in scientific practice. This continuous feedback between methodological problems and attempts at solving them, being essentially connected with the pursuit of validity, naturally leads to the pluralistic tendency found in empirical research in economics over the last decades and clearly manifest in the case of survey research. The methodological challenges within the latter mainly come from the pervading presence of framing effects in survey research, which, as argued here, prompts the application of new procedures able to improve the different kinds of validity.

8
9
10
11
12
13
14
15
16
17

Keywords Validity · Survey research · Framing effects · Methodological pluralism

18
19

The idea that science makes progress from old problems to new problems underlies methodological approaches to science as influential and diverse as those by R. K. Popper (1991/1999, chapter IX), T. S. Kuhn ([1962]1970, chapters IV, IX) or L. Laudan (1981, chapter VII). In traditional, contemporary philosophy of science, much attention has been paid to theoretical problems and questions on the relationship between theory and experience. It has been only relatively recently, over the last decades, that there has been a turn towards the study of scientific

20
21
22
23
24
25
26

This work has been financially supported by the research projects “Laws and Models in Physical, Chemical, Biological, and Social Sciences” (PICT-2018-03454, ANPCyT, Argentina), and “Stochastic Representations in the Natural Sciences: Conceptual Foundations and Applications (STOCREP)” (PGC2018-099423-B-I00, Spanish Ministry of Science, Innovation and Universities).

M. Caamaño-Alegre (✉)
Dpto. Filosofía, Facultad de Filosofía y Letras, Valladolid, Spain

Table 10.1 Feedback dynamics between problems and solutions leading to methodological pluralism

	First stage	Second stage	Third stage	Following stages	
Beginning of the feedback between problems and solutions	Empirical problems associated to procedure 1: Proc ₁	Solutions to Prob ₁ dependent on procedure 2: Proc ₂	Problems associated to Proc ₂ : Prob ₂	Feedback between problems and solutions continues	18.1
Pluralistic tendency	Employment of procedure 1 (Proc ₁)	Employment of Proc ₁ y Proc ₂		Employment of Proc ₁ , . . . Proc _n	18.2

practice, opening the possibility to explore the *praxical* and experimental side of scientific development. As it turns out, however, the inquiry into scientific practice reinforces the view that science is a problem-solving activity, one in which methodological refinement and innovation plays a major role. The abovementioned process of feedback between problems and attempts at solving them very often requires the devise of new methodological options, thereby favoring the pragmatic vindication of a certain methodological pluralism. In what follows I analyze how this dynamics unfolds within the field of empirical research in economics, particularly in the subfield of survey research. The schematic features of such dynamics are represented in Table 10.1 appearing next.

Over the last decades, the concern of economic researchers with the serious empirical limitations associated to traditional observational methods has led to a methodological pluralism, which in turn has given rise to new problems as well as new attempts at solutions. More broadly, in the case of social sciences, where intervention in the subjects' behavior is frequently verbal or at least dependent on language, the use of a wide variety of linguistic means to gather information about their beliefs, expectations, assessments or planned courses of action has significantly increased. In this respect, the reliance on various kinds of surveys and interviews has extended substantially, thereby widening the scope of linguistic interventions beyond the directions verbally conveyed to the experimental subject. In economics, survey research supplemented observation of choices to overcome the ambiguity of these observations, hence improving the validity of causal inferences about preferences. Yet, survey research faces its own validity challenges in the form of framing effects, which threatens all the main forms of validity together with the related feature of reliability. Consistently divergent answers to apparently the same questions concerning preferences preclude any possibility of determining robust correlations between questions regarding options and answers expressing preferences, making thus impossible to draw sound causal inferences and ultimately to attain further forms of validity that presuppose reliability, robust correlations and sound causal inferences.

The discussion of the above issues is structured as follows: first, the standard notion of validity in research methodology is characterized; second, the methodological pluralism connected with the increasing use of linguistic interventions

within economic methodology is examined; third, the challenge of detecting, explaining and controlling framing effects is analyzed; finally, some conclusions as to the validity challenges involved are drawn from the previous discussion.

10.1 The Standard Characterization of Validity in Research Methodology

Traditional philosophical approaches to validity were especially concerned with both theory testing and the attribution of logical rationality to science (see Messick's 1989 discussion of the subject). The two classical accounts of validity in Philosophy of Science, putting aside the more recent contributions by philosophers of experiment like I. Hacking (1983), A. Franklin (2005), P. Galison (1997), F. Steinle (1997) or D. G. Mayo (1996), respectively revolve around the notions of verifiability and falsifiability. A more extensive discussion of experimental validity was attempted instead within the emergent field of social science, specifically, within the methodology of empirical psychology. In this field such discussion embraces more than just those research components devised for the purpose of testing a theory. Here we pay attention to the enlarged view on validation coming from the social sciences. Within this area, validity and reliability are characterized both as logically independent notions and as commonly associated properties of measurements and procedures. Reliability concerns the extent to which an experiment, test, or any measuring procedure yields the same results in repeated trials under the same conditions (Pelham and Blanton 2003, 70–77, Carmines and Zeller 1979, 11–13), while the validity concerns the degree of success in attaining the purported outcome (that is, in determining the variable under study). The common association between reliability and validity is due to the fact that the first is usually required in order to establish the validity of the procedure or just to guarantee its useful applicability. Pelham and Blanton (2003, 70–75) distinguish three main forms of reliability: inter-observer agreement, internal consistency (or inter-item agreement in the same test), and temporal consistency (or test-retest reliability). A variety of statistical methods have been developed for the latter's careful assessment. The unreliability of a method has two possible, general sources: the uncertainty of the phenomena measured, and the errors of measurement (whether chance error, systematic or instrumental).

The notion of validity was originally developed from two different traditions in social science, namely, experimental and test research (Table 10.1). In the 1950s, the basic distinction corresponding to the first tradition was that between internal and external validity (Campbell 1957). As for the second tradition, the main kinds of validity were criterion, content, and construct validity (Cronbach and Meehl 1955). In an attempt to cope with different methodological challenges, test community gradually embraced an enlarged and unitary concept of validity (Angoff 1988, 25; Sireci 2009), one based on a comprehensive notion of construct validity which

comprises all sorts of empirical support for test interpretation and use (Messick 1989). Contrary to this, Shadish et al. (2002) kept the primary association of validity with the truth of knowledge claims, and integrated criterion, internal, construct, and external validity in other unitary framework applicable to either kind of social empirical research (i.e., experimental and test-based). Let us briefly consider each of these kinds of validity.

Statistical conclusion validity (or criterion-related validity as labeled by Carmines and Zeller 1979),¹ in the case of two variables, concerns the appropriate use of statistics to infer whether the presumed independent and dependent variables are correlated. It thus refers to how large and reliable is the co-variation between the presumed cause and effect (Campbell 1986).² That which is targeted by the measurement constitutes the criterion-variable to assess that validity of the method (or instrument). To establish criterion-related validity it is necessary to measure how well one variable (or set of variables), usually called “independent variable” or “intermediate variable”, predicts an outcome, usually called “dependent variable” or “ultimate variable”, based on information from other variables. Criteria validity depends on the extent to which the measures are demonstrably related to concrete criteria in the “real” world. When the criterion variable has current rather than future existence, the validity involved is called “concurrent validity”, otherwise, this kind of validity is referred to as “predictive validity”.

Internal validity refers to whether the co-variation between the presumed independent and dependent variable results from a causal relationship. It is then concerned with the causal interpretation of the criterion-related or statistical conclusion validity. Pelham and Blanton (2003, 62–64) point out that laboratory experiments prove to be very useful in providing information about causality, since they make it possible to isolate independent variables from potential sources of contamination, thereby providing better conditions for controlling individual differences.

In its standard presentation, construct validity is equated with the evidential basis of test interpretation (Messick 1989, 34). It concerns the extent to which a particular empirical indicator (or a set of indicators) represents a given theoretical concept, that is, the extent to which independent and dependent variables truly represent the abstract, hypothetical variables of interest to the scientist (Pelham and Blanton 2003, 66; Shadish et al. 2002, 65). The evaluation of construct validity involves close examination of the auxiliary theory or theories specifying the relationship between concepts and indicators. Such evaluation, therefore, entails examining whether a measure of a construct relates to other measures as established by sound auxiliary hypotheses concerning the construct’s empirical content (cf. Carmines and Zeller 1979, 23).

¹Carmines and Zeller have not been included in the above graphical representation because the notion of criterion-validity, even if named differently, had been introduced much earlier.

²We should insist here on Mayo’s contribution to this topic within the field of philosophy and her emphasis on the significance of statistics for the epistemology of experiment.

External validity refers to the appropriateness of generalizations from results obtained in an experimental setting to phenomena out of such setting. It thus concerns the extent to which a set of research findings provide an accurate description of what typically happens in the real world (Pelham and Blanton 2003, 64). If that which is generalized is a causal relationship, then construct validity consists in the validity of inferences about whether the cause-effect relationship holds over variation in samples, settings, and measurement variables. The two main sorts of generalizations pursued within experimental research are those with respect to some type of entity, and those with respect to some types of situations. On the other hand, the main restrictions to the generalizability of a finding are given by the boundary conditions restricting the attainability of these findings.

One way to address the problem of validation in qualitative terms is to analyze the factors jeopardizing the different kinds of validity. The capability of a method to avoid being affected by those factors can be considered as a sound indicator of how valid the method is. Noise and confounds are the main general threats to both statistical conclusion and internal validity, which are difficult to obtain even in experimental settings where variables can be partially isolated to test their impact on a single dependent variable. On the other hand, construct underrepresentation and construct-irrelevant variance in the test must be emphasized as the general types of threats faced by construct validity. In evaluating the adequacy of construct measurement, not only random errors but also systematic ones need to be considered. The latter may occur because of one or both of the following reasons: (i) tests leave out something that should be included (according to the construct theory), and (ii) they include something that should be left out. There are two main requirements of construct validity needed for protecting interpretations from these general types of threats, namely, convergent and discriminant evidence. The first enables us to assess the degree to which the construct's implications are realized in empirical score relationships, the second provides grounds to argue that these relationships are not attributable instead to distinct alternative constructs (Messick 1989, 34). Finally, some of the typical factors jeopardizing external validity (or generalizability of findings) are the contrived nature of the testing settings, and selection biases. Both issues are related to the general problem concerning the artificiality of laboratory experiments. This problem becomes evident not only in the difficulty to make experimental settings significantly similar to the targeted real situations, but also in the limitations to recruit a group of experimental subjects that are representative enough of the diverse population under study. Pelham and Blanton (2003, 66–67, 176) suggest several ways to minimize the effect of these threats: combining laboratory research with passive observational studies, randomizing subject's selection, and using manipulation checks with the subjects to make sure whether the intended experimental variation corresponds to the one accomplished in the experiment.

10.2 Linguistic Intervention as a Source of Pluralism in Economic Methodology

179

180

From a general point of view, verbal intervention raises two independent problems: the first concerns the possibility of determining whether the meaning assigned by the researcher to the utterances in the research context is the same as that assigned by the respondent; the second is related to the influence that different ways of presenting the same issue may bear on the respondent's response. Even though the focus of this paper is survey research in economics, a subject more often connected to the second problem, it may be worth it to mention the recent attempt to address the first issue on the grounds provided by several philosophers of language with crucial contributions within the field of pragmatics. Even if only very tentatively, several possible sources of discrepancy in understanding an utterance have been examined on the basis of John Searle's distinction between linguistic meaning and speaker's meaning – later supplemented by that between linguistic meaning and utterance meaning (cf. Searle 1978). Some of those discrepancies would be caused by the fact that each user of language would associate to an expression some contents beyond its literal meaning, some others would be due to contextual aspects affecting the use of expressions. Paul Grice's pragmatic approach relies on similar distinctions as those drawn by Searle and aims at explaining the same communicative difficulties pointed out by the latter. According to Grice, communicative intentions, which are the essential element in linguistic activity, are subject to the principle of cooperation, which in turn would unfold in different conversational maxims or principles (cf. Grice 1975). The application of this Gricean analysis to the study of empirical research in economics is still at a very early stage (cf. Schwarz 1996; Jones 2007; Geurts 2013), but certainly a pragmatic analysis of the communicative exchanges between the experimenter (or interviewer) and the experimental subject (or respondent) is required in order to determine the discrepancies between the message that the experimenter is trying to convey and the message grasped by the subject.

Linguistic intervention pervades the field of social science, whether by means of directions provided to the experimental subject or by interviews and surveys intended to collect information about the effect that certain issue has on the subjects' beliefs, choices or behavior. In parallel to this variety of verbal interventions, new difficulties concerning the so called "framing effects" are detected, which in turn prompt new attempts at sophisticating the methodological procedures in order to confront them.³ However, within the sphere of economic methodology, there have been two main prevailing assumptions whose endorsement has led respectively to either rule out or question the effectiveness of verbal empirical

³A traditional methodological principle endorsed in economics, namely, the one establishing that the experimental subject should receive written (not spoken) directions amounts to implicitly acknowledging the risk of introducing unwanted effects and possible confounds within the experimental context through the communicative interaction between the experimenter and the experimental subject.

procedures. The first assumption comes from the theory of revealed preference, committed to the methodological principle that inferences about a subject's future choices must be based on observations of previous choices made by the subject. The second assumption, by contrast, is one underlying the economic methodology expanded with survey research and interviews. Such assumption, usually referred to as the principle of extensionality or the invariance principle (Bourgeois-Gironde and Giraud 2009, 385–387), establishes that individuals' preferences should not be affected by variations in the description of a problem. Different ways of presenting the same set of possible options should thus not change the subjects' choices with respect to those options. As we will see next, some of the difficulties related to the above assumptions have motivated the pluralistic expansion of economic methodology.

With respect to the first assumption, it is worth emphasizing that a serious shortcoming affecting the theory of revealed preference stems from the ambiguity of subjects' observable behavior and the resulting inscrutability of expectations or radical under-determination of attribution of expectations. Since expectations, together with preferences, are acknowledged as crucially involved in the subjects' choices, the inscrutability of the former poses a major obstacle to explaining such choices. The problem emerges when researchers need to establish some suppositions about the subjects' expectations in order to make predictions. Given that expectations are not accessible through observation, information about them must be obtained by verbal means. The need to validate the verbal means employed in gathering information about expectations emerges from the very recognition that such means are needed in order to make progress. This methodological turn occurring in economics at the beginning of the 1990s results in the recognition of declared preferences, in addition to revealed preferences, as a legitimate evidential source in economics.⁴

As Charles F. Manski points out, from the early 1990s, economists who engaged in survey research have increasingly used questions regarding subjects' probabilistic expectations concerning significant personal events:

Observed choices may be consistent with many alternative specifications of preferences and expectations, so researchers commonly assume particular sorts of expectations. It would be better to measure expectations in the form called for by modern economic theory; that is, subjective probabilities. Data on expectations can be used to relax or validate assumptions about expectations. Since the early 1990's, economists have increasingly undertaken to elicit from survey respondents probabilistic expectations of significant personal events (Manski 2004, 1329).

Expectations have been determined for various kinds of events, among them: macroeconomic events (stock market returns), the risks faced by a person (job loss, mortality), future income (earning, Social Security profits), and choices made by a person (purchases, voting choices).

⁴An influential comprehensive criticism of the theory of revealed preference can be found in Daniel Hausman (2012).

Table 10.2 Pluralism in empirical economics emerging from a continuous feedback between methodological problems and attempts at solving them

	First stage	Second stage	Third stage	Forthcoming stage	
Beginning of the feedback between problems and solutions	Problems related to ambiguity of observed behavior: Inscrutability of expectations through observation	Solutions to ambiguity dependent on survey research	Problems associated to declared preference procedures: Framing effects	Solutions to framing effects dependent on self-reports, post-survey questionnaires...	t11.1
Pluralistic tendency	Employment of “revealed preferences”	Employment of both revealed and declared preferences procedures		Use of revealed and more refined, robust declared preferences procedures	t11.2

The cycle of methodological refinement described at the beginning of the paper appears here very clearly: the attempt to improve both the predictive effectiveness and the descriptive accuracy of economics goes hand in hand with its methodological widening. As reflected in the table below, this broadening of the methodological scope leads, like in other fields, to a methodological pluralism of a pragmatic kind (Table 10.2).

The pragmatic side of this tendency towards methodological pluralism should be understood along the lines of the methodological pragmatism put forward by Nicholas Rescher, which closely resembles the one tacitly embraced by researchers in their current practice (cf. Rescher 1977; Suppes 1998). According to this pragmatist standpoint, the question about the validity of procedures is not one to be answered *a priori* (cf. Wiener 1973–1974, 551–556; Haack 2006), but instead one to be assessed according to the usefulness of such procedures to attain certain epistemic ends (cf. Caamaño-Alegre 2013). It must be noted that the essential goal of increasing predictive power will be achieved to the extent that researchers manage to improve statistic, internal, construct and external validity of the procedures they employ. Similarly, descriptive accuracy is closely connected to the validity of the theoretical construct used in explaining behavior. Therefore, the different kinds of validity involve a specification of epistemic ends relative to which understand and evaluate the methodological developments. The growing interest raised by mixed methods and triangulation in economic methodology constitutes another clear sign of the pluralistic tendency in this field, a tendency with the underlying purpose of strengthening both methodological robustness and the empirical adequacy of theories (cf. Dellinger and Leech 2007; Downward and Mearman 2007; Starr 2014; Claveau 2011).

Let us go back to the side of this pluralistic trend that is the focus of this paper, namely, the use of surveys in the context of empirical research in economics. The shortcomings affecting the theory of revealed preference, in particular the need to

identify expectations, called for an empirical research by means of surveys, which was initially carried out according to the abovementioned principles of invariance and extensionality. However, despite the use of surveys in economic methodology, it took a long time until the problem of framing effects was properly noticed. Ivan Moscatti draws attention to this fact as he states:

Orthodox economists tend to discard framing effects as manifestations of the irrationality of individuals who simply fail to recognize that identical things are indeed identical. In opposition to this view, Tversky and Kahneman and other behavioral economists have argued that framing effects significantly influence economic behavior and therefore cannot be discarded without weakening the descriptive significance of economic theory; moreover, some framing effects seem to have a rational justification (Moscatti 2012, 6–7).

Behavioral economists have therefore diverged from the prevailing view of framing effects in economics, arguing that such effects should be approached, not as mere cognitive flaws in the recognition of identical options, but as signs of the subjects' attitudes towards different aspects involved in those options. So understood, framing effects turn out relevant for the description, explanation and prediction of the subjects' economic behavior. To put it clearly, the methodological problem of framing effects has encouraged the study of the role that language and communication play in subjects' understanding of the described options. As the use of surveys exponentially increases in the economic field, the need to pay attention to framing effects becomes more pressing. Michaela Nardo provides some interesting data in this respect:

The European Union, as well as the main OECD countries, regularly collect data from business and consumer surveys. The number of these surveys has substantially increased in the last three decades. If in the late 1960s they were less than 30 in 15 countries, in 1997 their number exceeded 300 in 55 countries. Only in the European Union more than 50.000 firms and 20.000 consumers are interviewed each month. Surveys address firms or agents directly, and rather than asking for exact figures, the questionnaires ask for assessment on the movement of short-term variables, such as output, prices, employment, trade, or investments (Nardo 2003, 645).

She warns us, however, that several difficulties underlie the use of (aggregates of results from) surveys with the purpose of building empirically valid representations of expectations as the basis for inferring the agents' future behavior. On the one hand, the agents themselves can fail to estimate their expectations; on the other hand, the frame in which a survey is presented can influence the expectations they declare (cf. Nardo 2003, 657–59).⁵ Even though the present paper highlights the use of surveys for research purposes, it is worth noting that also their practical use

⁵Nardo is of course aware of some other possible sources of error not related to the subjects' performance: "If survey data are a poor indicator of agents' expectations, then the quantified proxy will also poorly predict the behavior of the actual economic variable even if agents are perfectly rational. This is far from trivial, since being survey data approximations of unobservable expectations, they necessarily entail a measurement error. This error can be ascribed to the incorrect scaling of qualitative data, to sampling or aggregation errors and also to the general uncertainty attached to survey figures" (2003, 657).

entails methodological challenges that need to be addressed in order to guarantee the effectiveness of those procedures as tools for prediction and for gathering of information. In the next section, I examine in more detail what Nardo points out as the second source of difficulties affecting survey research.

10.3 The Problem of Framing Effects: Detection, Explanation and Control

Let us focus now on the specific problem of framing effects and the main attempts at accounting for them. After presenting a comprehensive classification of such effects, I will deal with the current attempts at explaining and controlling them, making special emphasis on the difficulties involved in the pursuit of validity in survey research.

10.3.1 The Detection and Classification of Framing Effects

As soon as the late 1990s, Levin et al. (1998) urged researchers to sophisticate the typology of framing effects so that it became possible to account for the apparently inconsistent results achieved when trying to detect such effects. The plurality of interventions, moreover, entails a corresponding plurality of framing effects whose treatment requires equally differentiated procedures. In the typology suggested by Levin, Schneider and Gaeth, three main kinds of valence framing effects are distinguished: the extensively discussed risky choice framing effect, and two other effects often overseen or mistaken for the latter, namely, attribute framing and goal framing. As explained by the authors (1998, 151, 181), each frame differs from the others in what is framed, what the frame affects, and how the effect is measured.

In the risky choice framing, the complete set of outcomes from a potential choice involving options with different levels of risk is described either in a positive or negative way. The framing effect is here measured comparing the rate of choices for risky options in each frame condition. Risk aversion would explain the fact that, when presented in negative terms, the riskier option is chosen by respondents more often than the safer one. A wide variety of experiments on risky choice, from bargain situations to medical treatments, shows that when the outcome is described in terms of gains (lives saved, earned income) subjects' tendency to take risks diminishes. By contrast, such tendency increases when outcomes are expressed in terms of losses (lost lives, incurred debts). The paradigmatic case of risky choice framing effect is illustrated by the so called "Asian disease problem" (cf. Tversky

⁶See Levin et al. (1998, 154–157) for a collection of experimental results obtained within the domain of risky choice framing effects.

and Kahneman 1981). In this task, the two equivalent pairs of independent options with different level of risk are the following: (a) a sure saving of one-third the lives versus a one-third chance of saving all the lives and a two-thirds chance of saving no lives; (b) a sure loss of two-thirds the lives versus a one-third chance of losing no lives and a two-thirds chance of losing all the lives. The majority of subjects select the first option in the positively framed version of the task, and the second option in the negatively framed version.

In the form of framing called “attribute framing”, the positive or negative description of some characteristic of an object or event affects item evaluation, which is estimated by comparing the attractiveness ratings for the single item in each frame condition. The associative processes based on valence is commonly assumed to explain the fact that positively described objects or events are more positively valued. This result has been established with much higher reliability and robustness than the other two kinds of framing effects compared by Levin et al. (1998, 160). The fact that evaluations vary as a result of positive or negative framing manipulation has been established for issues as diverse as consumer products, job placement programs, medical treatments, industry project teams or students’ level of achievement or the performance of basketball players.⁷ Ground beef, for example, was rated as better tasting and less greasy when it was described as 75% lean rather than as 25% fat. Similarly, students’ performance was rated higher when their scores were expressed in terms of percentage correct or percentage incorrect. Analogous results were obtained in the rest of cases.

Finally, in the case of goal framing, the same consequences of a conduct are specified either in positive or negative terms. The positive frame focuses attention on the goal of obtaining the positive consequence (or gain) associated with a given behavior, whereas the negative frame focuses attention on avoiding the negative consequence (or loss) associated with not performing such behavior. The variation in how persuaded an agent is to make or not make the decision to perform a certain conduct is regarded as an effect of the variations in the frames applied. The effect itself is measured by comparing the rate of adoption of such conduct under each frame condition. Experimental evidence shows that the negatively framed message, that is, the one emphasizing avoidable losses, proves more persuasive than the same message framed positively, and therefore stressing the potential gains. Real examples where goal frames are at use can be found in studies on the promotion of health, on endowment or on social dilemmas. Most subjects appear more inclined to adopt a certain conduct, –like for example, breast self-examination, use of public resources or of credit card–, when they receive information stressing the potential losses derived from not engaging in such conduct than when presented with information highlighting the potential profits resulting from engaging in it.

In the abovementioned examples, individuals show themselves more persuaded to adopt a given behavior when descriptions emphasize, respectively, the decrease

⁷See also Levin et al. (1998, 161–163) for a lengthy compilation of experimental results related to attribute framing effects.

in the probability of detecting a cancer if there is no self-examination carried out 396
 versus the increase of such probability in case a self-examination is performed, the 397
 losses suffered by the individual who contributes to the public goods versus the 398
 foreseen gains if the individual contributed to them, and the losses due to not using 399
 the credit card versus the benefits derived from its use.⁸ 400

10.3.2 *The Attempts at Explaining Framing Effects* 401

Despite the growing interest raised by the problem of framing effects, the majority 402
 of studies on these effects are focused on their diagnosis, while the attempts at 403
 explaining and controlling them are still extremely tentative and fragmentary. As 404
 already pointed out by Tversky and Kahneman (1981, 1991) in several of their 405
 influential studies on framing effects, the task of devising frames must be done 406
 taking into account individuals susceptibility to changes in reference points or in 407
 what is perceived as the status quo regarding some issue. Different frames would 408
 lead to different choices of reference points and, consequently, to a different way 409
 to encode the outcomes as gains or losses, which would accordingly bring about a 410
 different selection of options. This clearly calls for the development of procedures 411
 that can disclose such susceptibility on the side of the respondents. 412

In developing their prospect theory (cf. Kahneman and Tversky 1979), both 413
 authors appeal to the possible occurrence of highly intertwined phenomena like loss 414
 aversion and the endowment effect. These phenomena would emerge in most cases 415
 due to some framing conditions in which the reference point regarding the value of 416
 an outcome does not stay neutral but varies depending on what is induced by the 417
 frame itself. In their own words: 418

However, the location of the reference point, and the consequent coding of outcomes as 419
 gains or losses, can be affected by the formulation of the offered prospects, and by the 420
 expectations of the decision maker (Kahneman and Tversky 1979, 274). 421

Let us recall that prospect theory, as opposed to classical theory, is committed to 422
 the view that risk aversion is dependent on a reference point. Under that assumption, 423
 it is predicted that risk aversion is linked to the domain of gains and risk seeking to 424
 domain of losses. In their paper from 1979, Kahneman and Tversky established 425
 that the above tendency could be reversed depending on the framing employed 426
 for the same pair of options. An initial remark in that direction can be found in 427
 some of their comments on the isolation effect (Kahneman and Tversky 1979, 271), 428
 that is, individuals' inclination to ignore those components shared by alternatives 429
 and to focus on those making them different. Since there is more than one way to 430
 decompose a pair of alternatives in shared and distinctive components, the different 431
 ways of decomposition may also prompt different preferences. This point is made 432

⁸The wide range of real cases collected by Levin et al. (1998) can be found in 169–171.

more explicit as both authors identify the reference point assumed by individuals with those individuals' status quo or current state.

Kahneman and Tversky go into great detail as to how reference points may vary, emphasizing that those reference points fixed by the status quo may shift as a result of encoding losses and gains relative to expectations that differ from the ones determined by the status quo. They also mention more specific cases where different encodings of the same pair of options create discrepancies between the reference point and the actual situation. According to them, this is exactly what happens when the choice is encoded in terms of final outcomes, as suggested from decision theory, instead of in terms of losses and gains (cf. Kahneman and Tversky 1979, 286–287).

A variation in the way a message is encoded, therefore, entails a change of context that has both cognitive and motivational consequences. Such consequences will depend on the kind of encoding that is being used. Considering all forms of framing effects detected so far would go beyond the scope of the present paper, which is limited to the so called “valence framing effects”, that is, those effects resulting from a positive or a negative encoding of a message. Scientific research into these kinds of effects has led various authors to try to complete the list of variables involved in processing different encodings, thereby explaining the corresponding framing effects. In addition to loss aversion, endowment, preservation of the status quo and the tendency to ignore similarities –all of them trends acknowledged by Kahneman and Tversky in their studies on risky choice framing–, Levin and his collaborators point to the activation of positive associations in memory as the main mechanism responsible for framing effects (cf. Levin et al. 1998, 164–165). Positive stimuli generated by a frame would yield some associative responses that, in turn, would cause a clear increase in the level of approval that each individual assigns to the positively described option as opposed to that assigned to the negatively described one. It has even been demonstrated that the mere activation of positive associations with respect to one of the options presented for a given choice brings about substantial positive distortions of that option against the other one (cf. Russo et al. 1996, 103–107). In the experiment on distortion of alternatives carried out by Russo and his co-workers, positive descriptions of the owner of a restaurant or hotel remarkably influenced the more positive evaluation of the restaurant or hotel, despite the fact that such descriptions were logically independent of the attributes of the products offered. These experimental results reveal the same confirmation bias related to selective attention mechanisms as the one that has been observed in more general studies regarding the effect of expectations on judgment.

Turning now to the attempts at explaining goal framing effects, it is worth stressing the strong empirical support for the hypothesis of the negativity bias (cf. Taylor 1991, 68–71). According to this hypothesis, individuals pay more attention to negative information than to equivalent positive information, showing themselves more influenced by the former than for the latter. From the decade of the 1990s, some of the explanations for the different framing effects have been partially unified, more specifically, loss aversion is understood as a subclass of the negativity bias, and

the status quo bias is in turn regarded as a subclass of the loss aversion bias. In all these cases, the rejection caused by a loss is higher than the desire to obtain a gain of the same magnitude (cf. Levin et al. 1998, 177).

10.3.3 *The Pursuit of Control over Framing Effects*

If, especially during the 1990s, the detection, classification, and explanation of framing effects constituted a challenge only partially overcome despite the efforts made to that end, the challenge of controlling such effects has hardly been addressed. Yet, the identification of different bias that are activated according to the kind of frame in use sheds some light on the way individuals process information depending on how the latter is presented to them.⁹ The obvious consequence seems to be that, if a certain form of encoding the message is avoided, the bias caused by such encoding can be avoided too, and, together with it, the introduction of certain variable that detracts from the validity of the survey. All forms of validity – statistical as well as internal, construct and external – could be improved by avoiding the encoding responsible for the bias. Nevertheless, even if researchers decided to proceed this way, the question would remain of what the most neutral possible frame would be, or, to put it differently, what frame would be the least amenable of producing a biased response from the individual decoding the message.

From a pragmatic standpoint, that is, from a view primarily committed to methodological effectiveness, one option would be to examine, among those empirical findings obtained by experimenters working on framing effects, those which point to variables that diminish or prevent such effects. It is important to notice that the above findings have been very scattered and hardly ever replicated, since they have been obtained through studies not directly oriented to determine this sorts of variables, but rather focused on the detection of framing effects. However, despite the more basic goal served by these experiments, in some cases they included additions that turned out enlightening for the purpose of controlling framing effects. In the case of risky choice framing, for example, it was demonstrated that when some question about the subject's reasons for a certain choice was added to the survey, then the framing effect was diminished or even eliminated. It is what R. P. Larrick, E. E. Smith and J. F. Yates call "the reflection effect" (1992, 199), which, according to their results, would make it possible to reverse framing effects by means of reflection on the issue presented within the frame. In a similar vein, Stephen M. Smith and Irwin P. Levin experimentally showed that individuals with a lower need for cognition were more affected by framing effects than those with

⁹Even though references to "the bias" induced by frames can be very often found in research literature, it must be pointed out that in the present context such expression is employed in a broad sense and not with the more restricted of a difference between observable traits of the respondent and what she or he reports (cf. Groves and Singer 2004, 38–39).

a higher need for cognition, who in turn where almost immune to differences in framing (Smith and Levin 1996, 283).¹⁰

Experimental results suggest that factors other than the above also have a bearing on the scope of framing effects. Among these factors there are the domain of problems presented, the traits of experimental subjects, the magnitude or probability of potential outcomes, and the categories applied in verbalizing such outcomes (Levin et al. 1998, 153). For instance, subjects are more inclined to take risks related to health issues than related to finances. The other two cases referred above, however, could be covered by the general case where the amount of information handled by the subject is inversely proportional to the scope of the framing effects (Schoorman et al. 1994, 520). As already observed, the variations in such amount may be due to variations intrinsic to the frame, and basically dependent on how detailed the frame is, or to variations in the subjects, mainly related to their need for cognition or degree of competence on the kind of subject presented. With respect to the traits of experimental subjects, it has been found, for instance, that experts or students in a certain field tend to be less affected by framing effects when confronted with options evaluable from such field. Similarly, it has been verified that replacing expressions like “many” or “few” with numerical values lowers the intensity of framing effects. In the study by Schoorman et al. referred earlier, it has been experimentally established that the subject’s degree of involvement or responsibility concerning a given issue can also eliminate the bias produced by the framing of the issue. Moreover, some recent empirical findings show that the framing bias is eliminated when the implicit frame is presented explicitly (Gamliel and Kreiner 2013; Kreiner and Gamliel 2016), or when the addressee’s attention is drawn to it (Kreiner and Gamliel 2018). All these procedures would help reestablish reliability and validity by increasing the consistency in the answers collected, improving the robustness of statistical correlations, eliminating confounds, ultimately allowing for a better empirical grounding of constructs and a higher generalizability of both results and procedures.¹¹

The situation is somehow different in the case of the bias caused by the attribute frame, for, as noted earlier, the sort of effect produced by this frame is the most homogeneous and clearly verified among ones caused by the valence frames. Thus, despite the different domains of problems or the differences between subjects, the positive description of an item attribute, as opposed to its negative description, will almost always favor the more positive evaluation of both the attribute and

¹⁰Within the field of psychology, the need for cognition constitutes a personality variable reflecting the individuals’ disposition to perform cognitive tasks that require effort.

¹¹As Jiménez-Buedo points out, two different senses of generalizability –and thus of external validity– are usually mixed in the literature; one refers to the degree in which an experimental finding can be considered ‘representative’ of conditions outside of the experiment, and another points to the extent to which such finding can be applicable to parallel situations (cf. Jiménez-Buedo 2011, 276). Without questioning the problematic implications of such ambiguity, it seems that both statistical and internal validity are preconditions for external validity, since, otherwise, there would be nothing to generalize in either sense.

the corresponding item. However, also in the case of attribute framing, a lower intensity of the bias has been experimentally determined when there is, on the subjects' side, a high degree of involvement as to the issue being described (Marteanu 1989, 90–93; Millar and Millar 2000, 860–863). We find here again a phenomenon that suggests an inverse relationship between the intensity of the framing bias and the level of processing of information provided to the subject. This phenomenon might, therefore, support the hypothesis, backed up by the experimental work of Durairaj Maheswaran and Joan Meyers-Levy (1990, 365), according to which the more involved an experimental subject is in the issue described, the more detailed his or her processing of the information related to the issue. Moreover, several experimental studies have shown the occurrence of a closely related phenomenon, namely, that the evaluation of real items is less affected by framing bias than the evaluation of hypothetical items. Attribute framing effects are also diminished when subjects are asked to explain their answers or give reasons for them.

Let us finally briefly consider some possible factors relevant in the control of goal framing effects. Like in the former cases, the degree of involvement in the topic presented, together with the tendency of the subjects to make a cognitive effort, are inversely related to the intensity of the framing effect.¹² Perhaps because of the greater structural complexity of goal framing, there are more variations in operationalizing this framing, which ultimately entails a less homogeneous evidence for goal framing than for attribute framing (Levin et al. 1998, 176). More specifically, such operationalization can be done either through simple negation (not obtaining profits) or through alternative terminology (losing the possibility of obtaining profits). Even if it seems obvious that linguistic variation may influence the strength of all sorts of valence framing effects, there are more potential linguistic variations in the case of goal framing, since the latter involves describing the consequences ascribed to some behavior as opposed to those ascribed to not performing such behavior. As Levin and his co-workers emphasize, in order to clarify when the responses of the subjects are dependent on semantic variations, it is necessary to develop an empirical study on language itself (1998, 174).¹³ Here we find another instance of methodological development connected to newly recognized problems arising from methodological solutions to previous problems.

As for the need to focus on language, it is worth mentioning that there have been some attempts at explaining framing effects in general on the basis of the traditional semantic distinction between extension (what is designated by an expression) and intension (the way of determining extension). From the field of philosophy of economics, for example, Ivan Moscati has recently argued for understanding framing effects as doxastic effects caused by the intensional discrepancy between

¹²Numerous references to empirical studies that point to this issue can be found in Levin et al. (1998, 174).

¹³In his paper from 1992, Rolf Mayer provides some early clues to develop the kind of study suggested above. There he refers to some semantic aspects relevant in framing effects, such as the clustered nature of meaning, the impact of thematic roles or the distinction between discursive background and discursive front.

extensionally identical descriptions. Surveys employed by Tversky and Kahneman in their experiments included extensionally equivalent descriptions of outcomes and probabilities which, nevertheless, intensionally differed by virtue of the way uncertainty was presented, either in one stage games or in two stages games (Moscati 2012, 7). Moscati points to the problem of referential opacity in intensional contexts as that which would explain the apparent irrationality of subjects' tendency to prefer one option over the other:

If we look at framing effects using the notions of intension and extension, they no longer appear to be manifestations of irrationality. Rather, they seem to be just other instances of the failure of the substitutability principle in referentially opaque contexts. Therefore, when looked at from the intension-extension viewpoint, the relevant problem shifts from the issue concerning the individuals' rationality, to the question of whether standard, set-theoretic economic models are able to capture the intensional difference between extensionally equal objects (Moscati 2012, 8).

According to this author, the apparent manifestations of irrationality would be the consequence of an apparent co-extensionality, mistakenly taken as real by those researchers who overlook the opaque nature of intensional contexts such as that of subjects' beliefs.

Although following a different strategy, Sacha Bourgeois-Gironde and Raphaël Giraud (2009, 385–387) also make use of the distinction between intension and extension to explain how framing effects come to happen. Both authors draw attention to the fact that, in economic methodology, the principle of invariance or extensionality goes beyond the logical principle establishing the co-extensionality between expressions whenever the latter are interchangeable *salva veritate* (i.e., whenever truth-value is preserved). In the context survey research, what needs to be guaranteed by means of co-extensional descriptions is not only truth-value preservation but also the preservation of whatever information proves relevant for making decisions. What needs to be specified, therefore, is the kind of information regarded as relevant for purposes of deciding among the options presented. Only after such information had been specified, could framing effects be ascertained as violations of the extensionality principle in the contexts of decision under study. Violating extensionality would then imply that irrelevant information determines the choices or judgments made by the subjects.

Bart Geurts' 2013 article on framing offers another insightful discussion of the linguistic implications of framing effects. In his view, frames support counterfactual reasoning of the sort: if a state of affairs is positively or negatively described, then a different, respectively less or more advantageous state of affairs could have been the case. An important innovation of Geurts' approach is the explanation of framing effects, not only in terms of alternatives, but also in terms of what he calls "alignment". Expressions like 'too' or 'even' would depend on alternatives for conveying the speaker's intended message. For instance, 'even φ ' would mean that φ is true and that φ 's prior probability is low, relative to φ 's alternatives (Geurts 2013, 7). Such alternatives are ordered in a scale and being "stronger" in the scale could be expressed with '>'. According to Geurts, implicatures depending on ordered alternatives support automatic inferences about the correlation (alignment) between

prior probabilities and strength (Geurts 2013, 8). The definition of alignment states that, for any ψ , ψ' that are included among φ 's alternatives, if $\psi > \psi'$ then $\psi \gg \psi'$ (where ' $\psi \gg \psi'$ ' means that ψ is more improbable than ψ'). The intuition behind this definition can be expressed by saying that "‘more’ on the quantity scale entails ‘more’ on the improbability scale" (Geurts 2013, 9). An important point emphasized by this author is that the Alignment assumption is optional (thus not part of the lexical meaning) and operates by default on the basis of world knowledge (Geurts 2013, 10). Our regular exposure to correlations between quantitative and qualitative scales, together with our tendency to establish connections and pursuing coherence, would explain the emergence of alignment assumptions (Geurts 2013, 11). Framing effects would also be a manifestation of this combined phenomenon, they being the result of establishing connections between different frames and different counterfactual alternatives. In applying the above analysis to framing, Geurts arrives at an evaluative understanding of framing effects and, therefore, adds 'it is good that [φ]'in order to uncover the underlying alignment assumptions (with ' \gg ' now meaning 'is better than'). Imagine that an airplane with 600 passengers crashed and we hear that 300 people survived or, alternatively, that 300 people died.¹⁴ Our default alignment assumption would automatically yield the following interpretation for the positively frame description: 300 people survived $\gg n$ people survived (such that $300 > n$). Obviously, this interpretation would be inconsistent with our usual understanding of the negatively framed description, that is to say, we would reject that 300 people died $\gg n$ people died (such that $300 > n$). As Geurts concludes, far from being equivalent, both descriptions convey mutually inconsistent information about counterfactual states of affairs (2013, 12).

From a more pragmatic and pluralistic standpoint, Manski has explored the possibility of overcoming the flaws of economic survey research by following the same methods as those applied in cognitive psychology, which mainly rely on the determination of expectations. Thus, partly relying on a methodological tradition coming from empirical psychology, this author argues that research procedures should include questions about subjects' predictions concerning their own future behavior or self-reports on their own way of making decisions (Manski 2004, 1330–1331). An overall more robust treatment of subjects' patterns of decisions would be achieved by combining two different kinds of evidence, on the one hand, the observed conduct of subjects in making decisions, and, on the other, the self-reports made by respondents. As a result, there would be a wide plurality of procedures employed to determine not only subjects' expectations, but also their preferences, cognitive habits and intentions. Even if bringing with them a whole array of new difficulties, the use of self-reports would enable researchers to estimate several aspects involved in framing effects. In particular, it would make it possible to uncover several potential confounds and hence estimate the amount and kind of information actually processed by the subject, as well as the latter's need

¹⁴I am here slightly modifying Geurts' example for the sake of simplicity.

for cognition and underlying interests or preferences varying his or her attention mechanisms. 671
672

As a consequence of the above, construct validity could be highly improved, 673
since the theoretical explanation of how subjects make decisions would become 674
both theoretically and empirically more detailed by specifying how the postulated 675
preferences are constrained by different psychological aspects involved in the 676
interpretation of the options offered to them. These aspects could be empirically 677
determined to some extent on the basis of survey results, self-reports and observed 678
behavior. Improving construct validity would directly strengthen statistical and 679
internal validity as well. In both cases, it would be possible to isolate the effect 680
of different variables (expectations, understanding of options) that were previously 681
operating as confounds and, for this reason, were obscuring the possible statistical 682
correlation or causal link between the independent variable (preference) and the 683
dependent variable (decision). 684

10.4 Conclusions 685

The present paper has called attention to the continuous process of feedback 686
between empirical problems and procedural solutions taking place in scientific 687
research. It has been argued that such dynamics entails a methodological refinement 688
that naturally leads to a pluralistic methodological development. I have emphasized 689
how this development goes hand in hand with the possibility of improving the 690
validity of empirical research. The case of survey research in economics has served 691
to illustrate this kind of dynamics. Here the attempt at overcoming the shortcomings 692
of the theory of revealed preference, more precisely, the need to determine those 693
expectations involved in decision making, leads to the use of a wide variety of survey 694
procedures in economic methodology. Such use makes it possible to distinguish the 695
effect of preferences on decisions from the effect of expectations, thereby improving 696
statistical and internal validity of correlations between preferences (independent 697
variable) and decisions (dependent variable). This way of controlling variables, 698
however, requires researchers to face the methodological challenge of detecting and 699
controlling framing effects. 700

The previous discussion includes an overview of the main ways to account 701
for framing effects. Some of the most recurrent variables, like loss aversion, 702
positive associations triggered by positive descriptions or selective attention drawn 703
by negative information, point to well entrenched tendencies in most individuals. 704
Other variables, like subjects' degree of involvement, cognitive effort or situated 705
linguistic understanding, more directly reveal the importance of cognitive, semantic 706
and pragmatic factors. The improvement of statistical, internal validity depends 707
on the successful empirical determination of the above factors, and, therefore, on 708
their mitigation as possible confounds. Construct validity, more in particular, the 709
validity of the postulated causes for individuals' decisions, also depends on the 710
identification of the abovementioned factors. In particular, to account for framing 711

effects in terms of mere mistakes in understanding is a wrong approach, given
 all the evidence on such effects examined earlier. Instead of keeping a theoretical
 construct that empirically under-represents the causal factors directly involved in
 decision making, it would be necessary to enrich such construct by establishing an
 aggregate of variables able to empirically represent the phenomenon under study.
 Since preferences and expectations are not the only variables that prove causally
 relevant in decision making, the explanation of the latter should include a reference
 to the rest of variables already mentioned. Finally, since the above kinds of validity
 are preconditions for external validity, the latter can only be obtained after the
 former has been accomplished. The broad range of intricate survey problems to
 be addressed in the future no doubt will require a healthy dose of methodological
 pluralism.

References

- Angoff, W. H. (1988). Validity: An evolving concept. In H. Wainer & H. I. Braun (Eds.), *Test validity* (pp. 19–32). Hillsdale: Lawrence Erlbaum Associates.
- Bourgeois-Gironde, S., & Giraud, R. (2009). Framing effects as violations of extensionality. *Theory and Decision*, 67(4), 385–404.
- Caamaño-Alegre, M. (2013). Pragmatic norms in science: Making them explicit. *Synthese*, 190(15), 3227–3246.
- Campbell, D. T. (1957). Factors relevant to the validity of experiments in social settings. *Psychological Bulletin*, 54, 297–312.
- Campbell, D. T. (1986). Relabelling internal and external validity for applied social scientists. In W. M. K. Trochim (Ed.), *Advances in quasi-experimental design and analysis* (pp. 67–77). San Francisco: Jossey-Bass.
- Carmines, E. G., & Zeller, R. A. (1979). *Reliability and validity assessment*. London: Sage.
- Claveau, F. (2011). Evidential variety as a source of credibility for causal inference: Beyond sharp designs and structural models. *Journal of Economic Methodology*, 18(3), 233–253.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302.
- Dellinger, A. B., & Leech, N. L. (2007). Toward a unified validation framework in mixed methods research. *Journal of Mixed Methods Research*, 1(4), 309–332.
- Downward, P., & Mearman, A. (2007). Retrodution as mixed-methods triangulation in economic research: Reorienting economics into social science. *Cambridge Journal of Economics*, 31, 77–99.
- Franklin, A. (2005). *No easy answers: Science and the pursuit of knowledge*. Pittsburgh: University of Pittsburgh Press.
- Galison, P. (1997). *Image and logic: A material culture of microphysics*. Chicago: University of Chicago Press.
- Gamliel, E., & Kreiner, H. (2013). Is a picture worth a thousand words? The interaction of visual display and attribute representation in attenuating framing bias. *Judgment and Decision Making*, 8(4), 482–491.
- Geurts, B. (2013). Alternatives in framing and decision making. *Mind and Language*, 28(1), 1–19.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics* (Speech acts) (Vol. 3, pp. 41–58). New York: Academic.

- Groves, R. M., & Singer, E. (2004). Survey methodology. In J. S. House, F. T. Juster, R. L. Kahn, H. Schuman, & E. Singer (Eds.), *A telescope on society: Survey research and social science at the University of Michigan and Beyond* (pp. 21–64). Ann Arbor: University of Michigan Press.
- Haack, S. (2006). Introduction. In S. Haack (Ed.), *Pragmatism, old and new. Selected writings* (pp. 15–69). New York: Prometheus Books.
- Hacking, I. (1983). *Representing and intervening. Introductory topics in the philosophy of natural science*. Cambridge: Cambridge University Press.
- Hausman, D. (2012). *Preference, value, choice, and welfare*. Cambridge: Cambridge University Press.
- Jiménez-Buedo, M. (2011). Conceptual tools for assessing experiments: Some well-entrenched confusions regarding the internal/external validity distinction. *Journal of Economic Methodology*, 18(3), 271–282.
- Jones, M. K. (2007). A Gricean analysis of understanding in economic experiments. *Journal of Economic Methodology*, 14(2), 167–185.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Kreiner, H., & Gamliel, E. (2016). Looking at both sides of the coin: Mixed representation moderates attribute framing bias in written and auditory messages. *Applied Cognitive Psychology*, 30, 332–340.
- Kreiner, H., & Gamliel, E. (2018). The role of attention in attribute framing. *Journal of Behavioral Decision Making*, 31, 392–401.
- Kuhn, T. S. [1962](1970). *The structure of scientific revolutions*. Chicago: The University of Chicago Press.
- Larrick, R. P., Smith, E. E., & Yates, J. F. (1992). *Reflecting on the reflection effect: Disrupting the effects of framing through thought*. Paper presented at the meetings of the Society for Judgment and Decision Making, St. Louis, November, MO.
- Laudan, L. (1981). A problem-solving approach to scientific progress. In I. Hacking (Ed.), *Scientific revolutions*. Oxford: Oxford University Press.
- Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Process*, 76(2), 149–188.
- Maheswaran, D., & Meyers-Levy, J. (1990). The influence of message framing and issue involvement. *Journal of Marketing Research*, 27(3), 361–367.
- Manski, C. F. (2004). Measuring expectations. *Econometrica*, 72(5), 1329–1376.
- Marteau, T. M. (1989). Framing of information: Its influence upon decisions of doctors and patients. *British Journal of Social Psychology*, 28(1), 89–94.
- Mayer, R. (1992). To win and lose: Linguistic aspects of prospect theory. *Language and Cognitive Processes*, 7(1), 23–66.
- Mayo, D. G. (1996). *Error and the growth of experimental knowledge*. Chicago: The University of Chicago Press.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). New York: American Council on Education/Macmillan Publishing Company.
- Millar, M. G., & Millar, K. U. (2000). Promoting safe driving behaviors: The influence of message framing and issue involvement. *Journal of Applied Social Psychology*, 30(4), 853–866.
- Moscatti, I. (2012). Intension, extension, and the model of belief and knowledge in economics. *Erasmus Journal for Philosophy and Economics*, 5(2), 1–26.
- Nardo, M. (2003). The quantification of qualitative survey data: A critical assessment. *Journal of Economic Surveys*, 17(5), 645–668.
- Pelham, B. W., & Blanton, H. (2003). *Conducting research in psychology. Measuring the weight of smoke*. Belmont: Wadsworth/Thomson Learning.
- Popper, K. R. ([1991]1999). All life is problem solving. Reprinted in: K. R. Popper. *All life is problem solving* (pp. 99–104). London: Routledge.
- Rescher, N. (1977). *Methodological pragmatism. A Systems-theoretic approach to the theory of knowledge*. New York: New York University Press.

- Russo, J. E., Medvec, V. H., & Meloy, M. G. (1996). The distortion of information during decisions. *Organizational Behavior and Human Decision Process*, 66(1), 102–110. 810
811
- Schoorman, F. D., Mayer, R. C., Douglas, C. A., & Hetrick, C. T. (1994). Escalation of commitment and the framing effect: An empirical investigation. *Journal of Applied Social Psychology*, 24(6), 509–528. 812
814
- Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods, and the logic of conversation*. Hillsdale, NJ: Erlbaum. 815
816
- Searle, J. (1978). Literal meaning. *Erkenntnis*, 13, 207–224. 817
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin. 818
819
- Sireci, S. G. (2009). Packing and unpacking sources of validity evidence: History repeats itself again. In R. W. Lissitz (Ed.), *The concept of validity: Revisions, new directions, and applications* (pp. 19–37). Charlotte: Information Age Publishing. 820
821
822
- Smith, S. M., & Levin, I. (1996). Need for cognition and choice framing effects. *Journal of Behavioral Decision Making*, 9(4), 283–290. 823
824
- Starr, M. A. (2014). Qualitative and mixed-methods research in economics: Surprising growth, promising future. *Journal of Economic Surveys*, 28(2), 238–264. 825
826
- Steinle, F. (1997). Entering new fields: Exploratory uses of experimentation. *Philosophy of Science*, 64(supplement), s65–s74. 827
828
- Suppes, P. (1998). Pragmatism in physics. In P. Weingartner, G. Schurz, & G. Dorn (Eds.), *The role of pragmatics in contemporary philosophy* (pp. 236–253). Vienna: Hölder-Pichler-Tempsky. 829
830
- Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, 110, 67–85. 831
832
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458. 833
834
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 107, 1039–1061. 835
836
- Wiener, P. P. (1973–1974). Pragmatism. In P. P. Wiener (Ed.), *The dictionary of the history of ideas: Studies of selected pivotal ideas* (Vol. 3, pp. 551–570). New York: Charles Scribner's Sons. 837
838