

ORIGINAL RESEARCH ARTICLE

Experimentally Validated Surveys: Potential for Studying Cognitive and Behavioral Issues in Management

Paper accepted by former associate editor Alexandra Gerbasi

Anne-Gaëlle Figureau^{1,2}, Anaïs Hamelin^{3*}, and Marie Pfiffelmann³

¹HYDREOS, Tomblaine, France

²UMR G-EAU, Montpellier, France

³LaRGE, EM Strasbourg Business School, Université de Strasbourg, Strasbourg, France

Abstract

In management, addressing behavioral and cognitive issues empirically is particularly challenging, as it requires combining both internal and external validity with multilevel data collection. The traditional quantitative methodologies used in this area, which are surveys and laboratory experiments, are ill-suited for this purpose. This paper presents a promising, innovative methodology, the experimentally validated survey (EVS), which allows researchers to face this challenge. The EVS relies on the implementation of survey questions previously validated by a controlled experiment. This paper adopts a threefold approach: descriptive, achieved by proposing a structured review of the recent body of literature; practical, achieved by discussing technical issues that one may encounter while running an EVS; and critical, achieved by discussing the advantages and limitations of such methodology for management scholars. Our main contribution is initiating a dynamic movement toward the importation of this innovative methodology into management research.

Keywords: *Methodology; Validity; Experiments; Surveys; Behavior*

Handling Editor: Alexandra Gerbasi; Received: 5 December 2019; Accepted: 31 January 2020; Published: 16 December 2020

Have you heard about the latest reality TV adaptation of *The Truman Show*, in which real managers are unknowingly encased in an artificial world entirely controlled from the outside and permanently observed? Of course, you have not. Fortunately, it does not exist.¹ However, management researchers may sometimes envy biologists and entertain dark fantasies of putting managers or entrepreneurs under a glass

case to observe, measure, and ultimately understand how their behaviors and cognitive processes impact firms' outcomes. However, since this cannot be done, what is left for management researchers?

This methodological issue is of particular interest in the microfoundations movement, which aims to link individual processes to actual firms' outcomes (Abell, Felin, & Foss, 2008). More specifically, in microfoundations research, there is a strong interest in behavioral and cognitive issues. These considerations are not new and go back to Simon's (1947) early work on administrative behavior, which was concerned with how individual decision-making and motivation affect the organizational performance. Recently, there has been renewed interest in the behavioral foundations of organizations and decision-making (Felin, Foss, & Ployhart, 2015; Phan & Wright, 2018). Behavioral and cognitive theories and concepts are more and more systematically mobilized in diverse fields of the management literature, including in public policy with the

¹ There are multiple examples of managerial research experiments. One of the most famous attempts was the Hawthorne experiments conducted by Elton Mayo in the 1920s (Gillespie, 1991). More recently, 'lab-in-the-field' has built upon this tradition by implementing experimental games onsite with actual managers (Gneezy & Imas, 2017). Another approach consists of observing managers in their natural environment, as realized by Mintzberg (1973). However, those attempts are either very limited in terms of the timeline and scope of the observations or are always conducted in naturalistic environments where observations are noised by factors that the researcher cannot control from the outside. Indeed, conducting a *Truman Show* type of managerial experiment is neither feasible nor desirable from an ethical perspective.

*Corresponding author: Anaïs Hamelin, Email: anaïs.hamelin@unistra.fr

concept of 'nudges' (Thaler & Sunstein, 2008), organizational behavior scholarship (Moore & Flynn, 2008), strategy (Laroche & Nioche, 2015; Powell, Lovallo, & Fox, 2011), and entrepreneurship (Busenitz & Barney, 1997; Grazzini & Boissin, 2013; Phan & Wright, 2018).

However, empirically exploring the behavioral microfoundations of management² is challenging. First, when focusing on behavioral and cognitive aspects, the researcher is confronted with the nonobservability of some behavioral and cognitive characteristics, which can only be either revealed or self-reported. The choice between one and the other involves a tradeoff between the internal and external validity of the empirical constructs. Second, it requires the researcher to adopt a multilevel analysis. This implies that during the same data collection process, data on the individual and on the firm or organization level must be collected.

To collect data for exploring how behavioral and cognitive aspects affect organization outcomes, researchers essentially rely on survey data. For example, empirical papers on entrepreneurial cognition rely predominantly on surveys (58%), secondary data analyses (34%), and interviews and case studies (15%), while only 8% use experimental designs (Bird, Schjoedt, & Baum, 2012). This implies that "in most cases, the behaviors are self-reported and are broad and unspecific in nature" (Bird & Schjoedt, 2009, p. 334). In turn, this raises the issue of the appropriate measures by which to establish the validity and reliability of instruments used to measure behavioral or cognitive characteristics (Aguinis & Lawal, 2012; Bird et al., 2012).

One way to address this issue is to investigate the potential contributions of experimental economics, as done by psychology and behavioral economics scholars (Hisrich, Langan-Fox, & Grant, 2007). Indeed, within an experiment, preferences or behaviors are not self-reported by the subjects but are revealed through an incentivized choice. Experimental economics provide results with strong internal validity and reliability. However, despite its potential for application, this methodology is still underused in the management field (3% in Chandler & Lyon, 2001; 8% in Bird et al., 2012, 10.7% in Grégoire, Binder, & Rauch, 2019). This is mainly because experimental economics also suffers from limitations such as limited external validity, high cost, and limited sample sizes. In the specific context of empirically exploring the behavioral microfoundations of management, laboratory experiments do not allow for a multilevel analysis, as they are uncondusive to linking individual preference with actual managerial actions or outcomes within the same data collection.

² There are also qualitative approaches to address microfoundational issues. See, for example, Vo, Culié, and Mounoud (2016). However, as this paper focuses on quantitative methods, the reader interested in the issue of qualitative methods can find more information in the review paper on microfoundational research by Felin et al. (2015).

To overcome the challenges raised by the empirical exploration of the behavioral microfoundations of management, this paper presents and reviews a new methodology by which to elicit individual preferences while collecting information at the organizational level. This alternative methodology proposes using survey questions previously validated by an economic experiment to ensure their reliability (refereed here as EVS – for experimentally validated survey). Indeed, a growing body of experimental literature has shown that a well-chosen direct question validated by a prior experiment may be able to capture individual preferences as reliably as experimental economics (Ding, Hartog, & Sun, 2010; Dohmen et al., 2011; Falk, Becker, Dohmen, Huffman, & Sunde, 2016; Hardeweg, Menkhoff, & Waibel, 2013; Johnson & Mislin, 2012; Vieider et al., 2015; Vischer et al., 2013). Therefore, EVS might allow us to overcome empirical challenges in the exploration of the behavioral microfoundations of management. On the one hand, validation of the questions by an experiment ensures the internal validity of the behavioral constructs. On the other hand, implementation of a large-scale survey to the relevant population in a naturalistic environment ensures external validity. Moreover, EVS allows collecting data on the individual and organizational levels, such as the firm performance, during the same data collection process.

This paper presents this innovative methodology. In the process, it expands upon its main advantages and achievements (i.e., the questions that have already been validated and are 'ready for use'), gives some practical recommendations, and discusses specific implementation issues when running an EVS. Our main contribution is initiating a dynamic movement toward the importation of an innovative methodology into the research field of behavioral and cognitive approaches in management. In doing so, we respond to the call of management scholars, underlining that "experiments do not have to be employed alone" and that "using experiments in conjunction with other methods can be especially fruitful endeavor" (Williams, Wood, Mitchell, & Urbig, 2019, p. 217). The paper adopts a threefold approach: descriptive, achieved through a structured review of the recent body of literature; practical, achieved by discussing technical issues that one may encounter while running an EVS; and critical, achieved by discussing the advantages and limitations of the methodology for management scholars.

This paper is organized as follows. The 'Lab experiment or survey?' section compares the advantages and limitations of the two main quantitative methodologies: surveys and lab experiments. The 'EVS and hybrid methodologies' section presents the EVS's advantages and positions the EVS among the other hybrid methodologies used in management. The 'Experimental validation of survey questionnaires: a review' section reviews the literature on the EVS. The 'Experimental validation of a survey question: a practical guide' section tackles practical issues that one may face in the process of

experimentally validating a survey question. Finally, the 'Discussion and conclusion' section discusses the potential for use of the EVS in management studies and concludes.

Lab experiment or survey?

As they face the choice between surveys and lab experiments, researchers in management have to weigh the pros and cons of both methods. Survey techniques meet several of the managerial field's research requirements. They can be widely implemented on large samples at the national or even international level, and they yield massive amounts of data at a limited cost. Web-based and email questionnaires have the advantage of being quick and easy to implement. Surveys based on open questions allow for gathering subjective data, such as the preferences of large representative samples, thus enhancing the external validity. Nevertheless, this method is subject to several response biases (Zikmund, Babin, Carr, & Griffin, 2013). For instance, the hypothetical bias occurs when respondents answer questions in a way that differs from how they would have behaved in real life, expressly because a stated behavior would not have any real implication or consequence (Bohm, 1972). The social desirability bias – the tendency for respondents to answer questions in such a way that others, and particularly the interviewer, may view them favorably – may be the most studied response bias (for structural scholars, see Crowne & Marlowe, 1960; Paulhus, 1991).³ These biases are problematic, as they affect the way in which responses are provided and create systematic measurement errors (Lavrakas, 2008). Response biases impede the identification of the origin of variations of the explained variable, thus affecting the internal validity (Kirk, 1995).

An economic experiment consists of putting subjects (typically, a student population) in an artificial economic situation, in which every parameter is set and controlled by the experimentalist and wherein subjects have to make choices and decisions that are all observed and measured by the experimentalist. Experiments are, by definition, characterized by high control of the conditions and parameters of data production (Fehr et al., 2002). In fact, in order to reduce noise, context and background are removed from the experiments. This allows not only for making accurate and reliable measurements but also for performing diverse parameter manipulations (always one by one) that would be impossible in a real-life setting (Charness, 2010). The subsequent observations enable researchers to establish solid causality between external signals and individual behaviors (Charness, Gneezy, & Kuhn, 2013).

³ Many other response biases have been reported in the literature (acquiescence bias, overreporting bias, confusion regarding word meanings, strategic response bias, and retrospective bias).

In other words, experiments are characterized by high internal validity. Another significant advantage is the replicability of experiments. Indeed, laboratory experiments provide a platform that can be used by a wide range of researchers (Charness, 2010). The latter can test previous results in order to verify their robustness, manipulate an existing parameter in order to further explore its effects, or add a parameter in order to observe its additional effects (Fehr et al., 2002). However, the large-scale implementation of this method is extremely difficult, if not impossible. The major criticism against experiments concerns their external validity (Charness, 2010; Fehr et al., 2002; Harrison & List, 2004). Indeed, since results are produced in an artificial environment, they may be barely generalizable to the real world. A central question concerns which populations, settings, treatment, or measurement variables for which a lab-produced experimental effect can be generalized (Campbell & Stanley, 1966). Additionally, lab experiments involve high costs (game programming, equipment, and subject remuneration). Therefore, they are only suitable for small samples and are often made up of students, which raise the issue of sample representativeness and, more broadly, of external validity (Barbosa, Fayolle, & Smith, 2019).

Thus, in terms of the validity, surveys and lab experiments display opposite characteristics. Survey results have a higher external validity than experimental results, meaning that elicited preferences may be closer to reality and more robustly generalized. Inversely, the strict control of the data production process in experiments ensures higher internal validity than exists with survey data, making measurements of preferences and inferences of causal links more reliable. However, as mentioned earlier, the artificial environment of the experimental game may bring subjects far from the real environment, in which they are used to make decisions. Therefore, the results obtained in the lab may not be generalizable to reality (low external validity). Further, in terms of cost relative to the amount of data collected, surveys are much more advantageous, as they allow for collecting massive quantities of data from large samples, whereas experiments incur high implementation costs for gathering little information on small samples.

EVS and hybrid methodologies

Emerging literature in economics (Dohmen, Falk, Huffman, & Sunde, 2005; Dohmen et al., 2011; Falk et al., 2016) has proposed an innovative methodology, the EVS, to overcome the limits of lab experiments and surveys. This methodology relies on the implementation of survey questions previously validated by a lab experiment. On the one hand, the validation of the survey questions by a controlled experiment ensures the internal validity of the results. On the other hand, the implementation of the validated survey on large samples in a

naturalistic environment at the national or even international level enhances the external validity of the results. This methodology therefore combines the advantages of both lab experiments and surveys.

However, the EVS technique is not the only alternative to surveys and lab experiments proposed in the literature. In fact, as far as experiments are concerned, a wide range of hybrid methodologies has been used in management and economics. Field experiments are the most well-known alternative to laboratory experiments (Huber, Sloof, & Van Praag, 2014; Sadoff & Samek, 2019; Song, Liu, Lanaj, Johnson, & Shi, 2018). These experiments are conducted in naturalistic environments and use nonstudent populations. Often, the subjects of the experiments are not aware that their decisions are being studied. The main advantage of field experiments relates to the external validity of the results. In fact, as the experimentalist targets a specific population in its natural environment, the result will naturally be more applicable to the context studied. However, with field experiments, it is not possible to control the variables as closely as with lab experiments. Thus, field experiments yield a lower internal validity. To overcome this limitation, researchers (Gneezy & Imas, 2017; Harrison & List, 2004) have proposed a new experimental methodology called 'lab in the field', which combines elements of both lab and field experiments. The aim of 'lab in the field' is to provide a methodology that maximizes the pros and reduces the cons of lab and field experiments. This methodology consists of conducting a standardized validated lab experiment in a naturalistic environment using a population related to the study. Resorting to a validated lab experiment ensures tight control and the internal validity of the results. Targeting the relevant population and implementing the experiment in a naturalistic environment enhance the external validity. However, as with lab experiments, this methodology involves high costs (game programming, equipment, and subject remuneration) and is therefore only suitable for small samples. To implement experiments using large samples, researchers may turn to web-based experiments. For instance, Graf, König, Enders, and Hungenberg (2012) analyze how managers can reduce competitive irrationality in a sample of 934 managers using web-based experiments. In that case, the experiment is run in the form of a survey administered over the Internet. The large number of participants improves the external validity, as it can provide the "realism of field data" aspect (Falk & Fehr, 2003, p. 403). However, web-based experiments "are somewhat limited in their possibilities to control for the environment in which research participants must make their decisions" (Graf et al., 2012, p. 393) and therefore provide results with lower internal validity. Overall, the advantage of EVS in comparison with these methodologies is that it achieves a good balance between internal and external validity and implementation on large-scale samples

at a reasonable cost. This is of particular importance for empirically exploring the behavioral microfoundations of management. Indeed, it requires combining internal and external validity of the measures, but a multilevel analysis is also required. To link individual preferences with actual managerial actions or outcomes, the researcher should, within the same data collection process, collect data on the individual level and on the firm or organization level. With the EVS, the behavioral results of the surveys can be linked to organizational-level variables, such as the firm performance, whereas the limited sample size of lab experiments and 'lab in the field' results does not make this multilevel analysis possible. With the implementation of a traditional survey or a web-based experiment, linking micro- and macro-variables is possible, but the internal validity of the behavioral results collected with these methodologies is lower than with EVS.

Some researchers also propose combining different methodologies and using experiments to complement traditional management sciences methodologies such as observational research or surveys (Croson, Anand, & Agarwal, 2007; Zellmer-Bruhn, Caligiuri, & Thomas, 2016). Emerging literature (Barsade, 2002; Bekir, El Harbi, Grolleau, & Sutan, 2015; Chang, Lusk, & Norwood, 2009; Coppola, 2014; Hainmueller, Hangartner, & Yamamoto, 2015) aims at comparing questionnaire and experimental methodologies. For instance, Barsade (2002) measures emotional contagion using both participants' self-reports and observers' ratings of mood via video-tape ratings of the participants interacting in a group exercise. Bekir et al. (2015) investigate what individuals maximize: efficiency, equality, or positionality. To do so, they compare results obtained with incentive compatible choices (experiment) with results obtained by hypothetical surveys. Coppola (2014) compares three methods of risk preference elicitation (lottery-choice tasks, Domain-Specific Risk-Taking scale, and general multi-item questionnaire) and then compares the results to real respondents' behaviors in real life (such as having risky assets, smoking, and practicing risky sports). These studies use a combination of several methods, for example, experimental and self-reported measures, to triangulate or to compare the results. The aim of the EVS is neither to obtain two measures of the same factor nor to compare their differences. The EVS proposes experimentally validating survey questions through an economic experiment to ensure their reliability. Once the questions are validated, it is not necessary to perform a new experiment. Instead, the survey questions can be directly administered to a large sample by targeting a relevant population. Then, researchers no longer need to combine surveys and experiments. The internal validity of the results is ensured by the controlled experience, while the external validity is guaranteed by the implementation of the large-scale survey to the relevant population in a naturalistic environment.

Table 1. List of experimentally validated measures in the literature

Preferences tested	References	Experiments used for validation	Surveys' measures	Validation methods
Altruism	Falk et al. (2016)	Dictator game (1 st player)	Scenario transposed from an experiment + Self-assessment questions	Descriptive (correlations), Measure of predictive validity (regression analysis) Out of sample validity
Positive Reciprocity	Falk et al. (2016)	Investment game (2 nd player)	Scenario transposed from an experiment + Self-assessment questions	Descriptive (correlations), Measure of predictive validity (regression analysis) Out of sample validity
Negative Reciprocity	Falk et al. (2016)	Prisoner's dilemma or Ultimatum game	Scenario transposed from an experiment + Self-assessment questions	Descriptive (correlations), Measure of predictive validity (regression analysis) Out of sample validity
Risk aversion	Dohmen et al. (2011)	Lottery choices (Holt and Laury, 2002)	Self-assessment questions	Measure of predictive validity (regression analysis)
	Ding et al. (2010)	Lottery choices	Scenario transposed from an experiment + Self-assessment questions	Descriptive (correlations, comparison of distributions)
	Hardeweg et al. (2013)	Lottery choices	Self-assessment questions	Measure of predictive validity (regression analysis)
	Vieider et al. (2015)	Lottery choices	Self-assessment questions	Descriptive (correlations, comparison of distributions)
	Falk et al. (2016)	Lottery choices	Scenario transposed from an experiment + Self-assessment questions	Descriptive (correlations), Measure of predictive validity (regression analysis) Out of sample validity
Time preference: discounting	Vischer et al. (2013)	Time preference experiment	Self-assessment questions	Descriptive (comparison of distributions) Measure of predictive validity (regression analysis)
	Falk et al. (2016)	Time preference experiment	Scenario transposed from an experiment + Self-assessment questions	Descriptive (correlations), Measure of predictive validity (regression analysis) Out of sample validity
Trust	Johnson and Mislin (2012)	Investment game (1 st player)	Self-assessment questions	Measure of predictive validity (regression analysis)
	Falk et al. (2016)	Investment game (1 st player)	Scenario transposed from an experiment + Self-assessment questions	Descriptive (correlations), Measure of predictive validity (regression analysis) Out of sample validity

Experimental validation of survey questionnaires: A review

The EVS technique has been recently developed and discussed by a small number of scholars, mostly in the 2010s. Only eight papers on the subject have been identified,⁴ all of which are recent, reflecting the EVS's status as an emerging methodology. In a 2002 paper, experimentalists Fehr et al. began reflecting on how to transpose a robust and tried-and-tested economic experiment (the trust game) into a survey in order to implement it more broadly across larger samples and collect massive amounts of data. However, their aim was not to strictly compare experiment- and survey-elicited preferences in order to validate the latter using the former but was to adapt a classic economic experiment (the trust game) to the form of questions suitable for a survey. This was *not* an EVS, but their research question attests to the need for experimentalists to find ways to enlarge the samples based on which their experimental measures were implemented in the early 2000s.

⁴ A work in progress by Bauer & Chytilova has been attempting to experimentally validate a survey module to measure economic preferences in Kenya, but the results and protocol are not available yet. The preanalysis is available at <https://osf.io/pvf54/>.

The EVS technique genuinely emerged in the early 2010s, with scholars showing interest in validating survey questionnaires through lab experiments (Ding et al., 2010; Dohmen et al., 2005, 2011; Falk et al., 2016; Hardeweg et al., 2013; Johnson & Mislin, 2012; Vieider et al., 2015; Vischer et al., 2013). These studies are all from the economic field and focus on the measurement of economic preferences (risk aversion, time preference, trust, etc.) that are already reliably elicited by tried-and-tested experiments. These studies attempt to validate two types of surveys. In most cases, the validated survey consists of a series of self-assessment questions (asking participants to position themselves in relation to an assertion of the type 'generally, do you consider yourself a person ...'). In a few cases, the validated survey consists of self-assessment questions accompanied by scenarios that transpose an experiment into a questionnaire.

These eight studies (whose experimentally validated measures are listed in Table 1) all rely on the experimental validation of survey questions but are not constructed with the same research objective. The first category of studies (Ding et al., 2010; Dohmen et al., 2005; Hardeweg et al., 2013; Vieider et al., 2015) focuses on measuring risk aversion. First, Ding et al. (2010) and Dohmen et al. (2005) paved the way by exploring the possibility of measuring risk aversion using an

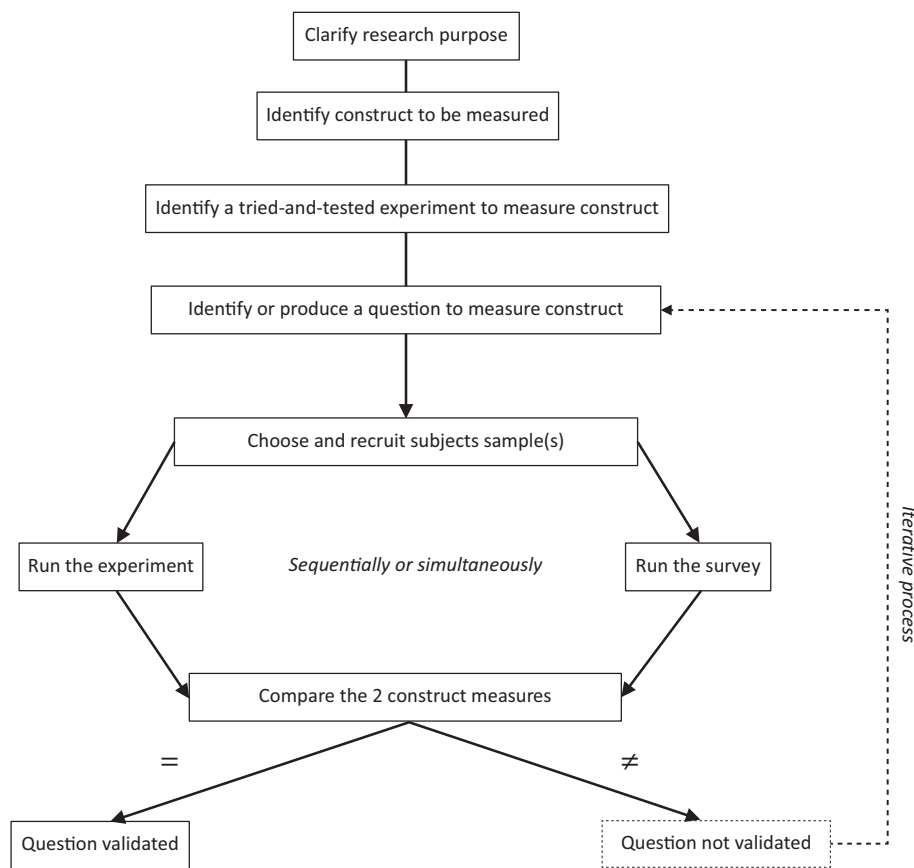


Figure 1. Work stages for the experimental validation of a survey question

EVS question. Their research seeks to determine which among a pool of questions (*ad hoc* or taken from other existing surveys) best reflect the risk attitude of the participants. Then, Hardeweg et al. (2013) experimentally validated survey questions using a particular target sample (rural households in a Thai region) in order to benefit from an EVS on risk aversion for future research with the same sample. Subsequently, Vieider et al. (2015) conducted a large-scale study in collaboration with universities in 30 different countries to compare risk aversion measurements from an experiment and a survey. They aimed at establishing whether there is a correlation between these measures (surveys vs. experiments) within and between countries in order to be able to use the same EVS in different countries with different cultures in the long run. The second category of studies (Johnson & Mislin, 2012; Vischer et al., 2013) aims at experimentally validating questions from existing national or international surveys such as the German Socio-Economic Panel (GSOEP) or the World Values Survey (WVS). Johnson and Mislin (2012) focus on trust, whereas Vischer et al. (2013) investigate time preferences. In both cases, the authors neither carried out the

experimental study nor administered the questionnaire but cross-analyzed data from past experiments with the responses collected by the GSOEP or the WVS. The third and last category of studies has a research-improving goal, aiming to propose questions, either *ad hoc* or from existing surveys, or scenarios that are as reliable as the experiments, so that the academic world can benefit from a set of reference questions that can be reused later (Dohmen et al., 2011; Falk et al., 2016). The work of Falk et al. (2016) seeks to produce what they call an experimentally validated 'survey module', measuring six different types of preferences. The use of a standard module of questions by the entire academic community will improve the comparability of studies.

Experimental validation of a survey question: A practical guide

Based on the literature on the EVS reviewed in the previous section, we represent in Figure 1 the work stages for experimentally validating a survey question. Figure 1 derives from a qualitative synthesis of the methodological and practical issues

found in our corpus of articles. First, we conducted a vertical analysis to extract the structure of the methodological part of each paper. Second, we met and compared our views, adopting a horizontal approach for the papers. Thanks to this transversal reading, we identified recurrences and regularities in the methodological and practical issues of the EVS papers. Third, we read each paper individually to see if our general structure was coherent with the elements present in our corpus of papers and make some minor adaptations. Overall, Figure 1 and the discussion later synthetically present the different steps in the experimental validation of a survey question, as it is most commonly executed in this emergent literature. Of course, in some cases, the names of the steps might differ, as some authors do not follow every step in the same order or use the same name.

The first step is to define the purpose of the research. In fact, different research purposes may bring researchers to resort to the experimental validation of a survey question, such as testing the validity of an existing and already widely used survey question or producing and testing several new questions and comparing their validity in order to select one for large-scale implementation. The concept to be measured has to be clearly and properly defined and distinguished from other related concepts in order to avoid confusion. Then, researchers have to identify the experiment and the questions which will be used to measure the construct. As the experiment is used to produce measures that will be considered as the reference for validation, it has to be a tried and tested experiment, with results whose robustness is unchallenged. The question used may be an existing question or an *ad hoc* question developed for specific research purposes. Last, we need to assess to what extent the two constructs, experiments and survey questions, are comparable through statistical analyses. If the survey measurement appears to be highly correlated with or predictive of the experimental measurement, the survey question is considered validated; otherwise, it is not, and an iterative process may begin, going back to the question identification stage and requiring reformulation.

However, at each of these steps, one may encounter difficulties and have to make methodological choices that will impact the final results. The main methodological issues in the experimental validation of survey questions are the recall effect, the choice of the subject pool, the order effect, and the sample size.⁵ The issue of the recall effect arises when associating an experiment with a survey questionnaire that aims at measuring the same preference or cognitive bias. Since questions are direct and explicit (e.g., asking people to rate their agreement with the statement 'I like taking risks'), people may understand

the purpose of the subsequent experiment and deliberately manipulate their answers. Such explicit framing is likely to raise response biases, in particular, social desirability and acquiescence biases. Thus, when combining an experiment with a question, there may be an influencing effect (recall effect) between the experiment and question. To control for the recall effect, one may choose among four possible solutions. The first is to have the experiment and question separated in time (see, e.g., Vischer et al., 2013 with a time interval of 2 years). A second solution is to design a 'between' experiment, which consists of changing subjects from the experiment session to the survey session in order to avoid a mutual influence effect between treatments. However, this requires controlling for external variables, such as age, gender, etc., unless one is working with a very large sample, which is considered representative of a population (Johnson & Mislin, 2012). This solution is implicitly used in most studies, for which researchers do not effectively manage the entire experimental validation process (in most cases, they validate a survey originally administered by a third party, such as WVS or SOEP surveys). A third solution that is only usable when studying several cognitive biases or preferences consists of combining questions and experiments on different preferences into the same session, with each session consisting of a question and an experiment that are not related to the same preferences (Falk et al., 2016). Finally, a fourth solution is to include many questions on different themes in the survey, so that the participants do not make the connection between the experiment and the objective of the questionnaire (Dohmen et al., 2011).

Another methodological issue in the experimental validation process is the order effect. The order effect arises when applying different treatments to the same pool of subjects. Subjects' behavior in the last treatment may be influenced by the first treatment, resulting in a lack of independence between the results. Such an issue emerges in the experimental validation process, since we apply two 'treatments' to the subjects (experiment and questionnaire) that may interact and bias the results. To control for the order effect, there are two possible solutions. The first method consists of adopting a *between* design using two distinct subject pools, with one for each of the experimental validation stages, the survey and the experiment. Implementing this method requires having two samples that are large enough to be considered representative of the whole population and thus are both similar or controlling for variables such as age, gender, etc. (see, e.g., Johnson & Mislin, 2012; Vischer et al., 2013). A second way to control the order effect is to adopt a *within* design and apply a counterbalancing technique that consists of reversing the order of the experimental and survey elicitation of preferences for half of the subjects. Half of the participants start with the experiment and then complete the survey, whereas the other half start with the survey and then participate in the experiment. If

⁵ Only practical issues specific to the experimental validation of survey questions are addressed here. Methodological issues common to classical experimental approaches are outside the scope of this discussion.

no difference is found in the final results between the two types of sessions, the order effect can be ruled out (see Falk et al., 2016 for an example).

A further methodological issue relates to the choice of sample subjects between a standard or specific population. When studying a specific population, in our case, entrepreneurs or managers, one might question the use of a standard population (in most cases, students) for the experimental validation of a survey question. While using a standard population has several advantages (ease of recruitment, ease of incentive, etc.), it alters the external validity. Some scholars call for using a pool of subjects from the population targeted for the survey to enhance external validity of the EVS, since results may differ among populations (Dohmen et al., 2011; Vischer et al., 2013). As targeting a specific population of managers or entrepreneurs might be difficult, one could take advantage of the development of eLancing for experimental validation for nonstudent participants.⁶ Overall, most EVS studies use standard subjects, who are students, and assume that the type of subject will not impact the comparison between survey-produced and experimentally produced data, provided that they stem from the same individual. Regarding this, Falk et al. (2016, p. 17) argued that “while the distributions of preferences may differ for students and non students, there is no particular reason to think that the correlation structure should differ”.

The last methodological issue pertains to the optimal sample size, which raises two issues. First, the sample size depends on the experimental design. More specifically, it depends on the number of cells of the experiment, which is determined by the number of treatments. In the case of the experimental validation of a survey question, treatments are related to the controls for recall and order effects. For example, if you control for the order effect by inverting the order between the experiment and survey questions, then the experimental design includes two cells. Second, there is the issue of the optimal subsample by cell. The literature on this issue is strikingly consistent, and most studies uniformly distribute at least 30 subjects into each cell (List, Sadoff, & Wagner, 2011). The EVS literature reviewed in this paper generally relies on samples that include 50–100 subjects per cell. However, relying on exceedingly large samples for experimental designs might be very costly. List et al. (2011) provide a more precise method, by which to calculate the optimal cell sample size, taking into account the structure of the experiment and the required significance and power.

The last step in the experimental validation process consists of assessing to what extent the two constructs, the

experiment and survey questions, are comparable. Different methods can be used to investigate whether the survey measurement can predict the actual behavior in the incentive-compatible experiment. Indeed, if it does, this justifies the use of the survey measure as a relevant proxy for the experimental construct in future surveys.⁷ In our review, we have identified three methods: descriptive methodologies, measures of predictive validity, and tests of out-of-sample validity.

First, descriptive methodologies rely on assessing the consistency of measurements. To do so, the researcher studies the significance of correlations between experimental and survey constructs by looking at whether the signs are going in the expected directions and whether the extent of the correlation across measures represents a meaningful pattern (Ding et al., 2010; Falk et al., 2016; Vieider et al., 2015). Alternatively or complementarily, some authors rely on the comparison of the distributions of the survey and experimental construct and on statistical assessments of distribution differences – for example, the Kolmogorov–Smirnov test (Ding et al., 2010; Vieider et al., 2015; Vischer et al., 2013).

Second, it is possible to compute measures of predictive validity. One measure is the mean square error (MSE) (Chang et al., 2009). The MSE is the mean of the squared difference between the reference value (experiment result) and the predictive value (survey results). The coupled experiment-survey question with the lowest MSE is deemed to have the best predictive performance. Another way to assess the predictive validity is to rely on regression analysis. With the latter method, the explained variable is the experimental measure of the construct, and the explicative variable is the survey measure. Then, after running a regression, we focus on the regression coefficients to assess the predictive validity of the survey measure. More specifically, we assess the survey question's predictive validity by looking at the significance of the coefficients and at whether signs point to the expected direction. The advantage of this more complex method is that it offers the possibility of introducing control variables, thus excluding the possibility of the result being affected by unobserved heterogeneity (Dohmen et al., 2011; Hardeweg et al., 2013; Johnson & Mislin, 2012; Vischer et al., 2013). Moreover, it allows for classifying candidate survey questions according to their predictive validity (Falk et al., 2016).

Finally, it is possible to test the predictive validity by estimating the out-of-sample validity. This can be done using two samples. The first sample is used to estimate a model of the predictive power of the explicative setup (survey question). Then, the predictive measurements are computed for the

⁶ However, one of the main limitations of eLancing is the selection issue. Indeed, in the case of the experimental validation of survey questions targeted toward management research, it might be difficult, if not impossible, to be sure that the eLancers are actual managers. In the case of entrepreneurs, this might be less of an issue.

⁷ In an experimental validation of a survey question, the survey and the experiment are conducted simultaneously; thus, the same sample is used. However, once the survey questions are experimentally validated, the survey questions can be implemented on a larger sample.

other sample and compared to the actual measurements. Stated differently, the authors use the subjects' survey responses to predict their choices in the experiment (based on the regression model previously estimated based on the first sample) and then regress the actual choices onto the predicted choices. If the survey reliably captures the preferences of individuals in this second sample, one would expect the intercept of the regression of the actual choices onto predicted choices to be zero and the coefficient of the predicted value to be one; thus, one must test for these hypotheses (Falk et al., 2016).

Discussion and conclusion

This paper has presented the innovative EVS methodology, which relies on the implementation of survey questions previously validated by a controlled experiment. First, the EVS ensures the survey questions' internal validity. Second, its implementation on the relevant population in a naturalistic environment ensures the external validity and multilevel data collection.

The EVS is of potential interest for researchers in management, particularly those interested in the behavioral and cognitive microfoundations of management. More specifically, Phan and Wright (2018, p. 179) underline "that cognition and behavior are at the core of management research. Research at the individual, organization, and system levels of analysis ultimately starts from theories of why and how individuals make decisions to compete or cooperate to achieve their goals". To investigate these questions, it is necessary to adopt a multilevel analysis and combine individual-level variables and firm-outcome variables within the same data collection process. The EVS offers a response to several empirical methodological challenges posed by the use of standard methodologies, which are surveys or experiments, to explore preferences (Aguinis & Lawal, 2012; Bird & Schjoedt, 2009; Bird et al., 2012; Chandler & Lyon, 2001; Charness, 2010; Fehr et al., 2002; Harrison & List, 2004; Hisrich et al., 2007; Zikmund et al., 2013). With the EVS, confidence regarding the causality (i.e., internal validity) is ensured by the experimental validation, while the generalizability and multilevel data collection are guaranteed by the large-scale implementation of the survey, ensuring sample representativeness. This is a distinguishing feature of the EVS compared to other validation techniques used in management. Indeed, these validation techniques rely on the use of complementary analyses that mostly add external validity to the existing measurement instrument (through knowledge on the context, the use of a first sample of subjects, etc.). However, these complementary analyses add less internal validity than laboratory experiments. For instance, marketing researchers often rely on Churchill's paradigm (Churchill, 1979). This validation technique is based on a confirmatory factor analysis to observe how the

model works outside of the sample. Another alternative is the mixed method, which consists of combining quantitative and qualitative methods (Tashakkori & Teddlie, 2003). Indeed, the implementation of qualitative approaches allows for a better understanding of the context in which the respondents make decisions and is accordingly conducive to a better design of the quantitative questionnaires by identifying appropriate, precise, and valid variables and measuring instruments (Molina-Azorin, Lopez-Gamero, Pereira-Moliner, & Pertusa-Ortega, 2012).

The EVS appears to be a promising methodology that allows for testing whether the theoretical behavioral and cognitive drivers proposed by the literature to explain firm performances are correct. A specific issue that can be addressed thanks to the EVS is the 'how?' question of the decision process. Psychological research on judgment and decision-making has largely suggested that only a minority of individuals actually structure their decision processes according to the normative suggestions of utility theory (Goldstein & Hogarth, 1997; Kahneman, 2003; Kahneman, Slovic, & Tversky, 1982; Simon, 1955, 1979). Indeed, the way in which information is ordered, framed, and organized considerably impacts the final decision (Hogarth, 1987; McNeil, Pauker, Sox, & Tversky, 1982). Research on entrepreneurial cognition suggests that this is even more pronounced for entrepreneurs (Allinson, Chell, & Hayes, 2000; Busenitz, 1999; Busenitz & Barney, 1997; Mitchell et al., 2002, 2007; Shepherd, Williams, & Patzelt, 2015). Thus, the EVS can help us understand how managerial decisions are made, given that it allows for empirical exploration of how information affecting a decision is gathered, framed, and organized at the individual level and how this affects organizational decisions and outcomes. Furthermore, the EVS can facilitate managerial applications of research, given that testing whether we have the correct theory by which to explain a phenomenon is a prerequisite to potential intervention in the organization. Finally, the EVS has a high potential to facilitate the replication and comparison of results, since the ultimate aim of reviewed papers on the EVS was precisely to develop 'ready-to-use' questions designed to be suitable for rapid and easy replication by the research community, without questioning the robustness of the results (Falk et al., 2016). An interesting potential research avenue would be to integrate experimentally validated questions in international surveys such as the World Value Survey. This would allow us to investigate the intercultural differences of several populations or to implement the survey on a large representative sample of the target population.

Overall, the EVS appears to be a promising methodology for management researchers. But what remains to be done? At this point in time, the literature in experimental economics has already developed the EVS for three categories of preferences – risk aversion, time preferences, and trust – which can be used by the research community as ready-to-use questions.

Falk et al. (2016) produce what they call an experimentally validated 'survey module', measuring six different types of preferences: risk aversion, time preferences, trust, positive reciprocity, negative reciprocity, and altruism. For instance, researchers can use the validated self-assessment questions proposed by Vischer et al. (2013) to investigate the link between firm performance and the time preferences of managers, giving new insights into the issue of myopic management (Mizik, 2010; Stein, 1988, 1989). Further, the questions validated by Johnson and Mislin (2012) allow for exploring how trust affects firms' financial decisions, which can be of great interest in research on family business financial management. To the extent of our knowledge, only one article in the management field has explicitly used, in a survey, a question previously validated by an experiment. In their work, Caliendo, Fossen, and Kritikos (2009) employ the direct question on risk assessment validated by Dohmen et al. (2005) to measure risk aversion in a representative population sample to study its presumed influence on self-employment. They justified their choice precisely on the grounds that the question was previously validated and that it captured the very construct that they were going to measure.

However, in managerial decision-making, other cognitive factors, such as overconfidence, availability bias, escalation of commitment, preference for skewness, and self-serving attribution, are widely cited. Thus, further research is needed in order to either validate existing questions using experimental designs or build and validate a direct question instead of an experiment for wider implementation purposes. To do so, researchers in management can leverage the literature in behavioral economics and psychology, which has yielded plenty of tried-and-tested lab experiments. Once the questions are experimentally validated using the procedure presented in the 'Experimental validation of a survey question: a practical guide' section, it will be possible to implement the survey on a specific population and to make the links between these cognitive factors such as firm performance. For instance, Moore and Healy (2008) propose a controlled experiment on the three forms of overconfidence that could be used to experimentally validate direct questions measuring overconfidence. Integrating these EVS questions on overconfidence into a survey would allow researchers to explore issues such as the link between innovation and overconfidence. This question has already been explored (Galasso & Simcoe, 2011; Chen, Ho, & Ho, 2014) but with an indirect measurement of overconfidence (stock option exercise).

Although we are convinced that the EVS method is a very promising avenue in managerial research, it suffers from some limitations. First, the EVS method assumes that experimentally elicited preferences will be closer to subjects' real behaviors and thus can be considered a reference for evaluating the reliability of survey-produced data. However, why not rely directly on observational data to test for validity? In the empirical

literature on bias and decision-making, some studies relate survey questions with real-world data (Chang et al., 2009; Coppola, 2014; Hainmueller et al., 2015). This is an alternative method to the EVS for validating survey questions. This real-world-based validation method is particularly useful when trying to validate a behavior or preference measurement instrument. However, real-world data are difficult to collect, and sometimes some behaviors of interest might be unobservable. In this case, researchers need to choose between experimentally revealed preferences and survey-stated preferences to proxy for real-world behavior. Take, for instance, overconfidence. It is a preference that needs to be revealed. You cannot directly observe overconfidence, you can only observe the consequences of overconfidence. Second, another limitation lies in the fact that experimental techniques are often implemented on student samples, meaning that the external validity of experiments for nonstudent samples could be undermined. This raises the question of whether to conduct the underlying experiment on the EVS target population. According to Falk et al. (2016), the module will be behaviorally relevant for nonstudents, as long as the correlations between survey items and experiments are similar to those in the student sample. While the distributions of preferences may differ for students and nonstudents, there is no particular reason to think that the correlation structure should differ. Finally, the EVS might be difficult to implement on topics for which there are no tried-and-tested experiments. Indeed, the experimental economics community's validation of the experimental design used in the experimental validation of survey questions ensures the EVS's internal validity. Therefore, interesting managerial topics, such as the preference for independence and against external investors typically displayed by family firms, which have been largely pointed out in business history literature (Douglas & Shepherd, 2000; James, 2006), cannot be readily explored through EVS techniques because, to date, there is no recognized experimental design for this issue. Thus, further research in the EVS scope should also encompass more traditional experimental economics research, aiming at developing tried-and-tested experiments on specific managerial topics.

Acknowledgments and funding

This research benefited from the financial support of l'Université de Strasbourg and Investissements d'avenir. Editorial support was provided by the Maison Interuniversitaire des Sciences de l'Homme d'Alsace (MISHA) and the Excellence Initiative of the University of Strasbourg. The authors thank Nicolas Eber, Amélie Boutinot, Jérôme Hergeux, Sophie Michel, the two anonymous reviewers, and the editor for helpful comments. Anaïs Hamelin and Marie Pfiffelmann are members of the research network Recherche & Expertise en Entrepreneuriat Grand Est.

References

- Abell, P., Felin, T. & Foss, N. (2008). Building micro-foundations for the routines, capabilities, and performance links. *Managerial and Decision Economics*, 29(6), 489–502. doi: 10.1002/mde.1413
- Aguinis, H. & Lawal, S. O. (2012). Conducting field experiments using eLancing's natural environment. *Journal of Business Venturing*, 27(4), 493–505. doi: 10.1016/j.jbusvent.2012.01.002
- Allinson, C. W., Chell, E. & Hayes, J. (2000). Intuition and entrepreneurial behaviour. *European Journal of Work and Organizational Psychology*, 9(1), 31–43. doi: 10.1080/135943200398049
- Barbosa, S., Fayolle, A. & Smith, B. (2019). Biased and overconfident, unbiased but going for it: How framing and anchoring affect the decision to start a new venture. *Journal of Business Venturing*, 34(3), 528–557. doi: 10.1016/j.jbusvent.2018.12.006
- Barsade, S. G. (2002). The ripple effect: Emotional contagion and its influence on group behavior. *Administrative Science Quarterly*, 47(4), 644–675. doi: 10.2307/3094912
- Bekir, I., El Harbi, S., Grolleau, G. & Sutan, A. (2015). Efficiency, equality, positionality: What do people maximize? Experimental vs. hypothetical evidence from Tunisia. *Journal of Economic Psychology*, 47(C), 77–84. doi: 10.1016/j.joep.2015.01.007
- Bird, B. & Schjoedt, L. (2009). Entrepreneurial behavior: Its nature, scope, recent research, and agenda for future research. In A. L. Carsrud & M. Brännback (Eds.), *Understanding the entrepreneurial mind* (pp. 327–358). Springer. doi: 10.1007/978-1-4419-0443-0_15
- Bird, B., Schjoedt, L. & Baum, J. R. (2012). Editor's introduction. Entrepreneurs' behavior: Zlucidation and measurement. *Entrepreneurship, Theory and Practice*, 36(5), 889–913. doi: 10.1111/j.1540-6520.2012.00535.x
- Bohm, P. (1972). Estimating demand for public goods: An experiment. *European Economic Review*, 3(2), 111–130. doi: 10.1016/0014-2921(72)90001-3
- Busenitz, L. W. (1999). Entrepreneurial risk and strategic decision making: It's a matter of perspective. *Journal of Applied Behavioral Science*, 35(3), 325–340. doi: 10.1177/0021886399353005
- Busenitz, L. W. & Barney, J. B. (1997). Differences between entrepreneurs and managers in large organizations: Biases and heuristics in strategic decision-making. *Journal of Business Venturing*, 12(1), 9–30. doi: 10.1016/s0883-9026(96)00003-1
- Caliendo, M., Fossen, F. M. & Kritikos, A. S. (2009). Risk attitudes of nascent entrepreneurs—new evidence from an experimentally validated survey. *Small Business Economics*, 32, 153–167. doi: 10.1007/s11187-007-9078-6
- Campbell, D. T. & Stanley, J. C. (1966) *Experimental and quasi-experimental designs for research*. Houghton Mifflin Company. doi: 10.4135/9781483346229.n137
- Chandler, G. N. & Lyon, D. W. (2001). Issues of research design and construct measurement in entrepreneurship research: The past decade. *Entrepreneurship, Theory and Practice*, 25(4), 101–114. doi: 10.1177/104225870102500407
- Chang, J. B., Lusk, J. L. & Norwood, F. B. (2009). How closely do hypothetical surveys and laboratory experiments predict field behavior? *American Journal of Agricultural Economics*, 91(2), 518–534. doi: 10.1111/j.1467-8276.2008.01242.x
- Charness, G. (2010). Laboratory experiments: Challenges and promise: A review of "Theory and Experiment: What are the Questions?" by Vernon Smith. *Journal of Economic Behavior & Organization*, 73(1), 21–23. doi: 10.1016/j.jebo.2008.11.005
- Charness, G., Gneezy, U. & Kuhn, M. A. (2013). Experimental methods: Extra-laboratory experiments—extending the reach of experimental economics. *Journal of Economic Behavior & Organization*, 91, 93–100. doi: 10.1016/j.jebo.2013.04.002
- Chen, S. S., Ho, K. Y. & Ho, P. H. (2014). CEO overconfidence and long-term performance following R&D increases. *Financial Management*, 43(2), 245–269. doi: 10.1111/fima.12035
- Churchill, G. A. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research*, 16(1), 64–73. doi: 10.2307/3150876
- Coppola, M. (2014). Eliciting risk-preferences in socio-economic surveys: How do different measures perform? *The Journal of Socio-Economics*, 48, 1–10. doi: 10.1016/j.socec.2013.08.010
- Croson, R., Anand, J. & Agarwal, R. (2007). Using experiments in corporate strategy research. *European Management Review*, 4(3), 173–181. doi: 10.1057/palgrave.emr.1500082
- Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4), 349–354. doi: 10.1037/h0047358
- Ding X., Hartog, J. & Sun, Y. (2010). *Can we measure individual risk attitudes in a survey?* IZA working paper, n°4807.
- Dohmen T., Falk, A., Huffman, D. & Sunde, U. (2005). *Explaining trust: Disentangling the roles of beliefs and risk preference*. IZA Working Paper.
- Dohmen, T., Falk, A., Huffman, D., Sunde, U. et al. (2011). Individual risk attitudes: Measurement, determinants, and behavioral consequences. *Journal of the European Economic Association*, 9(3), 522–550. doi: 10.1111/j.1542-4774.2011.01015.x
- Douglas, E. J. & Shepherd, D. A. (2000). Entrepreneurship as a utility maximizing response. *Journal of Business Venturing*, 15(3), 231–251. doi: 10.1016/s0883-9026(98)00008-1
- Falk, A., Becker, A., Dohmen, T. J., Huffman, D. & Sunde, U. (2016). The preference survey module: A validated instrument for measuring risk, time, and social preferences. *SSRN Electronic Journal*, 1–66. doi: 10.2139/ssrn.2725874
- Falk, A. & Fehr, E. (2003). Why labour market experiments? *Labour Economics*, 10(4), 399–406. doi: 10.1016/s0927-5371(03)00050-2
- Fehr, E., Fischbacher, U., Von Rosenblatt, B., Schupp, J. et al. (2002). A nationwide laboratory: Examining trust and trustworthiness by integrating behavioral experiments into representative surveys. *SSRN Electronic Journal*, 1–23. doi: 10.2139/ssrn.385120
- Felin, T., Foss, N. J. & Ployhart, R. E. (2015). The microfoundations movement in strategy and organization theory. *The Academy of Management Annals*, 9(1), 575–632. doi: 10.1080/19416520.2015.1007651
- Galasso, A. & Simcoe, T. S. (2011). CEO overconfidence and innovation. *Management Science*, 5(8), 1469–1484. doi: 10.1287/mnsc.1110.1374
- Gillespie, R. (1991) *Manufacturing knowledge: A history of the Hawthorne experiments*. Cambridge University Press.
- Gneezy, U. & Imas, A. (2017) Lab in the field: Measuring preferences in the wild. In A. Banerjee & E. Duflo (Eds.), *Handbook of field experiments* (vol. 1, pp. 439–464). Elsevier. doi: 10.1016/bs.hefe.2016.08.003
- Goldstein, W. M. & Hogarth, R. M. (1997). Judgment and decision research: Some historical context. In W. M. Goldstein & R. M. Hogarth (Eds.), *Cambridge series on judgment and decision making. Research on judgment and decision making: Currents, connections, and controversies* (pp. 3–65). Cambridge University Press. doi: 10.1002/9780470752937
- Graf, L., König, A., Enders, A. & Hungenberg, H. (2012). Debiasing competitive irrationality: How managers can be prevented from trading off absolute for relative profit. *European Management Journal*, 30(4), 386–403. doi: 10.1016/j.emj.2011.12.001
- Grazzini, F. & Boissin, J.-P. (2013). Managers' mental models of small business acquisition: the case of the SME French Transfert market. *M@n@gement*, 16(1), 49–87. doi: 10.3917/mana.161.0049
- Grégoire, D., Binder, J. & Rauch, A. (2019). Navigating the validity tradeoffs of entrepreneurship research experiments: A systematic review and

- best-practice suggestions. *Journal of Business Venturing*, 34(2), 284–310. doi: 10.1016/j.jbusvent.2018.10.002
- Hainmueller, J., Hangartner, D. & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *Proceedings of the National Academy of Sciences*, 112, 2395–2400. doi: 10.1073/pnas.1416587112
- Hardeweg, B., Menkhoff, L. & Waibel, H. (2013). Experimentally validated survey evidence on individual risk attitudes in rural Thailand. *Economic Development and Cultural Change*, 61(4), 859–888. doi: 10.1086/670378
- Harrison, G. W. & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, 42(4), 1009–1055. doi: 10.1257/0022051043004577
- Hirich, R., Langan-Fox, J. & Grant, S. (2007). Entrepreneurship research and practice: A call to action for psychology. *American Psychologist*, 62(6), 575–589. doi: 10.1037/0003-066x.62.6.575
- Hogarth, R. M. (1987). *Judgement and choice: The psychology of decision* (2nd ed.). John Wiley.
- Huber, L. R., Sloof, R. & Van Praag, M. (2014). The effect of early entrepreneurship education: Evidence from a field experiment. *European Economic Review*, 72(C), 76–97. doi: 10.1016/j.eurocorev.2014.09.002
- James, H. (2006). *Family capitalism: Wendels, Haniels, Falcks, and the continental European model*. Harvard University Press.
- Johnson, N. D. & Mislin, A. (2012). How much should we trust the world values survey trust question? *Economics Letters*, 116(2), 210–212. doi: 10.1016/j.econlet.2012.02.010
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, 93(5), 1449–1475. doi: 10.1257/000282803322655392
- Kahneman, D., Slovic, P. & Tversky, A. (1982). *Judgement under uncertainty: Heuristics and biases*. Cambridge University Press. doi: 10.1017/cbo9780511809477
- Kirk, R. (1995). *Experimental design, procedures for the behavioural sciences*. Brooks/Cole Publishing Company.
- Laroche, H. & Nioche, J. P. (2015). L'approche cognitive de la stratégie d'entreprise. *Revue Française de Gestion*, 8(253), 97–120. doi: 10.3166/rfg.160.81-108
- Lavrakas P. J. (2008). *Encyclopedia of survey research methods*. Sage. doi: 10.4135/9781412963947
- List, J., Sadoff, A. & Wagner, M. (2011). So you want to run an experiment, now what? Some simple rules of thumb for optimal experimental design. *Experimental Economics*, 14(4), 439–457. doi: 10.1007/s10683-011-9275-7
- McNeil, B., Pauker, S., Sox, H. & Tversky, A. (1982). On the elicitation of preferences for alternative therapies. *The New England Journal of Medicine*, 306(21), 1259–1262. doi: 10.1056/nejm198205273062103
- Mintzberg, H. (1973). *The nature of managerial work*. Harper & Row.
- Mitchell, R., Busenitz, L., Bird, B., Caglio, C. et al. (2007). The central question in entrepreneurial cognition research 2007. *Entrepreneurship Theory and Practice*, 31(1), 1–27. doi: 10.1111/j.1540-6520.2007.00161.x
- Mitchell, R., Busenitz, L., Lant, T., McDougall, P. et al. (2002). Toward a theory of entrepreneurial cognition: Rethinking the people side of entrepreneurship research. *Entrepreneurship, Theory and Practice*, 27(2), 93–104. doi: 10.1111/1540-8520.00001
- Mizik, N. (2010). The theory and practice of myopic management. *Journal of Marketing Research*, 47(4), 594–611. doi: 10.1509/jmkr.47.4.594
- Molina-Azorin, J. F., Lopez-Gamero, M. D., Pereira-Moliner, J. & Pertusa-Ortega, E. M. (2012). Mixed methods studies in entrepreneurship research: Applications and contributions. *Entrepreneurship and Regional Development*, 24(5–6), 425–456. doi: 10.1080/08985626.2011.603363
- Moore, D. & Flynn, F. J. (2008). The case for behavioral decision research in organizational behavior. *The Academy of Management Annals*, 2(1), 399–431. doi: 10.1080/19416520802211636
- Moore, D. A. & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2), 502–517. doi: 10.1037/0033-295x.115.2.502
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds.), *Measures of social psychological attitudes* (vol. 1, pp. 17–59). Academic Press. doi: 10.1016/b978-0-12-590241-0.50006-x
- Phan, P. & Wright, M. (2018). Advancing the science of human cognition and behavior. *Academy of Management Perspectives*, 32(2), 179–181. doi: 10.5465/amp.2018.0058
- Powell, T., Lovallo, D. & Fox, C. (2011). Behavioral strategy. *Strategic Management Journal*, 32(13), 1369–1386. doi: 10.1002/smj.968
- Sadoff, S. & Samek, A. (2019). Can interventions affect commitment demand? A field experiment on food choice. *Journal of Economic Behavior & Organization*, 158(C), 90–109. doi: 10.1016/j.jebo.2018.11.016
- Shepherd, D. A., Williams, T. A. & Patzelt, H. (2015). Thinking about entrepreneurial decision making: Review and research agenda. *Journal of Management*, 41(1), 11–46. doi: 10.1177/0149206314541153
- Simon, H. A. (1947). *Administrative behavior*. Macmillan.
- Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, 69(1), 99–118. doi: 10.2307/1884852
- Simon, H. A. (1979). Rational decision making in business organizations. *American Economic Review*, 69(4), 493–513.
- Song, Y., Liu, Y., Lanaj, M. W. K., Johnson, R. E. et al. (2018). A social mindfulness approach to understanding experiences customer mistreatment: A within person field experiment. *Academy of Management Journal*, 61(3), 994–1020. doi: 10.5465/amj.2016.0448
- Stein, J. C. (1988). Takeover threats and managerial myopia. *Journal of Political Economy*, 96(1), 61–80. doi: 10.1086/261524
- Stein, J. C. (1989). Efficient capital markets, inefficient firms: A model of myopic corporate behavior. *Quarterly Journal of Economics*, 104(4), 655–669. doi: 10.2307/2937861
- Thaler, R. & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Tashakkori, A. & Teddlie, C. (2003). *Handbook of mixed methods in social & behavioral research*. Sage.
- Vieider, F., Lefebvre, M., Bouchouicha, R., Chmura, T. et al. (2015). Common components of risk and uncertainty attitudes across contexts and domains: Evidence from 30 countries. *Journal of the European Economic Association*, 13(3), 421–452. doi: 10.1111/jeea.12102
- Vischer, T., Dohmen, T., Falk, A., Huffman, D. et al. (2013). Validating an ultra-short survey measure of patience. *Economics Letters*, 120(2), 142–145. doi: 10.1016/j.econlet.2013.04.007
- Vo, L.-C., Culié, J.-D. & Mounoud, E. (2016). Microfoundations of decoupling: From a coping theory perspective. *M@n@gement*, 19(4), 248–276. doi: 10.3917/mana.194.0248
- Williams, D., Wood, M., Mitchell, R. & Urbig, D. (2019). Applying experimental methods to advance entrepreneurship research: On the need for and publication of experiments. *Journal of Business Venturing*, 34(2), 215–223. doi: 10.1016/j.jbusvent.2018.12.003
- Zellmer-Bruhn, M., Caligiuri, P. & Thomas, D. C. (2016). From the editors: Experimental designs in international business research. *Journal of International Business Studies*, 47(4), 399–407. doi: 10.1057/jibs.2016.12
- Zikmund, G., Babin, B. J., Carr, J. C. & Griffin, M. (2013). *Business research methods*. Cengage Learning Custom.