

Teaching Evidence-Based Management With A Focus on Producing Local Evidence

JOERG DIETZ
JOHN ANTONAKIS
ULRICH HOFFRAGE
FRANCISKA KRINGS
JULIAN N. MAREWSKI
CHRISTIAN ZEHNDER
University of Lausanne

We present an approach to teaching evidence-based management (EBMgt) that trains future managers how to produce local evidence, causally interpretable data that is collected on-site in companies to address a specific business problem. Our teaching method is a variant of problem-based learning, a method originally developed to teach evidence-based medicine. Following this method, students learn an evidence-based problem-solving cycle for addressing business cases. In executing this cycle, students use and produce scientific evidence through literature searches and the design of local, experimental tests of causal hypotheses. We argue the value of teaching EBMgt with a focus on producing local evidence, how it can be taught, and what can be taught. We conclude by outlining our contribution to the literature on teaching EBMgt and by discussing limitations of our approach.

Maria advises a very well-known charity. Deeply concerned about an ever-decreasing volume of donations, this charity has partnered with several on-line shops. When purchasing books, cards, and other goods in these shops, customers can indicate their agreement to round up the amount charged to their credit cards (e.g., from US\$9.50 to US\$10) and donate the difference. Yet, the donation influx from these partnerships is negligible. After having reviewed the scientific literature on consumer behavior and donations, Maria proposes that the on-line shops implement a different default: Rather than setting up the websites such that customers do not donate unless they indicate they would like

to do so, *all* purchases should automatically be rounded up, *unless* customers indicate they prefer not to do so. Although research provides evidence for the effectiveness of this default principle (e.g., Johnson & Goldstein, 2003; Thaler & Sunstein, 2009), the president of the charity is skeptical: Will donors react negatively to a default that is set to their financial disadvantage and ultimately donate even less? Maria replies that an experiment could assess both the positive effects and the harmful outcomes of the new procedure. The charity could test that procedure against the existing one using as criteria the donations received as well as overall customer and donor satisfaction.

Maria is not a professional consultant, but a master's student in management who was part of a team that conducted a laboratory experiment on the donation project. She has benefitted from our course on evidence-based management (EBMgt), which is based on a simple idea: Students learn to identify management problems and to use scien-

We would like to thank Special Issue Editors Denise Rousseau and Sara Rynes-Weller as well as the anonymous reviewers for their feedback and encouragement. We are also grateful to Emmanuelle Kleinogel, Justin Olds, and Alain Salamin for their helpful comments and to Tobias Dennerlein and Thomas Fischer for their assistance.

tific methods to develop solutions. Central to the approach is that interventions ought to be tested with causally interpretable data that are collected on the site where managers execute their interventions. We refer to such causally interpretable data that are relevant to a specific problem at hand as *local evidence*. Local evidence is an example of what Rousseau (2006: 260) called "little e evidence"... "data systematically gathered in a particular setting to inform local decisions."

The purpose of this article is to advocate and present an approach to teaching EBMgt with a particular focus on producing such local evidence. The first section elaborates on *why* we advocate this pedagogical focus. The second explains *how* our teaching approach can be executed through a problem-based method. We propose that, contrary to concerns among some management scholars (e.g., Shugan, 2006), case studies can be powerful stimuli for learning about EBMgt and academic research. The third section details what is taught, notably an evidence-based problem-solving cycle that emphasizes local testing. We conclude by explaining our contribution to the literature on teaching EBMgt and by discussing limitations of our approach.

REASONS FOR TEACHING EBMGT WITH A FOCUS ON PRODUCING LOCAL EVIDENCE

Evidence-based managers (1) ask a focused question; (2) acquire the evidence that pertains to the question; (3) appraise the evidence; (4) apply it; and (5) analyze the efficiency and effectiveness of the decision and, if necessary, make adjustments (e.g., Jolley, Carroll, & Rousseau, 2012). More formally, EBMgt has been defined as "making decisions through the conscientious, explicit, and judicious use of four sources of information: practitioner expertise and judgment, evidence from the local context, a critical evaluation of the best available research evidence, and the perspectives of those people who might be affected by the decision" (Briner, Denyer, & Rousseau, 2009: 19). This definition is hard to digest for many students, who ask, for example, what exactly is the "best" available research evidence or when would they know that their evaluation was sufficiently "critical"? The definition signals that EBMgt is a complex craft. Only after students like Maria have tried to produce local evidence do they understand *how* complex a craft it really is; therefore, and for other pedagogical and practical reasons, we advocate teaching EBMgt with a concrete and narrow focus on producing local evidence.

Pedagogical Reasons for a Focus on Producing Local Evidence

There are two pedagogical reasons for focusing on producing local evidence: First, a relatively narrow focus makes the broad and difficult subject of EBMgt more tangible. Students have to absorb the logic of the scientific method and must learn to gather and evaluate evidence—whether existing or new data. Furthermore, in practicing EBMgt, students work on business problems that are often ill-defined and ill-structured; thus, the complex nature of EBMgt can easily overwhelm them. The focus on producing local evidence provides students with a vision and lets them steer their efforts toward developing scripts for EBMgt. Producing local evidence also allows learning by practicing on actual problems rather than by abstracting from knowledge.

The second benefit is that by learning to produce local evidence, students learn to evaluate other research. Producing local evidence helps develop causal reasoning skills, as students internalize the conditions of causality and acquire methods for establishing the causal interpretability of their evidence. They also gain a sense for the "nitty-gritty" of research, such as the assessment of variables through subjective or objective indicators. Eventually, students can use their skills on how to produce evidence to evaluate existing evidence, just as members of editorial boards of academic journals rely on their experience in reviewing manuscripts. Stefan, another one of our students who subsequently became a management trainee at a large multinational, offers an example. He explained that he literally chuckled when one consultant claimed that the evidence "proved" that emotional intelligence, a construct controversially debated by scientists (e.g., Harms & Credé, 2010), led to leadership effectiveness. This consultant appeared confused when Stefan asked whether the so-called evidence was causally interpretable; that is, if the possibility that cognitive ability and personality traits might affect both emotional intelligence and leadership skills had been ruled out.

Practical Value of Local Evidence for Managers

Focusing on producing local evidence is practical because it provides managers with useful information. In practicing EBMgt, managers first take into account and evaluate the currently available evidence before they embark on producing new local evidence. Once existing data have passed the

check for generic relevance to the business problem at hand, two key questions remain: Can the existing evidence be interpreted causally, and can it be applied to the local context? If the existing evidence falls short on these accounts, locally produced evidence becomes an important complement to evidence collected elsewhere.

For evidence-based managers, cross-sectional evidence that two variables are related—or even a study suggesting one variable “causes” another—is not sufficient for causal interpretability. What they need to know is that the evidence for a causal link is robust, meaning there is a causal effect that has been reproduced frequently and across different settings. For example, just as aspirin has been repeatedly found to reduce the risk of colon cancer in randomized controlled trials (Thun, Jacobs, & Patrono, 2012), a corporate trainer who evaluates pedagogical methods will find that there is robust experimental evidence of positive effects of active processing on learning (e.g., Bransford, Brown, & Cocking, 2000; Prince, 2004).

“For evidence-based managers, cross-sectional evidence that two variables are related—or even a study suggesting one variable “causes” another—is not sufficient for causal interpretability. What they need to know is that the evidence for a causal link is robust, meaning there is a causal effect that has been reproduced frequently and across different settings.”

Unfortunately, causally interpretable evidence might not be available because randomization is often not possible when studying management phenomena, or because the evidence might be outdated, inconsistent, contradictory, or poorly described (Pfeffer & Sutton, 2006). Published scientific evidence might also represent a biased sample of the entire body of evidence (Kepes & McDaniel, 2013), partly due to preferences for publishing data that support rather than invalidate theories. In these instances, local testing of a managerial intervention is often indispensable to show that an intervention has an effect where it is actually implemented.

A second practical benefit of producing local evidence is that it provides a check on local applicability. This issue translates essentially into a judgment of boundary conditions (i.e., in which contexts does a variable have a causal effect?) imposed by the local environment in which a manager seeks to implement an intervention. The identification of boundary conditions is typically characteristic of advanced stages of research on a concept (Reichers & Schneider, 1990). For example, the concept of service climate was introduced in the early 1980s (Schneider, Parkington, & Buxton, 1980), but it was only 2 decades later that the identification of boundary conditions commenced (e.g., Dietz, Pugh, & Wiley, 2004) and another decade after that before the first meta-analysis on service climate was published (Hong, Liao, Hu, & Jiang, 2013). Thus, depending on the stage of scientific progress, local evidence is needed to supplement existing evidence to assess local applicability. This is especially true when evaluating management fads and en-vogue ideas (e.g., selecting personnel based on emotional intelligence; Zaccaro & Horn, 2003).

When existing evidence is causally interpretable and boundary conditions have been established, managers might not need to establish local applicability, particularly when an effect has been found consistently in a context that is similar to theirs. Some effects, however, are highly context-sensitive. Classic examples are pay-for-performance schemes, which managers often view as a cure for motivation problems across contexts: On the one hand, the evidence indicates that these schemes improve employee performance (Rynes, Gerhart, & Parks, 2005); on the other, it also shows that they tend to increase performance on simple tasks, but not on more complex tasks that, for example, require creativity (e.g., Baer, Oldham, & Cummings, 2003; Byron & Khazanchi, 2012). The context sensitivity of management research has led scholars (Johns, 2006; Rousseau & Fried, 2001) to call for contextualizations to reflect that a cause may yield different effects depending on the cultural and organizational setting. The same holds true for EBMgt. To quote Pfeffer and Sutton (2006: 64): “Because companies vary so wildly in size, form, and age . . . it is far more risky in business to presume that a proven ‘cure’ developed in one place will be effective elsewhere.” If it is not clear that effects generalize, producing local evidence is necessary to replicate effects locally that have been found elsewhere.

Finally, local evidence equips managers with strong arguments for convincing other stakeholders that interventions work. At the casino chain Harrah's, managers conducted an experiment to compare the effects of different promotional packages on gambling revenue (Lal & Carolo, 2004). Customers in the control group received a standard package worth \$125 (a free room, two dinners, and \$30 in chips); whereas those in the treatment group were given only \$60 in chips. The revenue from the treatment group was higher, a finding that contradicted conventional wisdom at that time. Another example is that of Google's use of a data-driven approach to convince its engineers that people management indeed matters for retaining talented employees (Gavin, 2013). The opening example of Maria is also a case in point. The president of the charity will find it difficult to argue on the basis of his intuitive concerns if data collected on site from actual donors show the opposite.

In summary, both pedagogical reasons and the usefulness of local evidence for EBMgt justify teaching the skills necessary for collecting one's own causally interpretable data in one's company. A focus on producing local evidence helps students visualize EBMgt and also sharpens skills for evaluating existing evidence. Moreover, producing local evidence is a managerially useful complement to existing evidence when the latter is not sufficiently relevant to the problem at hand, lacks causal interpretability, or cannot be locally applied.

HOW TO TEACH EBMGT

We advocate using a variant of problem-based learning (PBL; Barrows, 1996), a method originally developed to teach evidence-based medicine. PBL typically opens with a case describing a problem of a patient. Aspiring doctors formulate a researchable question from the problem description (typically symptoms of a patient), search and evaluate the literature to identify quality evidence, and eventually on the basis of the evidence suggest how to treat a patient's symptoms (Rosenberg & Donald, 1995). Our approach to teaching EBMgt also starts with cases, but the patients are businesses and the symptoms are business problems. Students use an evidence-based problem-solving cycle to address these problems. This cycle is composed of four steps: first, students define the problem. Second, to analyze the problem and develop

solutions, they consult and evaluate academic and other evidence. Third, they practice designing and executing experimental tests of problem solutions. Fourth, they evaluate test results and recommend which solution should be implemented.

Three theoretical arguments can be made for using cases to facilitate learning EBMgt. First, as Briner and colleagues (2009: 20) noted, EBMgt "starts with the questions, problems, and issues faced by managers and other organizational practitioners." Working on cases allows students to simulate EBMgt in vivo, which eases the transfer of learning to the real world (Hmelo-Silver, 2004). Second, with a facilitated transfer comes an enhanced motivation. Students who see the practical relevance of their learning become intrinsically motivated (Ferrari & Mahalingam, 1998). Jolley and colleagues (2012: 340–341) noted that "students are hungry for cases and real-life illustrations of managers using evidence"—an observation that we can only confirm. Third, a PBL-based approach requires active processing by students, which, as much research has shown (e.g., Prince, 2004), enhances learning. A synopsis of the first case in our course, which is based on a study by Fehr and Goette (2007), serves as an example.

Case Example

The Bike Messenger Case

A bike messenger company in Zurich, Switzerland, employed mostly students, who were paid exclusively on commission. The company's owners planned an experiment to test what effect a 25% raise in commissions would have on the revenues generated by the bike messengers. On the basis of economic theory and research, the owners believed that the raise should motivate the messengers to work more shifts per week: With a higher commission rate, work would become more attractive compared to leisure, and messengers should show more effort during each shift. What would an experimental test of that hypothesis look like?

Prior to coming to class, our students read an article on EBMgt (Briner et al., 2009) as well as two papers on compensation (Kerr, 1995; Rynes et al., 2005). They are also asked to search for other problem-relevant literature. During the class discussion, students initially focus on the business problem, which can be defined as a gap between current revenues generated by employees and desired revenues. When analyzing the compensation system and its effects on performance, students

explain that the scientific evidence indicates the effectiveness of individual incentives for performance, in particular when it comes to relatively simple tasks such as those of bike messengers. Yet, they feel uncertain about the local applicability of this evidence, in part because few studies examine companies that pay their employees exclusively on a commission basis. Some students point out that the raise might not work as intended because bike messengers who are students might only work as much as needed to fund their studies. Course participants also suggest alternative interventions (e.g., other reward systems or the use of electric bicycles).

In discussing the experimental design, students typically converge on a between-participants design with random assignment of bike messengers to two groups, a treatment group in which the higher commission is paid and a control group in which the status quo is maintained. Eventually, the instructor reveals the design and results of the actual experiment carried out at the company:

For the experiment, the messengers were randomly assigned to one of two groups. In group A, the commission was increased by 25% in September; in group B, the increase took place in November, whereby in both months the non-treated group served as the control group. As it turned out, this increase motivated the bike messengers to work more shifts per week. At the same time, however, their revenues per shift dropped.

In evaluating the actual study and the results, students tend to recognize that having both groups experience the raise in commission strengthens the experimental design. The design ensures that effects cannot be explained by a unique reaction of one group to the treatment. It further eliminates ethical concerns that not all bike messengers would have been given a raise. The students are quite apt at explaining the unpredicted drop in performance per shift by citing the bike messengers' likely daily income targets, and their proposal not to introduce the raise permanently reflects what actually happened in the company.

In debriefing, the instructor primarily points to two issues: First, the owners followed a systematic process that included testing their proposed solution. Second, the owners could have improved how they executed the problem-solving

cycle. For example, in the analysis stage their search for existing evidence was too narrow, leading them to settle too quickly on raising the commission, and they should have involved the bike messengers. Still, the owners kept their task manageable, collected data, and hence, could uncover that in their company raising the commission had an unexpected effect on the bike messengers' performance per shift. Having given an example, we now introduce the teaching methodology in more detail.

Teaching Method

The bike messenger case illustrates our use of realistic problems as foci and stimuli for learning, to which scientific evidence and processes are applied. Other aspects of our approach are student-centeredness; instructors as facilitators; and a combination of individual, small-group, and classroom learning.

Use of Realistic Problems as Foci and Stimuli for Learning

The course relies on two main types of activities: case analyses and a project in which students produce local evidence while serving as managerial consultants. Turning first to cases, good cases for teaching EBMgt allow students to go through the evidence-based problem-solving cycle (for examples, see Table 1). Due to the lack of suitable cases (for exceptions, see Kovner, Fine, & D'Aquila, 2009, and Goodman & O'Brien, 2012), we have turned consulting experiences and academic articles (e.g., Fehr & Goette, 2007; Greenberg, 1990) into case studies which are available from us. We have also adjusted the teaching of cases that were written for a different purpose, such as the Elise Smart case (Gandz & Spracklin, 2007), to facilitate the use of scientific frameworks of employee performance (Boxall & Purcell, 2003; Campbell, McCloy, Oppler, & Sager, 1993).

In the EBMgt project, students team up with peers to work as consultants for a company. They have to find a client (i.e., a manager), identify and define a problem of interest to the manager, establish a plan for executing the evidence-based problem-solving steps, search and evaluate existing evidence to analyze and develop solutions to the problem, design and execute an experiment to test the proposed solution, analyze the data, and give a recommendation to the manager for ad-

TABLE 1
Teaching Plan: Evidence-Based Management (EBMgt) as Problem-Solving

Session	Case/problem	Possible readings
Introductory sessions		
1. Course introduction: EBMgt	Bike Messenger Case ^a based on Fehr & Goette (2007)	<ul style="list-style-type: none"> • Briner et al. (2009) • Kerr (1995) • Rynes et al. (2005)
2. Problem-solving	Taking Human Resources Seriously in Minneapolis (Johnson & Marietta, 2009)	<ul style="list-style-type: none"> • Childress & Marietta (2010) • Baron & Kreps (1999)
3. Reasons for and against EBMgt	Thirty questions based on Rynes et al. (2002) and Groysberg et al. (2007)	<ul style="list-style-type: none"> • Mientka (2013) • Rousseau & Barends (2011)
4. Examples of EBMgt	Guest speaker: Evidence-based manager	<ul style="list-style-type: none"> • Hodgkinson & Rousseau (2009) • Pfeffer & Sutton (2006)
Main sessions		
5. Defining a problem	Elise Smart (Gandz & Spracklin, 2007)	<ul style="list-style-type: none"> • Boxall & Purcell (2003) • Campbell et al. (1993) • Smith (1989)
6. Developing solutions	Room Cleanliness Case ^a	<ul style="list-style-type: none"> • Weick (1989) • Ghoshal (2005)
7. Types of evidence and criteria for evaluating them	Case on Work from Home ^a based on Bloom et al. (2013)	<ul style="list-style-type: none"> • Bergh, Hanke, Balkundi, Brown, & Chen (2004) • Antonakis et al. (2014) • Bailey & Kurland (2002)
8. Experiments and causality	Windermere Manor: Sustainability and Change (Joshi et al., 2013)	<ul style="list-style-type: none"> • Angrist & Pischke (2010) • Bandiera, Barankay, & Rasul (2011) • Goldstein et al. (2008) • Schultz et al. (2008)
9. Quasi-experiments	Case based on Greenberg ^a (1990)	<ul style="list-style-type: none"> • Greenberg (2009) • Shadish, Cook, & Campbell (2002)
10. Managerial evidence	Carter Racing (Brittain & Sitkin, 2006)	<ul style="list-style-type: none"> • Denrell (2005)
11. Evaluating solutions: basics	Kenexa (Dietz & Joshi, 2007)	<ul style="list-style-type: none"> • Rosenzweig (2007)
12. Evaluating solutions: false inferences from success	Ducati: The Making of a Great Motorcycle (Gino & Pisano, 2006)	<ul style="list-style-type: none"> • March (1991) • Gino & Pisano (2011)
Conclusion		
13. Course review	Project Presentations and Review	<ul style="list-style-type: none"> • Davenport (2009)

^aThese cases are available from the first author.

addressing the problem. At the end of the project, students give a presentation and write a report. Student teams work on the project throughout the semester and meet with the instructor several times. They also have access to advanced PhD students as mentors.

The EBMgt project is a critical supplement to the cases. Many partnering managers are not familiar with EBMgt, requiring the students to be consultants and educators. The first two steps of problem definition and analysis tend to be tremendous challenges. In contrast to the cases, it is often not clear whether there is a problem to be solved. Managers, for example, simply point to financial figures that could be improved. Once problem definitions become operational (e.g., a gap between the actual and desired sales of cell phone contracts), initial working hypotheses

about interventions are typically lacking. Hence, in contrast to the cases, student teams conduct systematic reviews for the project. During these reviews, the trade-off between analyzing the problem (including the search for existing evidence) and having to suggest interventions (i.e., formulating testable hypotheses) is regularly only resolved by project deadlines.

Another challenge is that in the design and testing phase, student teams frequently cannot collect data at a corporate location. Because collecting data is a crucial difference between the project and the cases, we also allow data collection at the university (e.g., in classes or using our laboratory). Student teams execute simple experiments, typically with no more than one independent variable and two or three conditions (treatment, placebo, and control). These simplified data collection ef-

forts still enable students to get their hands “dirty” (e.g., development of experimental materials; being an experimenter; entering, cleaning, and analyzing data). Furthermore, these efforts can serve as pretests (e.g., to determine the strength of an experimental treatment) for potential tests in the field, if the partnering manager so desires. Finally, laboratory experiments often stimulate follow-up research. For example, the experiment by Maria’s team on the effects of defaults on donations has resulted in an opportunity for a field experiment with the charity.

“Student teams execute simple experiments, typically with no more than one independent variable and two or three conditions (treatment, placebo, and control).”

Student-Centeredness

As the course progresses, we make the students increasingly responsible for their learning. At the outset, students learn an evidence-based problem-solving cycle as an organizing framework (see Goodman & O’Brien, 2012, on scaffolding and its fading over time). The initial two cases (the Bike Messenger Case and Johnson & Marietta, 2009) introduce this cycle in its entirety. Subsequently, the students apply this cycle to other cases typically with a focus on one step (e.g., defining the problem) while touching on the entire cycle. Moreover, students are fully responsible for the EBMgt project, just as they would be as managers in their companies.

To enable students to search for existing evidence, at the outset of the course we practice literature searches. Subsequently, students are charged with seeking evidence individually and in their teams. Our teaching experience, however, leads us to concur with Goodman and O’Brien (2012: 325) that “it is unrealistic to expect business students to search through, read, understand, and apply the academic management literature at the same level as academics who regularly consume and perform research.” Therefore, we support students in their literature searches by providing at least one scientific article per case. Our support reflects that evidence-based managers might consult experts who can point

them into the right direction (Rousseau & Barends, 2011).

Instructors as Facilitators

The other side of student-centeredness is that instructors have the role of *facilitators* rather than originators of learning. Facilitators ensure the use of the evidence-based problem-solving cycle and monitor students’ use of scientific evidence and processes, a critical difference from the research-ignoring case teaching lamented by Shugan (2006). By using questioning strategies, facilitators provide guidance toward enacting EBMgt and motivate students to reflect about their learning. They avoid, however, providing too much direction. For example, if students recommended a pre-post design without a control group for the Bike Messenger Case, we would ask them to explain their rationale for this design and to describe its advantages and disadvantages. We, however, would not prescribe the design that was actually used. Moreover, facilitators give specific and immediate feedback selectively, as it can undermine learning and its transfer to other contexts (Goodman & O’Brien, 2012).

Combination of Individual Learning, Small-Group Learning, and Classroom Learning

Initially, students prepare individually for the classes. Then, as members of student teams, they cooperate in working on the case studies. For example, for each case a different team member is responsible for identifying and consolidating scientific evidence and for sharing it with teammates. The shared responsibility for searching evidence facilitates teamwork and ensures that each student conducts literature searches, while keeping the workload manageable. In each 90-minute session, the student teams work independently on a case for about 40 minutes. The purpose of small-group work is to activate existing relevant knowledge and, through the exchange of perspectives, add new knowledge. Finally, the class discussion with all students elaborates on and consolidates the learning from the small-group work.

WHAT TO TEACH

Our approach places a strong emphasis on producing local evidence, but it is still an approach to teaching EBMgt. Hence, we have not designed a course in which students merely learn to conduct

causally interpretable tests in the field. Instead, students practice a complete evidence-based problem-solving cycle that includes the evaluation of existing evidence and emphasizes the testing of solutions. To present the content of our course, we have organized this section into four parts. In the first, we use another case study (Joshi, Bapuji, & Chandrasekhar, 2013) to illustrate that our problem-solving cycle is evidence-based and follows the scientific method. In the second and third parts, we describe the introductory and main phases of teaching our approach to EBMgt. The fourth part elaborates on the assessment of students' learning and the evaluation of the course.

Evidence-Based Problem-Solving Cycle

Case Example

Improving Towel Reuse. *The general manager of a 60-room hotel in London, Ontario (Canada), was concerned that only 25% of towels were reused, although 75% of guests indicated a willingness to reuse towels during check-in. Guests found the following signage in bathrooms: "Save our planet. Dear Guests: Everyday millions of liters of water are used to wash towels that have been used only once. You make the choice. A towel on the towel rack means "I will use again." A towel in the bathtub means "Please exchange." Please decide yourself. Thank you for helping us conserve the earth's vital resources." Housekeepers had observed that hotel guests left their towels more or less everywhere in the bathroom and the bedroom. Guests noted that they wanted to reuse towels, but that when they left them to dry on the shower rod or a chair, housekeepers replaced the towels. How should the procedure for towels be changed, and how could the effectiveness of this change be tested?*

The scientific method implies a systematic approach to problem solving that can be organized along four steps (Tilly, 2008):

Step 1: Problem definition. What is the problem, if any, that I would like to solve? What are the key outcome variables?

Step 2: Problem analysis. Why does the problem occur? What are variables that could be potential causes of the outcome? What information does existing evidence provide for analyzing the problem? What are potential solutions?

Step 3: Testing of solutions. How does one test the proposed solution or intervention?

Step 4: Evaluation of tested solutions. Do they have the expected effect?

Students encounter the Towel Reuse Case in their eighth session after they have already become familiar with the above cycle. Along with the case, students read two articles on field experiments about towel reuse in U.S. hotels (Goldstein, Cialdini, & Griskevicius, 2008; Schultz, Khazian, & Zaleski, 2008). Concerning step 1, problem identification and definition, students quickly settle on defining the problem as a gap between the current low rate of towel reuse and a higher desired rate. For step 2 (problem analysis including the identification of solutions), students draw on the existing evidence by Goldstein et al. and Schultz et al., which is based on theories about social influence. Being experiments, each of the five studies in these two articles can be interpreted causally. These field experiments indicate that a potential solution lies in the use of normative messages (e.g., "join the 75% of your fellow guests who reuse their towels"). The five studies, however, also partially contradict each other about whether both descriptive and injunctive norms should be used and whether local framing (e.g., "join the 75% of your fellow guests in this room who reused their towels") further increases towel reuse.

Students debate intensely about the local applicability of these studies for two reasons. First, there are concerns about cultural differences and changes in environmental attitudes since 2008. Second, as detailed as the scientific articles are, information on some potential boundary conditions, such as the customer segment and the size of the towel racks, is not provided. Students realize that only an empirical test can settle this debate.

For step 3, the testing, students benefit from knowing about Goldstein and colleagues (2008) and Schultz and colleagues (2008). They discuss, for example, why these researchers used experiments with control groups. They also pick up on details, such as the measurement of towel reuse (e.g., should only those towels be counted that customers place on the towel racks or should all towels be counted that customers do not put in the bathtub?), the training of the housekeeping staff, and the duration of the test. Although the possibility of a pre-post design without a control group is raised, the preferred test pits a normative message

(treatment group) against the existing message (control group) with random assignment of messages to rooms. For this case, information on what the general manager actually did is not available and, hence, students can only discuss how they would evaluate the proposed solution.

The problem-solving cycle proposed above varies from typical managerial problem-solving cycles in three critical aspects. First, the deliberate checking for existing evidence contrasts with what managers usually do. On the basis of a study of 356 managerial decisions, Nutt (1999) found that managers in most decisions (80%) intuitively picked one course of action without considering other options or systematically examining the evidence. The success rates of multiple-option-based decisions, however, were higher. In an evidence-based problem-solving cycle, decisions are not based on intuitive preferences but on an analysis of the existing evidence. In the towel reuse case, the evidence included causally interpretable field experiments that were based on research about normative social influence (e.g., Cialdini & Goldstein, 2004). This research has repeatedly shown that the behavior of others affects individuals' own behaviors, particularly in novel, ambiguous, or uncertain situations. The facilitator can pose the following questions to stimulate reflection about existing evidence (obviously to be adjusted for cases other than the Towel Reuse Case):

- What are common assumptions about towel reuse in hotels and interventions to increase it? Which interventions have already been tested to increase towel reuse? What does the evidence suggest about the effectiveness of these interventions? What are sources for this evidence? Is the evidence based on data, and can these data be causally interpreted? Is the evidence locally applicable?
- What are the basic arguments of theories of normative social influence? What is your evaluation of these arguments? How and why have you arrived at this evaluation?
- Are you possibly ignoring other sources of evidence, for example, from other disciplines?

A second feature that distinguishes the evidence-based problem-solving cycle from other problem-solving cycles is the emphasis on rigorous causal analyses. Managers pose causal questions such as "if I do this (i.e., the cause), will that happen (i.e. the effect)?" but their causal analyses are often flawed. For example, retrospective analyses of performance problems are typically biased by knowl-

edge about the performance itself (e.g., Rosenzweig, 2007). An evidence-based problem-solving cycle includes a rigorous analysis of causality of both the existing evidence and new test designs for producing local evidence. Facilitators can stimulate analyses of causality with questions about its conditions (e.g., Antonakis, Bendahan, Jacquart, & Lalive, 2010) and its direction:

- Does the presumed cause temporally precede the effect?
- Is the presumed cause systematically related to the effect?
- Are there other variables that are causally related to the same effect and, if so, how are these other variables related to the presumed cause?
- Might the effect also predict the presumed cause?
- Are there other threats to causality (e.g., selection effects)?

Third, in practice, solutions are often not explicitly tested after their implementation. This shortcoming can be due to lack of time, financial or personal resources, or other obstacles, but it can also result from methodological issues. In an evidence-based problem-solving cycle, a robust test of the proposed solution is crucial, and facilitators can probe for testability with the following questions:

- Can different solutions be distinguished from each other?
- What are measurable criteria for assessing solutions? How do you gauge the criteria?
- Is your test set up such that empirical evidence could not only confirm but also contradict your hypotheses?

The testability of solutions requires a rigorous application of all steps of the problem-solving cycle. If, for example, a problem or the variables that constitute the problem are not well-defined, it is not clear against which criteria solutions should be evaluated. Having discussed the evidence-based problem-solving cycle, we now describe in more detail the two phases of our approach to teaching EBMgt, the introductory phase and the main phase (see Table 1 for a list of the thirteen 90-minute sessions including cases and readings).

Introductory Phase: Laying the Foundation for EBMgt

The introductory phase consists of four sessions that aim to make students familiar with the ba-

sics of EBMgt, notably literature searches and the evidence-based problem-solving cycle, address concerns about EBMgt, and raise students' motivation for it. In the first session, after an introduction to the course and to literature search, students work on the Bike Messenger Case described before to practice the four steps of the problem-solving cycle. They do so again in the second session with another case (Johnson & Marietta, 2009). Students are very engaged in working on the cases.

In the third session, students respond to a survey of 30 questions concerning managerial myths about leadership and managing people (Groysberg, Lane, & Knoop, 2007; Rynes, Colbert, & Brown, 2002; for a session similar to our third one, see Jelley et al., 2012). The pattern of responses by students is similar to that of managers as reported by Rynes et al. (2002), showing that common beliefs and scientific evidence are often inconsistent. Although some students react negatively to having their false beliefs exposed, in general, the survey allows students to see the usefulness of consulting the existing evidence.

“The pattern of responses by students is similar to that of managers as reported by Rynes et al. (2002), showing that common beliefs and scientific evidence are often inconsistent.”

In the fourth session, a guest speaker, who completed a PhD in organizational behavior and then became a human resource director, discusses with students pros and cons of practicing EBMgt and presents a case from his own experience. Through this guest speaker, students gain a window into the reality of EBMgt, for example, that he also has learned that at times his beliefs were at odds with the data. EBMgt, however, has helped him to make better decisions, which has led him to cherish rather than bemoan situations in which the data contradicted his assumptions. The guest speaker also reports of challenges, such as resistance by colleagues to evidence-based practices, suggesting that managers also need to know how to cope with such attitudes.

Main Phase: Practicing the Four Steps of the Problem-Solving Cycle

With the exception of the last session, students work in the remaining nine sessions on cases to practice the evidence-based problem-solving cycle or parts of it, whereby we emphasize a particular step in each class. We focus on step 1 (problem definition) and step 2 (problem analysis) in one session each, followed by four sessions that primarily address step 3 (testing solutions) and two sessions on step 4 (evaluation of tested solutions). Below, we use cases from our course to illustrate the learning of each step without presenting the full cycle for these cases as we have done for the bike messenger and the towel reuse case.

Step 1: Problem Definition:

What Is the Problem, If Any, That I Would Like to Solve? What Are the Key Outcome Variables?

Students learn to define problems as measurable discrepancies. For an evidence-based manager, defining a problem amounts to determining the outcome variables (or the dependent variables) and their assessment. In addition, evidence-based managers must critically evaluate the problem definition. For example, how would other stakeholders define the problem? Does the problem definition imply a violation of pertinent (e.g., organizational) values? Is the problem limited to an individual or does it pertain to larger units? Would the problem definition change if other disciplinary perspectives had been adopted (e.g., an economic point of view vs. a sociological point of view)?

We use a case study on an employee who underperforms after having taken maternity leave (Gandz & Spracklin, 2007) and provide academic readings on human resource management and job performance (e.g., Boxall & Purcell, 2003). Students struggle to arrive at a problem definition, although (or because) the definition as a discrepancy between actual and previously agreed-upon job performance seems evident. Instead, students commonly confound symptoms and causes of the problem, for example, suggesting that the problem should be defined as a lack of motivation rather than a lack of performance. To stimulate reflection, the facilitator might ask students to assume the role of a doctor who first has to know the symptoms of an illness before uncovering the cause and prescribing a treatment. Once the definition of the problem as a performance discrepancy is in place, students see its advantages: Now causes

and effects are separated from each other, and effects of potential treatments can be assessed. Facilitators also use questions for allowing students to further probe the definition. For example, do all employees in the unit underperform or do all employees who took pregnancy leaves underperform?

Step 2: Problem Analysis:

Why Does the Problem Occur? What Are Variables That Could Be Potential Causes of the Outcome? What Information Does Existing Evidence Provide for Analyzing the Problem? What Are Potential Solutions?

In this step, evidence-based managers consult many sources about problem causes and solutions, including their own or others' experiences, internal company information, and external evidence, such as academic journals. To illustrate the learning of this step, we present below another hotel case (see class 6 in Table 1).

The Room Cleanliness Case. *One of us once was asked by a hotel to raise the motivation of its room attendants whose poor performance had led to many customer complaints. Management had already introduced more room checks, a standard procedure in the best hotels, and one that, according to a manager, was necessary. In a meeting, room attendants complained about their picky supervisors and a lack of time to clean the rooms. They explained that they tussled over the room trolleys that were used to transport cleaning materials: Some trolleys were larger and easier to maneuver. On a usual workday, which lasted until 5:00 p.m., the attendants had to clean the rooms at full speed to finish by 3:00 p.m. so that supervisors could do their checks, and give attendants enough time to address the shortcomings that supervisors had identified. In a subsequent meeting to analyze the situation, the hotel managers received information on scientific frameworks (e.g., Boxall & Purcell, 2003) that conceive of job performance as being a function of ability, motivation, and external constraints.*

The Room Cleanliness case illustrates challenges in uncovering existing evidence and conducting systematic analyses. Students acknowledge that managers seemingly considered the problem to be a lack of room attendants' motivation. Yet, many students prefer a definition as a gap between actual and desired room cleanliness because it allows for a more comprehensive analysis of causes and solutions (i.e., a lack of motiva-

tion is only one possible reason for a lack of room cleanliness). Another challenge lies in identifying immediately problem-relevant evidence, such as studies on the causal determinants of room attendants' performance or those of room cleanliness. Academic frameworks for analyzing employee performance, however, do exist. For example, Boxall and Purcell (2003) suggested an ability–motivation–opportunity model. The rooms might not be clean because the room attendants do not have the necessary ability, are not motivated, or do not have the means (e.g., poor equipment) to service them.

Partly due to the paucity of directly relevant scientific evidence, the involvement of stakeholders during the first two steps of the cycle is critical, and the consultants in the case spoke with room attendants, their supervisors, and managers. The input from stakeholders might trigger further literature search, for example, on the impact of monitoring on job performance. Typically, students settle on faulty equipment and a lack of time as key causes of the room cleanliness problem. Ultimately, on the basis of the information provided by the stakeholders and through the existing evidence (albeit seemingly sparse), students arrive at potential solutions, such as replacing smaller trolleys by larger ones and rearranging the room checks by supervisors. The facilitator challenges the students to assure that these potential solutions have predictive power, show fit with stakeholder concerns, and are testable and feasible.

In debriefing, the facilitator points to the importance of executing the first two steps of the evidence-based problem-solving cycle as properly as possible—as difficult as the reality of doing so might be. Discussing the differential impact of defining the problem as lack of motivation versus lack of cleanliness on the likely solution to the problem powerfully illustrates this argument. The facilitator also points to the need to arrive at potential solutions even when the existing evidence seems sparse and stakeholders contradict each other.

Step 3: Testing of Solutions: How Does One Test the Proposed Solution or Intervention?

"Do we think this is true? Or do we know?" This is a question that Gary Loveman, CEO of *Caesars Entertainment*, poses (Davenport, 2006). If evidence-based managers want to know whether prospective solutions are truly effective in their companies, they need to conduct causally interpretable tests on-site to produce local evidence. While our students are asked to

design and evaluate tests of solutions throughout the semester (and execute their own test in the EBMgt project), four cases are used to zoom in on testing (see sessions 7–11 in Table 1). We use a case based on Greenberg's (1990) theft study as an illustration. This case is taught with an emphasis on quasi-experiments after students have already worked on two other cases that focus on the design of true field experiments.

The Theft Case. *Professor Smith could not believe this unique opportunity. Because of the loss of two large contracts, a manufacturing company had to cut salaries in two of its three plants. Smith would be allowed to study the effects of these pay cuts on employee theft. Theories of distributive justice proposed that employees would react negatively to pay cuts and, for example, make up for the lost pay by stealing inventory. The particular interest of Smith, however, was in the effects of adequate versus inadequate explanations of these pay cuts on employees' reactions. Theories of procedural justice suggested that proper explanations of pay cuts by management might attenuate negative reactions. The company had already selected the two plants where the pay would be cut. Now Smith had to develop a test that met his high academic standards while also being feasible for the company and allowing the managers to learn about mitigating negative reactions to pay cuts.*

The problem in this case can be defined as employee theft, assessed by inventory shrinkage rates. This outcome variable is suggested already in the case, but several other alternative assessments of negative reactions to injustice are possible (e.g., other deviant behaviors). The solution in the form of adequate explanations, which is well-grounded in existing evidence (Greenberg, 2009), is also mentioned in the case. In opening the discussion about the design of the test, we use two questions by Angrist and Pischke (2008: 3): "What is the causal relationship of interest?" and what experiment could ideally be used to test the causal effect under scrutiny? Students recognize that the causal relationship concerns effects on employee theft of pay cuts and adequate explanations. The ideal experiment would manipulate pay cuts (present/absent) and adequateness of the explanation (yes/no) and assess their impact on employee theft, and employees would be randomly assigned to conditions.

Having designed the ideal experiment, students move on to designing the actual study. Because of the preselection of plants where the pay would be cut, a random assignment of employees to pay cut

conditions is not possible. Therefore, unknown differences across the plants, in addition to the pay cut manipulation, might also cause effects on employee theft. A pre-post design in which theft is assessed before, during, and after the experiment allows for controlling for systematic differences to the extent that they already exist prior to the study. In addition, differences between the plants (e.g., demographic profile, size, and location-specific variables) might be assessed. If students do not see the challenges arising from a lack of random assignment, we ask them about the consequences for the study if only men worked in one plant and only women in the other.

Having discussed potential differences among plants due to the impossibility of random assignment, students then turn to the manipulated variable: the adequateness of explanations. Because there are only two plants where the pay is cut, it is obvious that an adequate explanation will be given in one plant, but not the other. The question, however, arises as to how to design the control condition: Should no explanation at all be given or should an explanation, but a less adequate one, be given? Eventually, students usually agree that in both conditions the same top manager should give an explanation to avoid confounding between adequateness of explanations and presence of a top manager.

By now, students have covered the basics of the problem-testing step, having determined the outcome variable and its assessment, the causal variables, control variables, and the study design. They can now launch into details of the test execution, in particular how the adequate explanation variable should be operationalized. Students discuss whether (a) only one aspect of an adequate explanation should be varied, which would allow for specific causal inferences, or (b) several aspects of an adequate explanation should be varied, which would enhance chances that an effect might be found. In addition, they argue about the measurement of the outcome variable (theft); for example, the frequency of measurement and whether other outcome variables (e.g., perceptions of equity) should also be assessed. Another issue, the potential loss of participants over time, is typically addressed by students only after they have been prompted. Then they see that another reaction to a pay cut might be voluntary resignations, particularly if the pay cut was not adequately explained.

In summary, The Theft Case illustrates the teaching of step 3 (i.e., testing solutions). We start with identifying the causal relationship on the basis of steps 1 and 2 of the evidence-based problem-solving cycle before establishing the ideal experiment for testing this relationship. Then we move on to determining the study design (e.g., a between- or within-participants design) and finally go into details, such as operationalizing experimental conditions. Across the four sessions regarding testing of solutions, students are confronted not only with experiments and quasi-experiments, but also nonexperimental studies, which if properly designed can be used to establish causal links (Antonakis et al., 2010). We consistently retain a focus on causal interpretability, emphasizing that managers should pose causal questions when they reflect about potential interventions ("if I do X, will Y change?").

Step 4: Evaluation of Tested Solutions: Do They Have the Expected Effect?

Across two sessions, we highlight five aspects of evaluating locally tested solutions: *causal interpretability*, *effectiveness*, *efficiency*, *robustness across stakeholders and frames*, and *sufficiency*. A sixth aspect, *robustness across contexts*, becomes relevant if tested solutions are also to be implemented on sites other than the test site. *Effectiveness* refers to the reduction of the discrepancy identified in the problem definition. When being paid a higher commission, did the bike messengers work the desired number of shifts per week or at least a higher number of shifts than before? *Efficiency* takes the costs and the benefits of an intervention into account. *Robustness across stakeholders and frames* refers to the extent to which the solution creates value for all stakeholders (e.g., financial, social, and ethical acceptability) and draws on different disciplines. Last, *sufficiency* of a tested solution must not prevent managers from seeking better solutions. A review should be conducted after a full run through the cycle.

Two cases focus on the evaluation of tested solutions. The Kenexa case (Dietz & Joshi, 2007) describes a linkage research project in a retail bank. Students receive the data to assess the effects of service climates on customer satisfaction. Using Excel spreadsheets, they run correlation and regression analyses and have to judge managerial success on the basis of statistical parameters. The

Ducati case (Gino & Pisano, 2006) allows the comparison of different testing strategies (e.g., simulations vs. tests of prototypes and laboratory vs. field tests) and their respective outcomes.

Then, during the last session, student teams briefly present their EBMgt projects. The range of topics is remarkable, including marketing projects (such as the impact of product labeling on evaluations of product quality), operations management projects (e.g., the charitable donation project or a project on reducing the use of plastic bags in a supermarket chain), and HR projects (e.g., the impact of frequent updates on turnover of members in a student association or the effect of a training on the use of rhetoric communication tactics on ratings of charisma).

After the presentations, a class discussion follows in which students discuss differences between the cases and the project. Students acknowledge that the project is initially a big fuzzy challenge despite the clear structure of the evidence-based problem-solving cycle. They also note the labor intensity and difficulty of the data collection process. Finally, students report on their key learning points from the course, such as the value and power of empirical evidence for executing managerial interventions.

By the end of the course, students have worked on 10 cases and the EBMgt project. Time after time, they have practiced in full or partially the evidence-based problem-solving cycle, helping them to acquire a structured approach to address business issues and to sharpen their critical thinking (cf. Rousseau & Barends, 2011). Through repeated practice and through comparisons across cases and between cases and the project, students develop a repertoire on the uses of EBMgt. Below, we elaborate how we assess students' learning and the course.

Assessment of Students and the Course

Assessment has to fit the approach and, hence, must capture the students' practice of EBMgt. In our course, students are assessed on two individually conducted case analyses and the team-based EBMgt project. Grading criteria involve the mastery of each step of the evidence-based problem-solving cycle. For example, for step 1 (problem definition), we verify that the problem (1) refers to the effect (and avoids confounding causes and effects); (2) is formulated as a gap between an actual and desired state; (3) is quantifiable; (4) can be operationalized, and (5) spec-

ifies the level of analysis. Further checks include that the definition considers stakeholder concerns and other disciplinary perspectives. In addition to assessing students' performance for each step, the case analyses and the project are evaluated in their entirety with regard to coherence across the four steps, usefulness to the company, and awareness of limitations.

Ideally, EBMgt courses would be assessed in terms of their effects on the performance of students as evidence-based managers, for example, in a longitudinal cohort study over several years and relative to a comparable control group of students who did not take the course (e.g., using a propensity score analysis or a treatment effects model). Proxies would include the extent to which managers execute projects using the evidence-based problem-solving cycle, whether their company adopts and benefits from EBMgt, and whether EBMgt skills translate into career success (e.g., promotion speed, hierarchical level, salary).

As our course is still in its nascent stage, the data on it are sparse. Quantitative evaluations by students are positive, but what causes this feedback is not clear. The qualitative data fall along two dimensions: (1) relevance and (2) pedagogy. Students, for example, appreciate that the course is relevant to strengthening their critical thinking skills. It also prepares them for their thesis projects, albeit there is only an occasional not a mandatory connection to course projects. Students mostly laud the pedagogy, such as the degree of interaction and the variety of cases, although some criticize the workload. Some alumni report an increase in their confidence in questioning the evidence for consultants' suggestions for organizational changes. For the next version of the course, we will collect premeasures (see Jelley et al., 2012) and implement a design with a control group.

For assessing the effects of PBL in general on the use of evidence-based practices, research on student samples (e.g., students of medicine) and practitioner samples (e.g., physicians) is available. Meta-analyses (e.g., Albanese & Mitchell, 1993; Dochy, Segers, Van den Bossche, & Gijbels, 2003; Vernon & Blake, 1993) have indicated positive effects of PBL (vs. traditional approaches) on clinical performance, as measured by the treatment of real and simulated patients and by scores on tests of clinical knowledge. Vernon and Blake also mentioned several studies according to which PBL-trained students made more indepen-

dent and frequent use of academic sources than did traditionally trained students. For physicians, a review of studies by Koh, Khoo, Wong, and Koh (2008) showed positive weak effects for problem-based learning (vs. traditional approaches) for the use of information resources, research skills, and understanding of evidence-based medicine. The few studies on PBL in business and economics (e.g., Maxwell, Mergendoller, & Bellisimo, 2005; Mergendoller, Maxwell, & Bellisimo, 2000; Son & van Sickle, 2000; van den Bossche, Segers, Gijbels, & Dochy, 2004) are consistent with these findings.

Most of the research on PBL, however, is based on quasi-experimental studies, in which participants were not randomly assigned to treatments. The better performance of PBL-trained students and physicians might have resulted from their training or have been due to selection effects. Moreover, the operationalizations of PBL, traditional learning approaches, and outcome variables varied starkly across studies, making it difficult to isolate which PBL practices contributed to effective learning. In sum, whereas the causal interpretability of the existing evidence and its applicability to our context have to be established, the evidence indicates that PBL has positive albeit often weak effects on learning evidence-based practices.

DISCUSSION

We have advocated an approach to teaching EBMgt that focuses on producing local evidence. The approach uses a variant of the PBL method, and its key teaching tools are cases of actual business problems and a group project that emphasizes the design, execution, and analysis of experimental tests. As a first contribution, our approach illustrates one way that students can become producers of causally interpretable research; whereas numerous approaches to teaching EBMgt merely aim to make students "savvy consumers" (e.g., Jelley et al., 2012: 341) of it. Our approach emphasizes producing local evidence ("little e evidence") in addition to searching, evaluating and applying scientific evidence ("big E evidence"). Producing evidence allows students to (1) better understand existing theory, which is otherwise often hard to digest, (2) see the challenges of implementing even seemingly straight-forward theories, and, (3) better judge the quality of existing evidence. To use an analogy, to become a surgeon, one can read books

and articles on the best evidence in the library. Or, one can practice on corpses in the morgue. Studying books and cutting up corpses provide complementary insights, and medical training does not only require the absorption of literature, but also mandates hands-on practice for good reasons.

As a second contribution, our approach shows how to teach academic research and EBMgt with cases, contrary to the position that, as Shugan (2006: 1009) claimed, case teaching "helps destroy the link between academic research and classroom learning." This is not necessarily the case, as we have illustrated with four examples on how to intertwine teaching academic research with teaching cases of actual business problems. The teaching process, not the cases per se, determine the impact of academic research on students. Key to our approach is that cases are tackled by taking advantage of existing evidence and of applying scientific methods. Indeed, our students do not just read and analyze case studies, but in executing an evidence-based problem-solving cycle they search for scientific articles, evaluate the evidence presented in these articles, and design studies.

If instructors teach EBMgt only with cases, they face several challenges. For example, students might deal with the complexity of cases through oversimplification (e.g., mono-causality and single-criterion evaluations). Instructors can evoke comparisons of cases to the EBMgt project, in which errors of oversimplification become more readily apparent to students. Similarly, instructors must ensure that students avoid the mistake of false generalizations from single cases. The cases are used to illustrate the application of scientific theories, methods, and evidence to the solution of business problems, but they do not allow the discovery of general principles from individual business problems.

Despite the value of teaching EBMgt with a focus on producing local evidence, our approach is not without limitations. One concern is that students learn relatively little about identifying and evaluating existing evidence. At our school, other courses (e.g., on organizational behavior and marketing) introduce students to topic-relevant scientific content. These courses complement, as would those described by Erez and Grant (2014), our course, which largely trains students on scientific methods for managerial use. Another concern is that our approach requires basic statistical knowledge, which our students learn in another class. Erez and Grant also describe how they introduce

students to scientific methods in a few sessions, while we do so over an entire course.

Last but not least, the question arises whether our approach can be taught across business schools. We have successfully taught a version with an even narrower focus on experimental logic at a less research-oriented business school. The purpose was to provide students with a different lens on critical thinking while solving business problems. Another possible adaptation is to reduce the scope of the project, for example, by requiring the design but not the execution of studies. Instead, students might participate as experimenters and subjects in small in-class experiments; for example, to determine the effect of an organic label on the perceived taste of orange juice.

To conclude, we have presented an approach to teaching EBMgt that allows connecting scientific methods and academic research with actual business problems. This approach brings academic research to life, builds skills for evidence-based problem-solving, as well as for designing and executing local tests. Ultimately, once students master the evidence-based problem-solving cycle, they are in a position not only to benefit from the rigor of scientific evidence and methods, but also to produce relevant and causally interpretable evidence by themselves. And it is perhaps only then that they start to understand what the term "management" can really mean.

REFERENCES

- Albanese, M. A., & Mitchell, S. 1993. Problem-based learning: A review of literature on its outcomes and implementation issues. *Academic Medicine*, 68(1): 52–81.
- Angrist, J. D., & Pischke, J.-S. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton: Princeton University Press.
- Angrist, J. D., & Pischke, J.-S. 2010. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24(2): 3–30.
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2010. On making causal claims: A review and recommendations. *The Leadership Quarterly*, 21(6): 1086–1120.
- Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. 2014. Causality and endogeneity: Problems and solutions. In D. V. Day (Ed.), *The Oxford handbook of leadership and organizations*: 93–117. New York: Oxford University Press.
- Baer, M., Oldham, G. R., & Cummings, A. 2003. Rewarding creativity: When does it really matter? *The Leadership Quarterly*, 14(4): 569–586.
- Bailey, D., & Kurland, N. 2002. A review of the telework research:

- Findings, new directions and lessons for the study of modern work. *Journal of Organizational Behavior*, 23: 283–400.
- Bandiera, O., Barankay, I., & Rasul, I. 2011. Field experiments with firms. *Journal of Economic Perspectives*, 25: 63–82.
- Baron, N. J., & Kreps, D. M. 1999. Consistent HR practices. *California Management Review*, 41(3): 29–53.
- Barrows, H. S. 1996. Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*, 68: 3–12.
- Bergh, D. D., Hanke, R., Balkundi, P., Brown, M., & Chen, C. 2004. An assessment of research designs in strategic management research: The frequency of threats to internal validity. In D. J. Ketchen & D. D. Bergh (Eds.), *Research methodology in strategy and management*: 349–371. Oxford: Elsevier.
- Bloom, N., Liang, J., Roberts, J., & Ying, Z. J. 2013. *Does working from home work? Evidence from a Chinese experiment*. NBER Working Paper no. 18871. Cambridge, MA: National Bureau of Economic Research.
- Boxall, P., & Purcell, J. 2003. *Strategy and human resource management*. New York: Palgrave Macmillan.
- Bransford, J. D., Brown, A. L., & Cocking, R. R. 2000. *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academies Press.
- Briner, R. B., Denyer, D., & Rousseau, D. M. 2009. Evidence-based management: Concept cleanup time? *Academy of Management Perspectives*, 23: 19–32.
- Brittain, J. W., & Sitkin, S. B. 2006. *Carter racing*. Chapel Hill, NC: Delta Leadership.
- Byron, K., & Khazanchi, S. 2012. Rewards and creative performance: A meta-analytic test of theoretically derived hypotheses. *Psychological Bulletin*, 138: 809–830.
- Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. 1993. A theory of performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations*: 35–70. San Francisco: Jossey-Bass.
- Childress, S., & Marietta, G. 2008. *A problem-solving approach to designing and implementing a strategy to improve performance*. Boston, MA: Public Education Leadership Project at Harvard University.
- Cialdini, R. B., & Goldstein, N. J. 2004. Social influence: Compliance and conformity. *Annual Review of Psychology*, 55: 591–621.
- Davenport, T. H. 2006. Competing on analytics. *Harvard Business Review*, 84(1): 98–107, 134.
- Davenport, T. H. 2009. Make better decisions. *Harvard Business Review*, 87(11): 117–123.
- Denrell, J. 2005. Selection bias and the perils of benchmarking. *Harvard Business Review*, 83(4): 114–119.
- Dietz, J., & Joshi, C. 2007. *Kenexa*. Ivey Publishing: 9B07C04. London, Canada: Western University.
- Dietz, J., Pugh, S. D., & Wiley, J. W. 2004. Service climate effects on customer attitudes: An examination of boundary conditions. *Academy of Management Journal*, 47(1): 81–92.
- Dochy, F., Segers, M., Van den Bossche, P., & Gijbels, D. 2003. Effects of problem-based learning: A meta-analysis. *Learning and Instruction*, 13(5): 533–568.
- Erez, A., & Grant, A. M. 2014. Separating data from intuition: Bringing evidence into the management classroom. *Academy of Management Learning and Education*, 13(1): 104–119.
- Fehr, E., & Goette, L. 2007. Do workers work more if wages are high? Evidence from a randomized field experiment. *American Economic Review*, 97(1): 298–317.
- Ferrari, M., & Mahalingam, R. 1998. Personal cognitive development and its implications for teaching and learning. *Educational Psychologist*, 33(1): 35–44.
- Gandz, J., & Spracklin, E. 2007. *Elise Smart*. Ivey Publishing: 9B03C010. London, Canada: Western University.
- Gavin, D. A. 2013. How Google sold its engineers on management. *Harvard Business Review*, 91(12): 74–81.
- Ghoshal, S. 2005. Bad management theories are destroying good management practices. *Academy of Management Learning and Education*, 4(1): 75–91.
- Gino, F., & Pisano, G. P. 2006. *Ducati case: The making of a Grand Prix motorcycle*. Harvard Business School Case 605-090. Boston, MA.
- Gino, F., & Pisano, G. P. 2011. Why leaders don't learn from success. *Harvard Business Review*, 89: 68–74.
- Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. 2008. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, 35: 472–482.
- Goodman, J. S., & O'Brien, J. 2012. Teaching and learning using evidence-based principles. In D. M. Rousseau (Ed.), *The Oxford handbook of evidence-based management*: 309–336. New York: Oxford University Press.
- Greenberg, J. 1990. Employee theft as a reaction to underpayment inequity: The hidden cost of pay cuts. *Journal of Applied Psychology*, 75(5): 561–568.
- Greenberg, J. 2009. Everybody talks about organizational justice, but nobody does anything about it. *Industrial and Organizational Psychology*, 2(2): 181–195.
- Groysberg, B., Lane, D., & Knoop, C. I. 2007. People management. Harvard Business School Teaching Note 407-016.
- Harms, P. D., & Credé, M. 2010. Remaining issues in emotional intelligence research: Construct overlap, method artifacts, and lack of incremental validity. *Industrial and Organizational Psychology*, 3: 154–158.
- Hmelo-Silver, C. E. 2004. Problem-based learning: What and how do students learn? *Educational Psychology Review*, 16: 235–266.
- Hodgkinson, G. P., & Rousseau, D. M. 2009. Bridging the rigor-relevance gap in management research: It's already happening! *Journal of Management Studies*, 46: 534–546.
- Hong, Y., Liao, H., Hu, J., & Jiang, K. 2013. Missing link in the service profit chain: A meta-analytic review of the antecedents, consequences, and moderators of service climate. *Journal of Applied Psychology*, 98: 237–267.
- Jelley, R. B., Carroll, W. R., & Rousseau, D. M. 2012. Reflections on teaching evidence-based management. In D. M. Rousseau (Ed.), *The Oxford handbook of evidence-based management*: 337–355. New York: Oxford University Press.

- Johns, G. 2006. The essential impact of context on organizational behavior. *Academy of Management Review*, 31(2): 386–408.
- Johnson, E. J., & Goldstein, D. 2003. Do defaults save lives? *Science*, 302: 1338–1339.
- Johnson, S. M., & Marietta, G. 2009. Taking human resources seriously in Minneapolis. Public Education Leadership Project at Harvard University, PEL No.-055.
- Joshi, C., Bapuji, H., & Chandrasekhar, R. C. 2013. *Windermere manor: Sustainability and change*. Ivey Publishing: 9B13C044. London, Canada: Western University.
- Kepes, S., & McDaniel, M. 2013. How trustworthy is the scientific literature in I-O psychology? *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 6(3): 252–268.
- Kerr, S. 1995. An academy classic: On the folly of rewarding A while hoping for B. *Academy of Management Executive*, 9(1): 7–16.
- Koh, G. C.-H., Khoo, H. E., Wong, M. L., & Koh, D. 2008. The effects of problem-based learning during medical school on physician competency: A systematic review. *Canadian Medical Association Journal*, 178: 34–41.
- Kovner, A. R., Fine, D. J., & D'Aquila, R. 2009. *Evidence-based management in healthcare*. Chicago, IL: Healthcare Administration Press.
- Lal, R., & Carrolo, P. 2004. *Harrah's Entertainment Inc.* Harvard Business School Case 502-011.
- March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2: 71–87.
- Maxwell, N. L., Mergendoller, J. R., & Bellisimo, Y. 2005. Problem-based learning and high school macroeconomics: A comparative study of instructional methods. *Journal of Economic Education*, 36(4): 315–329.
- Mergendoller, J. R., Maxwell, N. L., & Bellisimo, Y. 2000. Comparing problem-based learning and traditional instruction in high school economics. *Journal of Educational Research*, 93(6): 374–382.
- Mientka, M. 2013. With lawyers and junk science, big tobacco sought to influence regulatory policy. *Medical Daily*. Retrieved from <http://www.medicaily.com/articles/15944/20130528/tobacco-industry-institute-of-medicine-public-health-policy-smoking-regulation.htm>. June 28, 2013.
- Nutt, P. C. 1999. Surprising but true: Half the decisions in organizations fail. *Academy of Management Executive*, 13(4): 75–90.
- Pfeffer, J., & Sutton, R. I. 2006. Evidence-based management. *Harvard Business Review*, 84(1): 62–84.
- Prince, M. 2004. Does active learning work? A review of the research. *Journal of Engineering Education*, 93(3): 223–231.
- Reichers, A. E., & Schneider, B. 1990. Climate and culture: An evolution of constructs. In B. Schneider (Ed.), *Organizational climate and culture*: 5–39. San Francisco: Jossey-Bass.
- Rosenberg, W., & Donald, A. 1995. Evidence based medicine: An approach to clinical problem-solving. *British Medical Journal*, 310(6987): 1122–1126.
- Rosenzweig, P. 2007. Misunderstanding the nature of company performance: The halo effect and other business delusions. *California Management Review*, 49(4): 6–20.
- Rousseau, D. M. 2006. 2005 Presidential Address: Is there such a thing as “evidence-based management”? *The Academy of Management Review*, 31(2): 256–269.
- Rousseau, D. M., & Barends, E. G. R. 2011. Becoming an evidence-based HR practitioner. *Human Resource Management Journal*, 21(3): 221–235.
- Rousseau, D. M., & Fried, Y. 2001. Location, location, location: Contextualizing organizational research. *Journal of Organizational Behavior*, 22(1): 1–13.
- Rynes, S. L., Colbert, A. E., & Brown, K. G. 2002. HR professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management*, 41(2): 149–174.
- Rynes, S. L., Gerhart, B., & Parks, L. 2005. Personnel psychology: Performance evaluation and pay for performance. *Annual Review of Psychology*, 56(1): 571–600.
- Schneider, B., Parkington, J. J., & Buxton, V. M. 1980. Employee and customer perceptions of service in banks. *Administrative Science Quarterly*, 25: 252–267.
- Schultz, P. W., Khazian, A., & Zaleski, A. 2008. Using normative social influence to promote conservation among hotel guests. *Social Influence*, 3: 4–23.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. 2002. *Experimental and quasi-experimental designs for generalized causal inference*: 1–33. New York: Houghton Mifflin.
- Shugan, S. M. 2006. Editorial: Save research—Abandon the case method of teaching. *Marketing Science*, 25(2): 109–115.
- Smith, G. F. 1989. Defining managerial problems: A framework for prescriptive theorizing. *Management Science*, 35(8): 963–981.
- Son, B., & Van Sickle, R. 2000. Problem-solving instruction and students' acquisition, retention and structuring of economic knowledge. *Journal of Research and Development in Education*, 33(2): 95–105.
- Thaler, R., & Sunstein, C. R. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. London: Penguin.
- Thun, M. J., Jacobs, E. J., & Patrono, C. 2012. The role of aspirin in cancer prevention. *Nature Reviews. Clinical Oncology*, 9(5): 259–267.
- Tilly, W. D. III. 2008. The evolution of school psychology to science-based practice. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology V*: 18–32. Washington, DC: National Association of School Psychologists.
- Van den Bossche, P., Segers, M., Gijbels, D., & Dochy, F. 2004. Effects of problem-based learning in business education: A comparison between a PBL and a conventional educational approach. In R. Ottewill, L. Borredon, L. Falque, B. Macfarlane, & A. Wall (Eds.), *Educational innovation in economics and business*: 205–227. Dordrecht, The Netherlands: Springer.
- Vernon, D. T., & Blake, R. L. 1993. Does problem-based learning work? A meta-analysis of evaluative research. *Academic Medicine*, 68(7): 550–563.
- Weick, K. E. 1989. Theory construction as disciplined imagination. *The Academy of Management Review*, 14(4): 516–531.
- Zaccaro, S. J., & Horn, Z. N. J. 2003. Leadership theory and practice: Fostering an effective symbiosis. *The Leadership Quarterly*, 14(6): 769–806.

Joerg Dietz is professor of organizational behavior at the Faculty of Business and Economics, University of Lausanne. Dietz obtained his doctorate in business administration from Tulane University's A. B. Freeman School of Business. His research interests include workforce diversity, service climates, and contextual antecedents of organizational behavior.

John Antonakis is professor of organizational behavior at the Faculty of Business and Economics, University of Lausanne. Antonakis obtained his doctorate in applied management and decision sciences from Walden University. He currently does research on leadership, individual differences, and research methods.

Ulrich Hoffrage is professor of decision theory at the Faculty of Business and Economics, University of Lausanne. Hoffrage received his Dr. phil. from the University of Salzburg. He is interested in simple heuristics, bounded rationality, ethical decision making, consumer decision making, and group decision making.

Franciska Krings is professor of organizational behavior in the Faculty of Business and Economics, University of Lausanne. She received her doctorate in psychology from the University of Bern. Krings' primary research interests focus on diversity and discrimination, with issues related to age and generations, immigrants, and gender.

Julian N. Marewski is assistant professor at the Faculty of Business and Economics, University of Lausanne. Marewski obtained his doctorate from Free University Berlin in 2009 and was a postdoctorate and research scientist at the Max Planck Institute for Human Development (Berlin). His research focuses on decision-making processes.

Christian Zehnder is professor of organizational decision making at the Faculty of Business and Economics, University of Lausanne. Zehnder received his doctorate in economics from the University of Zurich. His research is in the field of behavioral and experimental economics and combines insights from psychology and economics.