Experimental Methods in International Management Research

Anne Bartel-Radic, Philippe Mouillot, Danielle A. Taylor

Bartel-Radic, Anne, Mouillot, Philippe, & Taylor, Danielle (2019) Experimental Methods in International Management Research.
In : Bartel-Radic, Anne (coord.) *Méthodes de recherche innovantes et alternatives en économie et gestion / Innovative and alternative research methods in economics and Business Administration*. Luxemburg : EIKV, pp. 61-75.

Experimental Methods in International Management Research

ABSTRACT

Experimental methods have long been used in the natural sciences to test a hypothesis by controlling experimental conditions. These methods have also been used in economics and management in recent decades, and have become a reference in some sub-disciplines such as marketing. However, experiments are rare in international management research. In this chapter, we present the difficulties, but also the advantages of the experimental methodologies in this sub-discipline through the example of the InterCCom project, which deals with intercultural competence. Thanks to the development of a computerized serious game for this project, the behaviors of the members of a virtual, international team are measured. Various research questions are outlined and answered.

KEYWORDS: Experimental methods, experimentation, international management, intercultural competence

INTRODUCTION

The purpose of this chapter is to provide researchers and social science students with an introduction to experimental methodologies. These methodologies originally come from the natural sciences such as physics, chemistry or medicine. In recent decades, they have been widely used in economics and management science, particularly in the sub-discipline of marketing. On the other hand, in international management, contributions remain relatively rare, whereas the best journals of the discipline have called for a significant increase of this research.

This chapter is structured in four parts:

1. Clinical experimentation as used in medicine or neuroscience.
2. The establishment and use of experimental methods in economics and management science.
3. Application to the international management field and the inherent difficulties in this sub-discipline.
4. The presentation of a research project in intercultural management, and how it mobilizes experimental methods.

1. EXPERIMENTAL METHODS: SOME FUNDAMENTALS

In a broad sense, any empirical data collection that aims to test a hypothesis, whatever the mode of data collection (observation, interviews, focus group, or questionnaire), can be described as experimentation. The hypotheses to be tested include links between one or more causes, and a consequence. A link between two phenomena can be qualified as causal if the following criteria are met (Hill, 1965):

- the correlation must be strong;
- this correlation must be shown in different contexts, in studies conducted on different populations;
- the cause should have specific consequences;
- the cause should precede the consequence in time;
- it must be possible to find a plausible explanation of how the cause influenced the consequence;
- and, in general, it must be possible to find a monotonic function between the cause and the consequence (*more* causes should result in *more* consequences).

There are various experimental methods including "real" experimental designs, quasi-experimentation and other data collection techniques. According to Greenwood (1972: 178), true experimental designs must meet the following criteria:

- they aim to test a hypothesis;
- this test goes through a precise and controlled research protocol;
- and the experimental conditions are controlled using variable elimination techniques, constant conditions and random sampling.

Only research designs that include two measures, such as before and after an event or treatment proposed to the test group (e.g. inclusion of a variable), and that have a control group that has not undergone the same treatment abide by these conditions. Thus, in pharmacology, for example, a drug meant to cure a certain pathology can be tested through a process where the test group of patients is given the drug while the control group is given a placebo. Both groups are given an evaluation of their state of health before and after the treatment.

Thus, we distinguish the experimental protocols that take place in someone's normal environment ("field experiment") from experimental protocols that take place in a dedicated place such as a laboratory (i.e. "clinical experimentation"). Studies taking place in a person's normal environment have a higher external validity because people's behaviors will be less biased. But not only are these field trips costlier in terms of time and resources, they also do not allow controlling contextual variables with the same precision. Clinical experiments are thus usually given greater internal validity, even if they also have limitations (Schnell, Hill & Esser, 2018).

The very foundation of clinical experimentation rests on three essential pillars:

- An artificial reproduction that is as faithful as possible to the environment that the experiment is testing.
- A judicious choice of tools and methods to mobilize to achieve one's goals.
- A research question that is clear, measurable and relevant.

For the first pillar to be stable, a noted technique in experimentation is that reality should first be observed at length. Indeed, in terms of intercultural management, immersion in the country's culture, frequent sharing with its representatives and specific work on stereotypes can prove to be very effective. For the second pillar to be stable, it is sufficient to justify each decision; to do so, we should explain the reasoning for each choice by constantly employing the word “because.” As for the third pillar, only one question merits recurrence: what do we want to measure?

Quasi-experimental designs are qualified as research designs in which an experimental group and a control group are compared, but where the classification of an individual in one group or the other is not randomly distributed. Generally, the observed individuals choose the group they want to belong to, which creates biases. Within these quasi-experimental designs, we distinguish “natural experiments” from “quasi-experiments.” In the former, manipulation of the variable occurs without action by the researcher (such as the decision to reduce speeds from 90 km/h to 80 km/h on national roads in France in 2018), while in the latter, the researcher deliberately manipulates the variable (by introducing a new public policy in only one or a few geographical areas that are expressly chosen to verify its impact, for example).

The use of experimental methods in the social sciences faces a number of limitations. We need very large samples (for example 1500 people if we want to be able to generalize the results to the population of a country like Germany, Schnell, Hill & Esser, 2018: 203). Samples should be comprised of people ideally chosen at random, which requires significant resources. Furthermore, and most importantly, many variables studied in the social sciences, such as culture, intelligence, or socio-professional categories cannot easily be manipulated during the experiment, and especially do not allow people to be randomly assigned to one of the two groups.

Take the case of a measure of intelligence. The general factor (or g factor - Spearman, 1904) consists of agreeing on a conceptualization of the object of research, both in its definition and in its recurrent structure, in order to be able to quickly identify any dissonance. Here, it is essential to distinguish the subject capable of solving a problem because s/he has already solved it, because s/he has observed its solution, or because s/he is using his/her unique intelligence. Thus, we distinguish memory, learning, and “pure” intelligence (Mouillot, Drillon, & Montargot, 2018). In a study on intelligence, we could lament that the only factor tests still designed today are matrices. But the matrices do not eliminate the cultural aspect; they implement many cognitive dimensions (which is their strength), among which are working memory and verbalization, for example. There is no doubt that such an approach leads to differential psychology. This makes the object of study the existence of stable characteristics in individuals, which distinguish one individual from another.

Indeed, creating a clinical experiment not only makes it possible to reveal differences (even

concerning less apparent dimensions than hair color, for example), but also to measure this stability. The stability of these differences is the basis for the prognosis. Before any creation of a clinical experimentation tool, it is therefore essential to agree on the essence and the boundaries of what is being measured. If, in the context of intelligence for example, researchers define intelligence as the ability to solve a problem that has never before been encountered, then they could vary the clinical elements of the test according to degrees of difficulty or the g factor of the novelty of the problem to be solved being systematically clinically present.

Let's talk about the test itself. This technique is generally used to establish numerical differentiations, thanks to statistical tools such as dispersion, standard deviation, etc. However, it is also possible to use testing methods to place the subject in an atypical experimental situation, which is nevertheless effective. Take the case of the Thematic Aperception Test (T.A.T.) that Henry Murray created in 1935. In clinical psychology, the T.A.T., frequently used in parallel with Herman Rorschach's Test, consists of a series of storyboards showing scenes from everyday life. Unlike the Rorschach Test where the patient is questioned in order to know what s/he "sees" in the inkblots presented to him/her, in the T.A.T. the patient is asked to tell a story about what s/he sees. S/he therefore uses a multitude of complex and subconscious elements that guide him/her in his daily life, but which, in this case, are mobilized artificially through a clinical suggestion. It then becomes very interesting to use variants of this type of test in order to know the personal and private feelings of a subject. So, for example, rather than ask him/her if s/he thinks that a Mercedes is a younger or more mature brand, or if milk makes people fat, presenting the image of two different silhouettes and asking which drinks milk and which drives a Mercedes is enough. Subconscious associations will do the rest.

Finally, from the results of one test, factor analysis allows researchers to isolate factors that are common to several tests, which in turn allows the interpretation of numerical or qualitative findings. In this case, we are really at the heart of comparative clinical analysis with the control group in opposition to the experimental groups. Factor analysis makes it possible to clearly identify the variants and their impact on the subjects belonging to the control group. For example, if one wishes to set up a protocol that makes it possible to artificially and subconsciously influence subjects to provide certain answers to a questionnaire, it is essential to start by submitting the questionnaire to a control group to be able to then indicate the variations between the groups.

2. EXPERIMENTATION IN ECONOMICS AND BUSINESS ADMINISTRATION

In 1993, Davis and Holt wrote: "A small but increasing number of economists have begun to make use of experiments under economic conditions." The situation has dramatically changed since then, and behavioral economics has become a vast and widely recognized field of research. Experiments are used to test economic theories, such as financial decision-making, public goods provision, market equilibrium, public policy and other variations in the environmental economics (Jacquemot, l'Haridon & Morin, 2013; Lunn & Choisdealbha, 2018; Ferraro and Price, 2013; Binet, Denant-Boemont & Hammiche, 2019). In economics, experimentation has attained such success because lab and field experiments allow researchers to observe and analyze the specific determinants of economic behavior (Jacquemot, l'Haridon & Morin, 2013). Harrison and List (2014) further specify that a lab

experiment is an efficient method in order to create counterfactual scenarios. Lunn and Choisdealbha (2018) go so far to state that a well-designed lab study can be the best method for answering policymakers' research questions.

In management research, the use of experimental methods has also been reinforced (see Igalens & Roussillon Soyer, 2019), but later, and to a lesser extent. Specifically, marketing and behavioral finance are the fields of research where they are the most frequently mobilized. In both, research questions may reflect consumer or purchasing choices (for example, stocks).

Research designs that only provide measurement at a time t, and therefore no measurement before or after the event, are qualified as pre-experiments. A method of collecting empirical data that is very common in management science is the questionnaire. In these "ex-post-facto-designs," the researcher can certainly study correlations. However, s/he cannot manipulate the independent variable, which does not allow the elimination of a certain number of alternative explanations as in true experimentation.

Contrarily, clinical experimentation consists of reproducing a real situation in a setting or artificial environment while maintaining the maximum fidelity for the reality from which it comes in order to test a variable or to measure a variation (*delta*). For this, the researcher will either start from real statistical data collected from longitudinal observations, or s/he will observe the behavior of a group called the "control" in order to have a comparative basis before observing so-called "experimental" groups, thus with modifications of the measured variable. For example, based on consumption statistics, the researcher can recreate an environment close to that of a supermarket and provide the subjects with a shopping list before observing the way they walk around, their hesitations, etc. Another example, if you want to measure the coherence of a website's architecture, the researcher could simply place topics on a screen and ask the subjects to find some certain information. The researcher can then observe the path the subjects take, this time virtually.

Clinical experimentation can only help study the usual laws of behavior by systematically varying the factors that influence it. However, some of these factors can only be changed by using the differences between individuals and by comparing the answers given by a certain group of subjects to the answers given by another correctly chosen group. The latter must be of equivalent quantitative and qualitative composition. It is therefore essential to identify the important factors that are at the origin of the studied behavior in order to either modify them, withdraw one or more of them, or even to add one or more of them. This is why the use of control and experimental groups, which must be similar in their representations and content, is so important. Without systemic, stable and recurrent reference, it is impossible to observe a salience, a digression or an anomaly.

The following example is taken from the doctoral research of one of the authors who specialized in neuromarketing (Mouillot, 1999). The aim was to measure the degree of suggestibility of two experimental groups that received a subliminal suggestion, in other words a subconscious influence. Why two groups? Because it was necessary to check the results of the first one with a second experiment. Indeed, if the control group serves as a reference before the clinical variable is mobilized, the results from an experimental group must be confirmed by a second experiment organized under the same circumstances but with different subjects in order to highlight any result variation. If both experimental groups produce data,

whose proximity is considered acceptable - this interval must be defined upstream of the experiment - then the influence of the clinical variable becomes comparable to the control group. On the other hand, if the results provided by the two experimental groups are frankly divergent or convergent but judged to be insignificant, then the use of a third experimental group is necessary to identify the element(s) that caused this divergence.

In the present example, the research hypothesis was that natural behavior could be subconsciously influenced, thus modified, for at least 30% of the subjects participating in the experimentation, with more or less 5 points of variation. For this purpose, we used a control group consisting of 159 subjects and two experimental groups consisting respectively of 95 and 58 subjects. Methodologically, the number of control subjects was sufficiently close to the number of experimental subjects (159 vs. 153) to allow comparison. A questionnaire was submitted to the control group (159 subjects) in order to know the degree of market penetration of non-alcoholic drink brands without subconscious suggestion and independently of any sensory suggestion. The subjects completed the questionnaire in a room where there was no reference to any brand. Unsurprisingly, for this study, which took place in France and among French consumers, Coca-Cola, Evian, Orangina, Perrier and Pepsi-Cola collected the highest number of answers.

Then, the first experimental group was mobilized (95 subjects). This group was presented three video clips aimed at subliminally suggesting Teisseire, a French flavored syrup brand. The brand image was inserted and projected at a speed of $1/25^{th}$ of a second. The group was invited to watch these clips in order to choose which one would be most relevant to accompany an advertising campaign aimed at fighting AIDS. While both the first experimental group and the control group named most of the same brands, Teisseire reached 4^{th} position and Pepsi-Cola was taken out of the rankings. With 33.68% of the answering behaviors having been modified, our hypothesis was validated. It only remained to verify with a second experimental group that these results were methodologically acceptable. Seven months later, a second experimental group was formed (58 subjects) and submitted to the same protocol. After viewing the clips, Teisseire reached 5^{th} position in the questionnaire answers, this time replacing Perrier. In this second case, 27.33% of the responses were modified. We hypothesized an artificial behavioral change within a range of 25 to 35%, and our two results were within this range. Since Teisseire was never mentioned by the control group subjects, and a Chi-square test had validated our statistics, this result encouraged us to think that artificial subconscious suggestion had a real impact on the randomly-chosen 18 to 25-year-old subjects, whatever their gender.

3. EXPERIMENTATION IN INTERNATIONAL MANAGEMENT

Certain fields such as economics, psychology and marketing increasingly use experimental research designs, to the point where they have become almost mandatory for high-level publications. By contrast, in international management, Zellmer-Bruhn, Caligiuri, and Thomas (2016) found that less than 1% of more than 900 empirical studies published in the *Journal of International Business Studies* (JIBS) use experimental methods. This is mainly due, among other things, to both the difficulty of designing appropriate research protocols that can be randomly assigned to participants and to the challenges related to sampling.

By 2016, experimentation was reportedly only present in a limited number of business and management subjects, including marketing/ advertising consumer behavior, sales communication, venture capitalists' (VCs') decision-making, cultural differences in decision-making, and empowerment and job satisfaction (Zellmer-Bruhn, Caligiuri & Thomas, 2016). However, it must be noted that few of these experiments were true experiments due to their non-randomized sampling approaches.

In the international business (IB) context, particular challenges to conducting experiments lie in the difficulties related to randomization and the nature of the samples. In other words, it is impossible for researchers to completely randomize their samples. As Zellmer-Bruhn, Caligiuri and Thomas (2016) convey, researchers cannot simply randomly assign countries to political economies, companies to globalization strategies, or country of origin to individuals, for example. Furthermore, student samples have been long viewed as inferior in the IB community, whereas they are often used in other disciplines such as psychology. While student samples may be appropriate for certain studies in IB, undergraduates cannot demonstrate proficient background and experience to respond sensibly to all experiments (Van Witteloostuijn, 2015). Furthermore, on the subject of samples, IB requires specified and varying cultural and institutional backgrounds, often with people from different geographic locations (Zellmer-Bruhn, Caligiuri & Thomas, 2016). This can represent a challenge simply in recruiting the appropriate participants.

In IB research, the most common experiments relate to individual or team-level outcomes. In international marketing, where experiments have long taken place, the study by Pornpitakpan (1999) provides an interesting example. Pornpitakpan (1999) concludes that Americans who adapt to two cultures and languages, in this case Thai and Japanese, have more positive sales outcomes even within their national contexts. In international economics, Roth, Prasnikar, Okuno-Fujiwara & Zamir (1991) employ an experimental design to compare bargaining behaviors among different cultures. They demonstrate that what an "acceptable" offer is depends on cross-country differences. In management, Caligiuri and Phillips (2003) conducted an experiment, randomly assigning expatriate participants to use or not a self-assessment decision-making tool. They showed that participants who received the tool, along with a realistic job preview (RJP), reported increased confidence in their decision to accept an international assignment and had greater success during said assignment.

In *inter*cultural management (that is if one wants to go beyond the comparison of different cultures), clinical experimentation is traditionally interactional. Indeed, since interculturality necessarily implies the existence of a relationship between at least two people with different personal, social and societal repositories, this approach aims to demonstrate the existence of

interactions between two subjects. Any response is then considered to be a reaction to one or more stimuli, knowing that the behavior of a person cannot be changed without concomitant changes in his entourage.

However, in terms of clinical experimentation, it is all about interpretation, especially when one is working within the broad field of intercultural analysis. Without interpretation, any measure is meaningless, since it is neither put into perspective nor placed in a context, which in this case is experimental. Thus, when one speaks of the test method, without analysis of the standard deviation, an average could appear indicative of a tendency whereas it is not so. It is the same with dispersion. From a single cloud of points on which a large concentration would appear separated from a few isolated points, a study of consumption behaviors would be tempted to focus on the cloud of points whereas a psychological or intercultural study would be inspired to observe isolated points, which could potentially be respectively representative of psychological deviations or cultural specificities. In research, the very nature of the information that needs to be gathered to achieve the objective therefore commands the means employed to do so. Indeed, since 1993 when her book *Methods of Social Sciences* was published, Madeleine Grawitz has invited us to think about the fact that if it is possible to consider catching fish with a butterfly net, we certainly do not catch butterflies with hooks. Moreover, Friedrich Nietzsche also always maintained that the most precious riches were methods, this being a viewpoint that Rene Descartes could not deny either, for it is not the object that makes science but the method.

Despite the small number of studies mobilizing experimental methodologies in international management, multiple authors highlight its bright future. Zellmer-Bruhn, Caligiuri and Thomas (2016) conclude that experimentas are underrepresented in JIBS, but that they provide a clear opportunity to develop the evidence for causal relationships in international business research. Van Witteloostuijn (2015) encourages the development of an “experimental IB tradition” through the creation of web-based tools to further research, but also to increase the effectiveness and fun of IB teaching. In this way, we will next discuss such a tool created by the InterCCom project, that studies as well as teaches about intercultural teamwork.

4. THE INTERCCOM PROJECT: A PLATFORM FOR INTERNATIONAL MANAGEMENT EXPERIMENTATION

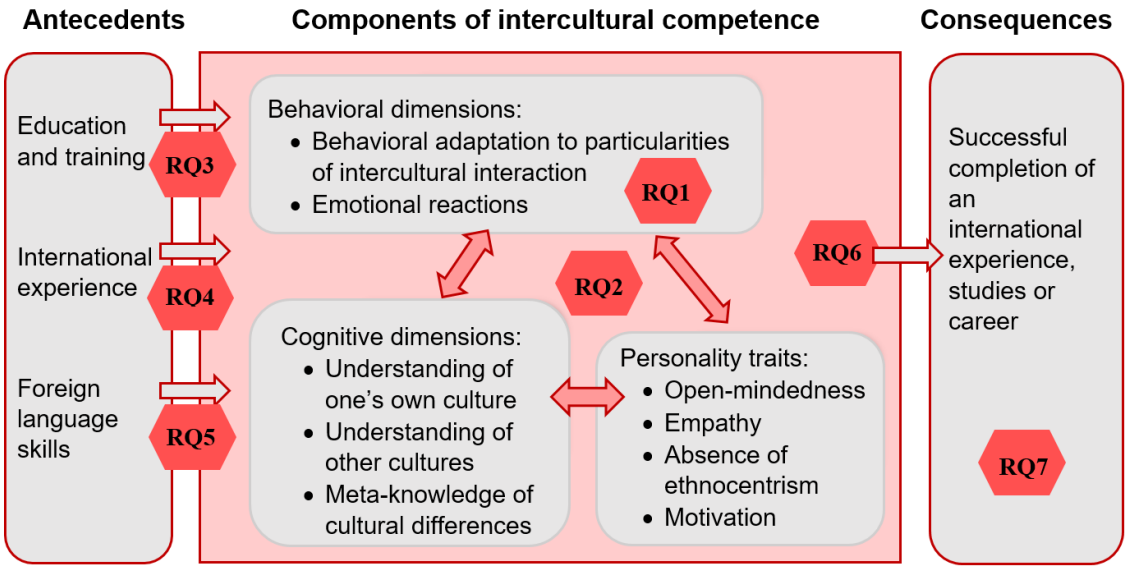
In this last part of the chapter, we present a project that has been conducted since January 2019 at Sciences Po Grenoble and University Grenoble Alps, which aims at building a serious inter-cultural management game called LINK the serious game®. The game be used as a quasi-experimentation tool in research because the choices made by the players are recorded. The game is implemented by a digital platform also developed within the project, the GenaGame® platform. On this platform, content (documents, images and videos) can be uploaded, and an interactive journey of the player through a content, depending on a scenario and underlying theoretical dimensions, can be defined. Other games in international management are thus under development. All of these games are grouped under the i-Team Games® brand.

The relevance of the InterCCom project within the field of intercultural management is threefold. First, it will allow an objective measure of the behavioral dimension of intercultural competence, which has not yet been achieved. Second, it responds to the call for more experimental research in international management (Zellmer-Bruhn, Caligiuri, & Thomas, 2016). Finally, it helps develop a game-based approach to intercultural management, which has only rarely been done, whereas serious games are attracting ever-increasing attention as innovative and effective tools for learning and management knowledge (Michel, Kreziak & Heraud, 2009, Vallat et al., 2016).

The context imagined for LINK the serious game® is that of a virtual team working on an innovation project. Innovation projects are more and more often carried out within international teams that are scattered throughout the world. For collaboration within these teams to be successful, the Intercultural Competence (IC) of those team members is essential (Molinsky et al., 2012, Stahl et al., 2010). The concept of IC has given rise to a vast literature in communication sciences, languages and civilizations, and of course international management. A great diversity of approaches and vocabulary coexist, but a conceptualization that has gradually imposed itself addresses the IC through its components. These can be grouped into three categories: personality traits seen as IC-related (e.g. open-mindedness, empathy, lack of ethnocentrism, etc.), knowledge of cultural differences, and the behavioral adaptation to the latter. Training, but especially international experience, is generally considered to contribute to IC. Figure 1 summarizes the background, components and consequences of IC.

The most common existing measures in IC all have important biases. This casts doubt on the concept itself, which is insufficiently empirically based (Van de Vijver & Leung, 2009):

- Very often, intercultural competence is reduced to only one of its components: personality traits and attitudes related to intercultural competence, measured by methodological tools such as the MPQ (Multicultural Personality Scale, Van Oudenhoven & Van der Zee, 2002). However, the link between this component and the others is not obvious. In fact, in recent research, only 17% of intercultural knowledge is explained by these personality traits (Bartel-Radic & Giannelloni, 2017).
- Measures on the notion of cultural intelligence (Ang et al., 2007, Thomas et al., 2008), defined in the same manner as IC, are certainly multidimensional (and therefore also include cognitive and behavioral dimensions). However, the small number of questions per dimension is self-assessed and not located.
- In practice more than in research, intercultural competence is often seen as “automatically” resulting from international experience. However, our research results in a demonstrated link of up to about 5%, which is extremely weak (Bartel-Radic, 2014).



Research Questions:

RQ 1: How should **behavioral dimensions** of intercultural competence be measured **experimentally**?
RQ 2: To what extent do **personality traits** influence intercultural competence?
RQ 3: What types of **education and training** impact intercultural competence?
RQ 4: Under what conditions does **international experience** increase intercultural competence?
RQ 5: When and to what extent do **foreign language skills** increase intercultural competence?
RQ 6: Does intercultural competence **differ among different cultures**?
RQ 7: To what extent does intercultural competence lead to success in a person's **international or intercultural experience, studies or career?**

Figure 1: Antecedents, components and consequences of intercultural competence

At the research level, the objective of the InterCCom project is to answer the above research questions through the development of a serious computer game. This game avoids the methodological bias of existing methods and evaluates knowledge and intercultural behavior in various realistic situations. The results collected through this game and subsequently analyzed will answer the aforementioned research questions. Each of the research questions addresses current theoretical gaps in the field of intercultural management.

Specifically, in LINK the serious game®, the player assumes the role of the coordinator of both a fictional and virtual team of an innovation project, a team composed of people of various nationalities and scattered around the globe. The player must coordinate the work of the team, adapting their behavior to their respective national cultures. Depending on the communication and the management styles adopted, the virtual members of the team will more or less contribute to the team's work and its performance. The scenario will be built through a succession of scenes corresponding to the technique of "critical incidents" (Flanagan, 1954) or brief stories recounting intercultural situations. These scenes are likely to be differently interpreted according to different cultures because they are based on the cultural dimensions developed by Hofstede (1980, 2001), Hall and Hall (1990) and Hampden-Turner & Trompenaars (2003), which relate to the vision of the group, hierarchy, time, competition, rules, etc. The critical incident technique captures tacit knowledge related to IC (Johnson et al., 2006). The use of five critical incidents in the preliminary project research (Bartel-Radic, 2014, Bartel-Radic & Giannelloni, 2017) has been viewed as a promising prospect for the field. In

this project, the goal is to go beyond and use them as a basis for a quasi-experimental methodology.

The method used to develop the storyboards and the game scenario is based on three complementary processes:

- Interviews and focus groups with leaders and members of international innovation project teams were conducted and analyzed to identify “critical incidents” of interest to the game.
- During the collaborative workshops which brought together an interdisciplinary and very international project team, improvisation theater was used as a methodology to develop more scenes, and to discuss the reactions of the bearers of this or that culture to these situations.
- For the two phases of collaborative work on the game, that is to say also during the platform and game design, we mobilized design thinking methodologies (Dorst, 2011, Chanal & Merminod, 2018). The project team defined four personas of players and worked on the respective expectations and motivations before describing the game from the perspective of the users’ experiences.

The commonality between the different applications of the i-Team Games® collection is that the player is a coordinator and/or member of an international virtual team. A specificity of the project compared to others that exist in the field is such that the player plays his/her own role (in terms of age, sex, nationality, place of work, etc.), which greatly contributes to anchoring the realism of the game. The virtual exchanges with the other members of the team simulated in the game are done via digital communication tools, electronic or instant messaging, documents sharing and videoconference (simulated with the help of short videos).

If these experiments can capture the behavior of participants in virtual international teams vis-à-vis various research issues, we cannot qualify them as experiments in the strict sense of the term. Let's go back to the principles of experimentation mentioned in this chapter. It is clear that this is the case of a clinical experiment, where the environment of the participant is artificially reconstituted, in order to be able to control the various variables as much as possible and to compare a large number of individuals. Two other key principles are “the administration of a treatment or experiment,” a variable that is varied during the experiment, as well as the existence of a control group for which this variable does not change. The game LINK the serious game® includes an assessment of the behavioral dimension of intercultural competence at the beginning and the end of the game, as well as game sequences followed by explanations in the form of included lessons. A possibility of use is therefore to separate the sample in two, with one part going through the whole game, and the other one which only carries out the evaluations in both the introduction and the conclusion. This protocol would allow testing the hypothesis that training in intercultural collaboration in the form of serious play increases the participants’ intercultural competence. But the most interesting issues (see Figure 1) are much more difficult to test in an experimental form. Research questions 1 and 2 focus on the behavioral nature of intercultural competence, but do not imply “administered processing” or even a systematic hypothesis test. For research questions 3 and 4 on factors increasing intercultural competence, there are two possibilities: 1) to use the measure of intercultural competence through one-time experimentation, and to compare subgroups according to their prior international experience or their training. We are then in the case of

pre-experimentation as well as a quasi-experiment. Or, 2) raise intercultural competence through experimentation twice, before and after international experience or training. This research design is then an (almost) real experiment, with a control group (but not assigned in a randomized manner) and "processing" that is administered between the two measures.

CONCLUSION

Whatever the variable that the researcher wishes to test, the latter will always be stuck between the devil and the deep blue sea, either as a prisoner of his will to reproduce a reality in a clinical environment, or as a slave to his predisposition to determine what real environment would be the most conducive to welcome a given experiment. There is no "best solution" in experimental research. The key is to maintain a strong intellectual integrity and to accept the need to find strength in weakness, in other words to see a source of future exploration within the limits of a research study. A clinical environment is remarkable because it allows the researcher to be in control of all endogenous and exogenous variables that s/he wants to mobilize to measure what s/he ultimately aims to measure or understand. And the field is outstanding because it allows the researcher to test what s/he wants to test with the confidence that all effects will be represented. As early as the beginning of the 18th century, Gottfried Wilhelm Leibniz was referring to "small perceptions," or all of those influences, suggestions, interferences, and other stimuli that govern our perceptions and choices without us being able to detect them. With the progress of science, we now know that a human being is able to capture 40 stimuli at a time from his/her environment among the thousands that yet reach him/her. S/he will never see all that is exposed to his/her eyes, s/he will never feel all the perfumes and odors that reach him/her, s/he will only taste what is sweet, salty, bitter or sour, s/he will only hear certain sounds, and s/he will only touch what will be within his reach. The essential will be out of reach, as has always been the case. This is the reason why experimental research is so important. The researcher must therefore go through a phase of frustration in which s/he must first of all determine what s/he *cannot* measure. The number of variables will be reduced but the researcher will reach results with a greater amount of precision. In reality, s/he will increase the number of variables but necessarily reach biased results. In international management, the dilemma is perhaps even more tangible because to the complexity of the human is added to that of culture. Testing cultural variables through a clinical approach can help identify stereotyped behaviors. However, those behaviors can be fake as they may be overplayed. Testing cultural variables in a real environment creates a space in which natural reactions can have free rein. On the other hand, individual personalities often take precedence over culture, which makes it difficult to clearly determine what comes from what. The ultimate solution probably lies in a hybrid approach, a mixed methodology combining clinical and field approaches. But above all, only time and the tenacity of researchers will make it possible to better understand Humanity, in its diversification but likewise in its similarities. To believe in this future, it is perhaps already right to rule in favor of Arthur Schopenhauer who, in 1819, wrote, *"thus the task is not to contemplate what no one has contemplated but to meditate like no one has ever meditated on what everyone has in front of them."*

REFERENCES

Ang, S., Van Dyne, L., Koh, C., Ng, K. Y., Templer, K. J., Tay, C., & Chandrasekar, N. A. (2007). Cultural Intelligence: Its Measurements and Effects on Cultural Judgment and Decision Making, Cultural Adaptation and Task Performance. *Management and Organization Review*, 3(3), 335-371.

Bartel-Radic, A. (2014). La compétence interculturelle est-elle acquise grâce à l'expérience internationale ?" *Management International*, 18 (Special Issue), 194-211.

Bartel-Radic, A., & Giannelloni, J.-L. (2017). A renewed perspective on the measurement of cross- cultural competence: An approach through personality traits and cross-cultural knowledge. *European Management Journal*, 35(5), 632-644.

Binet, Denant-Boemont & Hammiche (2019), The Incidence of Price Information on Water Consumption Under Complex Tariff Scheme: a Laboratory Experiment, 6th French Association of Environmental and Resource Economists, University of Rennes 1.

Caligiuri, P. M., & Phillips, J. (2003). An application of self-assessment realistic job previews to expatriate assignments. International Journal of Human Resource Management, 14(7), 1102-1116.

Chanal, V., & Merminod, V. (2018). Comment adresser les problèmes pernicieux de manière créative avec le design thinking? *Management international* (in press/online first).

Cortes-Cornax, M., Front, A., Rieu, D., Verdier, C., & Forest, F. (2016). ADInnov: An Intentional Method to Instil Innovation in Socio-Technical Ecosystems. In International Conference on Advanced Information Systems Engineering CAiSE 2016: *Advanced Information Systems Engineering* (pp. 133-148). Springer, Cham.

Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton university press.

Dorst, K. (2011). The core of 'design thinking' and its application. *Design studies*, *32*(6), 521-532.

Ferraro, P. J., & Price, M. K. (2013). Using nonpecuniary strategies to influence behavior: evidence from a large-scale field experiment. *Review of Economics and Statistics*, *95*(1), 64-73.

Flanagan, J. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327-358.

Forest, F., Lavoisy, O., & Chanal, V. (2009). Integrated Scenario-based Design Methodology for Collaborative Technology Innovation, *The Future of Innovation* (*Int. Journal of Innovation Management*), June, Vienna, 99.

Grawitz, M. (1993). *Méthodes des sciences sociales.* Précis Dalloz, 9^{e} édition.

Greenwood, E. (1972). Das Experiment in der Soziologie. In: König, R. (Ed.) *Beobachtung und Experiment in der Sozialforschung*. 8^{e} ed. Köln: Kiepenheuer & Witsch, 171-220.

Hall, E. T., & Hall, M. R. (1990). *Guide du comportement dans les affaires internationales : Allemagne, Etats-Unis, France*. Paris, Seuil.

Hampden-Turner, C., & Trompenaars, F. (2003). *Riding the waves of culture*. London, Nicholas Brealey.

Harrison, G. W., List, J. A. (2004). Field experiments. *Journal of Economic literature*, *42*(4), 1009-1055.

Hill, A. B. (1965). The Environment and Disease: Association or Causation? *Proceedings of the Royal Society of Medicine,* 58(5), 295-300.

Hofstede, G. (1980). *Culture's consequences: International differences in work related values*. London, Sage.

Hofstede, G. (2001). *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations across Nations*. Thousand Oaks, CA, SAGE.
Igalens, J., & Roussillon Soyer, C. (2019) Quasi experimental design. In : Chevalier, F., Cloutier, L.M., & Mitev, N., *Research methods for the DBA*, Paris : EMS, 245-260.
Jacquemet N., L'Haridon, O., & Morin, P. (2013) Economie expérimentale et comportements. Introduction, *Revue française d'économie*, 28(2) : 3-8. DOI : 10.3917/rfe.132.0003
Johnson, J. P., Lenartowicz, T., & Apud, S. (2006). Cross-cultural competence in international business: toward a definition and a model. *Journal of International Business Studies,* 37(4): 525-543.
Leibniz, G. W. (1703). Les petites perceptions. In *Nouveaux essais sur l'entendement humain*, II, chapitre 1. Paris: Garnier, 96-97.
Levitt, S. D., & List, J. A. (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic perspectives*, *21*(2), 153-174.
Meyer, M. (2008). La méthode des scénarios : un outil d'analyse et d'expertise des formes de communication dans les organisations, *Études de communication*, 31, 2008, 133-156.
Michel, H., Kreziak, D., & Heraud, J.-M. (2009). Evaluation de la performance des Serious Games pour l'apprentissage : Analyse du transfert de comportement des éleveurs virtuels de Vacheland. *Systèmes d'information et management*, 14(4), 71-86.
Molinsky, A., Davenport, T. H., Iyer, B., & Davidson, C. N. (2012) Three Skills Every 21st Century Manager Needs. *Harvard Business Review*, 90(1-2), 139-143.
Mouillot, Ph. (1999). *Fondements épistémologiques, théoriques et expérimentaux de la suggestion subliminale par le biais des images subliminales : Une approche de communication commerciale.* Thèse de doctorat, Université de Nice - Sophia-Antipolis.
Mouillot, Ph., Drillon, D., & Montargot, N. (2018). Mosaïque des intelligences dans l'organisation : Éléments de distinction entre hauts potentiels, talents et HQI. *Revue Internationale de Psychosociologie et de Gestion des Comportements Organisationnels*, XXIV(58), 171-196.
Pornpitakpan, C. (1999). The effects of cultural adaptation on business relationships: Americans selling to Japanese and Thais. *Journal of International Business Studies*, 30(2), 317-337.
Roth, A., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and Marketing Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study. *American Economic Review,* 81(5), 1068-1095.
Schnell, R., Hill, P. B., & Esser, E. (2018). *Methoden der empirischen Sozialforschung*. Oldenbourg: De Gruyter.
Schopenhauer, A. (1819). *Le monde comme volonté et comme représentation*. Traduction A. Burdeau, PUF, 1992, p. 58.
Spearman, C. (1904). "General Intelligence", Objectively Determined and Measured, *American Journal of Psychology*, 15(1), 72-101.
Stahl, G. K., Maznevski, M. L., Voigt, A., & Jonsen, K. (2010). Unraveling the effects of cultural diversity in teams: A meta-analysis of research on multicultural work groups. *Journal of International Business Studies*, 41(4), 690-709.
Thomas, D. C., Elron, E., Stahl, G., Ekelund, B., Ravlin, E., Cerdin, J.-L., Poelmans, S., Brislin, R., Pekerti, A., Aycan, Z., Maznevski, M., Au, K., & Lazarova, M. (2008). Cultural Intelligence: Domain and Assessment. *International Journal of Cross-Cultural Management* 8(2), 123-143.
Vallat, D., Bayart, C., & Bertezene, S. (2016). Serious games in favour of knowledge management and double-loop learning? *Knowledge Management Research & Practice*, 14(4), 470-477.

Vijver, F. J. R., & Leung, K. (2009). Methodological Issues in Researching Intercultural Competence. In Deardorff, D. K. (Ed.), *The SAGE Handbook of intercultural competence*. (404- 418), Thousand Oaks, SAGE.
Van Oudenhoven, J. P., & Van der Zee, K. I. (2002). Predicting multicultural effectiveness of international students: The Multicultural Personality Questionnaire. *International Journal of Intercultural Relations* 26(6), 679-694.
Van Witteloostuijn, A. (2015). Toward experimental international business: Unraveling fundamental causal linkages. *Cross Cultural Management*, *22*(4), 530-544.
Zellmer-Bruhn, M., Caligiuri, P., & Thomas, D. (2016). From the Editors: Experimental designs in international business research. *Journal of International Business Studies* 47(4),399-407.

AUTHORS

Anne Bartel-Radic
Full Professor in International Management
Univ. Grenoble Alpes, Grenoble INP, CERAG, F-38000 Grenoble, France

Philippe Mouillot
Associate Professor in Marketing & Strategy - ACR
IAE de Poitiers, France

Danielle Taylor
PhD Candidate in International Management
Univ. Grenoble Alpes, CERAG, F-38000 Grenoble, France