ISSN 1808-057X
DOI: 10.1590/1808-057x20231594.en

ORIGINAL ARTICLE

Experimental scenarios in accounting research: Highlights in methodological development

Myrna Modolon Lima[1]
https://orcid.org/0000-0003-2084-481X
Email: myrnalima@alumni.usp.br

[1] Universidade de São Paulo, Faculdade de Economia, Administração, Contabilidade e Atuária, Departamento de Contabilidade e Atuária, São Paulo, SP, Brazil

Received on 12.08.2021 – Desk acceptance on 01.06.2022 – 5th version approved on 03.12.2023
Editor-in-Chief: Fábio Frezatti
Associate Editor: Jacqueline Veneroso Alves da Cunha

ABSTRACT

The main objective of this article is to guide the researcher, through the exposition of highlights and examples from previous literature, in the building of an experimental scenario for the investigation of his research question. The study focuses on the use of experimental scenarios in accounting research, bringing specific highlights that must be observed regarding the theoretical and methodological development of the instrument. Highlights of a methodological nature, relevant to the use of experimental scenarios aimed at investigating accounting research questions focusing on individuals' judgment and decision-making, are systematized. The impact of this research consists of systematizing aspects relevant to the building of experimental scenarios that can affect methodological rigor and internal, construct, and statistical conclusion validity. It consists of a method proposal approach, with examples brought through a narrative review. The study systematizes relevant aspects to be considered in the researchers' choices regarding the building of experimental scenarios, such as the role of randomization, the choice of participants, the minimization of social desirability bias, specific ways of collecting the dependent variable, and possibilities of collecting the participant pool. Additionally, the role of using experimental scenarios is discussed in view of the theoretical nature of the research question, mainly regarding the assessment of judgment and decision making.

Keywords: experiment, experimental scenarios, accounting research, causality, experimental games.

Correspondence address

Myrna Modolon Lima
Universidade de São Paulo, Faculdade de Economia, Administração, Contabilidade e Atuária, Departamento de Contabilidade e Atuária
Avenida Professor Luciano Gualberto, 908 – CEP 05508-010
Butantã – São Paulo – SP – Brazil

This is a bilingual text. This article was originally written in Portuguese, published under the DOI https://doi.org/10.1590/1808-057x20231594.pt
Paper presented at the XV Congresso ANPCONT, December 2021.

1. INTRODUCTION

This article aims to systematize highlights to build experimental scenarios in accounting research in the Brazilian context. Experimental scenarios, or vignettes, are commonly used in international accounting studies, just as in Hartmann and Maas (2010) and Kadous and Zhou (2019), however, they are still poorly explored in the national literature (Aguiar, 2017; Frezatti et al., 2015). Scenarios allow researchers to benefit from internal validity, consistent with the use of an experimental approach, and the testing of theories through situations closer to reality (Eckerd et al., 2021). Also, experimental scenarios are an efficient way to reconcile the experiment's internal validity with the external validity of situations explored in the scenarios (Aguinis & Bradley, 2014). In this sense, it is assumed that the use of scenarios can bring greater contributions to areas with an applied nature, such as accounting research.

Studies like Aguiar (2017) make substantial contributions by discussing central aspects of using an experimental approach in accounting research. The author discusses aspects relevant to the experimental method, exemplifying possible research designs and its execution. Additionally, Suave et al. (2021) explore, specifically for research in accounting education, the use of experimental methodologies as promoters of improvements in teaching practices. Despite the contributions to Brazilian accounting research in terms of exposing the method and notes on future research perspectives, both studies address the use of experiments as a whole, without delving into research building and choices in the experimental scenario.

In the international context, studies like Aguinis and Bradley (2014), Collett and Childs (2011), Eckerd et al. (2021), and Lonati et al. (2018) addressed the use of the experimental scenario or vignette in research. The authors sought to point out concepts, build best practices, and address possible problems in using experimental scenarios. However, these studies have a broader focus, addressing the use of scenarios as a method for social sciences or organizational research.

Thus, this article seeks to complement the research gap introduced by these two Brazilian national studies and by international research, in order to deepen the knowledge about experimental scenarios. That said, the main objective of this text is to systematize points in the building of experimental scenarios in Brazilian accounting research. In order to achieve its main objective, it relies on the following specific objectives: define experimental scenarios; bring the main differences regarding the choice of the experimental scenario in view of the experimental game; and discuss pertinent highlights in the building of experimental scenarios.

In short, we seek to contribute to Brazilian researchers in the building of experimental scenarios as a research method, providing examples of the use of the method in previous research in the international and national literature. Also, since non-compliance with highlights brought in the article can harm the internal, construct, and external validity of the study, we seek to bring to light specificities and examples that can help researchers to minimize validity issues in their studies.

2. EXPERIMENTAL SCENARIOS OR VIGNETTES

Experimental scenarios, also known as vignettes, are descriptions of realistic situations, often hypothetical, introduced to respondents so that they can infer about their behavior, intents, and attitudes in the face of exposed information (Aguinis & Bradley, 2014; Trevino, 1992; Wason et al., 2002). Given their experimental nature, the scenarios introduce manipulated and randomized information, while keeping all the other scenario items constant, which allows the verification of causality between the research variables (Shadish et al., 2002). Nevertheless, their rather realistic characteristics bring greater generalization of results, compared to other types of experimental methods such as the experimental game (Aguinis & Bradley, 2014). In short, the use of experimental scenarios allows researchers to create or reproduce actual scenarios in hypothetical situations, combining the realism needed to ensure the study's credibility and the respondent's identification with the research subject.

Experimental scenarios can also be thought of as a way of simulation, in which conditions are created "within an experiment that simulate or closely duplicate the natural environment in which the behaviors being examined would normally occur" (Gravetter & Forzano, 2012, p. 219). For instance, in the case of audit research, one can check the auditor's judgment in a fair value assessment of

the impairment test of a client's goodwill, in which his/her motivation (intrinsic versus extrinsic versus control) (Kadous & Zhou, 2019).

Finally, it is understood that the objective of experimental scenarios is for the respondent to feel, psychologically, within the exposed situation, in order to emulate the type of behavior or intended behavior that he/she might experience in that situation or in similar situations (Gravetter & Forzano, 2012). The scenarios created do not serve to assess the respondent's behavior for that single scenario, but so that, in the face of the simulated environment, it is possible to verify the respondent's behavior in a similar scenario (Gravetter & Forzano, 2012). Going back to the examples mentioned, an audit firm needs to make decisions about complex estimates and the impairment test is a faithful representation of one of these decisions. That is, the scenario has specificities, but it emulates a context that is beyond the situation described.

2.1 The Main Differences between the Use of Experimental Scenarios *versus* Experimental Games and Their Limitations

One of the researchers' main decisions when building their experiment is whether to use experimental scenarios and/or experimental games. The choice should be driven by the research question (Kinney, 2019), thus, it is worth knowing the differences and constraints of using experimental scenarios when compared to experimental games, since this will make researchers able to build one and/or the other in a more informed way.

The first big difference between experimental scenarios and experimental games concerns the context, more precisely, how real the context is. A point of convergence between the two types of experiment is about experimental realism. In experimental realism, elements are introduced that manage to capture the individual's attention to a given topic and that cause impact, but that do not provide details emulating reality (Liyanarachchi, 2007). However, additionally, experimental scenarios also use greater details, which emulate a real context, called *mundane realism*. Mundane realism is defined as the possibility that events occurring in the experiment are likely to occur in practice ('real world') (Swieringa & Weick, 1982).

As for the building of the experiment, mundane realism is responsible for the level of detail in the context of the experimental scenario, which should be at researcher's discretion. This level of detail is commonly found in participant characterization, e.g. company manager or employee or in the details provided about the company that help the participant to put himself in that position and grasp the context more easily, as seen in Aguiar (2021), who asks participants to put themselves in the position of managers in a mid-sized industry responsible for the emission of toxic gases. On the other hand, experimental games tend to be rather artificial (Arnold & Triki, 2018), including the names of participants, who are often referred to as *players*, as seen in Santos et al. (2021).

There are discussions in the literature as to the extent of application of mundane realism. Certain researchers claim that one should try to mitigate the possibilities of mundane realism, focusing primarily on experimental realism, since the former can interfere with the quality of inference by adding distractions and contexts that would be ignored in real-life situations (Gravetter & Forzano, 2012; Kadous & Zhou, 2018). For instance, the addition of extremely detailed contexts, which may confuse the respondent and saturate them with the amount of information, are not recommended as they only create an unreal scenario, focusing on aspects that are relevant to grasping of the proposed situation (Kadous & Zhou, 2018).

On the other hand, other researchers address a need for a balance between mundane realism and experimental realism, or that researchers should assess the need for adding specific details for the building of their theory (Haynes & Kachelmeier, 1998), considering the need for the respondent to feel, psychologically, within the exposed situation in order to emulate the type of behavior that they would experience in that situation or in similar situations (Gravetter & Forzano, 2012). Thus, it is understood that the researcher should make an assessment of the benefits and harms of including contextual elements in the scenario, focusing mainly on the type of research carried out (Haynes & Kachelmeier, 1998; Kadous & Zhou, 2018).

There are positive points in using more abstract experiments, more common to experimental games, mainly for testing theories (Lucas, 2003), since a theory "cannot possibly incorporate every single factor that is present in an individual's environment, nor can it predict how all of these factors can simultaneously impact individuals' behaviour" (Arnold & Triki, 2018, p. 290). However, there is no unanimity in the literature or in the academic community about the superiority of one type of experiment over another.

The discussion between mundane realism and experimental realism intertwines with discussions between researchers in the experimental economics community and the psychology community. In the case of accounting research, we see the same discussion between

these different communities regarding the building of the experiment. This rationale is corroborated by Haynes and Kachelmeier (1998), who state that experimental economics researchers tend to classify experimental scenarios as very confusing, while psychology researchers classify experimental games as too abstract. In short, it is understood that accounting researchers who choose to use experimental scenarios must create a context that is adequate to test their theory, making use of both experimental realism, needed for any type of experiment, and mundane realism (with parsimony).

The second major difference refers to the research focus, i.e. whether the focus is economic and/or behavioral/cognitive. Like the discussion between mundane realism and experimental realism, it is also usual, in accounting research, to see similar discussions between researchers who focus on economic theories or psychological theories, who have various stimuli to be investigated (Libby et al., 2002). Experimental scenarios are commonly used in research questions focusing on judgment and decision-making processes, given that the scenarios are hypothetical and can be designed to capture these issues (Trevino, 1992).

On the other hand, experimental games stem, for the most part, from Classical Economic Theory, mainly from Game Theory, which was headed by Von Neumann and Morgenstern (2004) and Nash (1950,1951) – responsible for the concept called Nash Equilibrium (Onderstal, 2014, p. 60). The basic premise of the various models of Game Theory is that of Rational Choice Theory, which predicts that "[...] a decision-maker chooses the best action according to her preferences, among all the actions available to her" (Osborne, 2000, p. 4). Game Theory seeks to solve economic problems that go beyond economic theory, more specifically, it focuses on grasping economic behavior and "[...] in the attempts to find an exact description of the endeavor of the individual to obtain a maximum of utility, or [...] a maximum of profit." (Von Neumann & Morgenstern, 2004, p. 1).

For instance, the Agency Theory has premises of the Classical Economic Theory, which addresses the man/woman as a rational individual ('*homo economicus*'), capable of seeking to maximize their gains in each choice (Jensen & Meckling, 1976), which facilitates its adaptation to various experimental designs. In practical terms, it is usual for accounting research that addresses economic theories to resort to games adapted as experimental tasks, such as the dictator game (Ham et al., 2017), the public goods game (Tayler & Bloomfield, 2011) and the prisoner's dilemma (Grasser et al., 2021), rather than using experimental scenarios.

On the other hand, the experimental scenarios focus on bringing contextual aspects that can emulate behavioral/cognitive stimuli that are needed for various types of accounting research, since, probably, the respondent would not be able to be stimulated in the same way in a game (Libby et al., 2002). Thus, studies that require specific knowledge or that focus on the respondent's various motivations, such as auditors or investors, may be more interesting when using experimental scenarios, given their rather realistic nature.

It is worth noticing that, although accounting studies mostly focus on economic theories (Hesford et al., 2006), the explanations for the phenomena rely on behavioral/cognitive perspectives. In this case, it can be seen that the entire literature on budgetary slack, for instance, involves the use of two or more theories, one being economic and the other(s) behavioral. As seen in Hartmann and Maas (2010), which unites the perspective of agency conflict of the business unit's controller when preparing the budget proposal relying on behavioral perspectives of social pressure and the controller's Machiavellian level.

The third major difference concerns how participants are characterized in both experiment modalities and the consequences of this characterization. As previously mentioned, participants in experimental scenarios and experimental games are classified, respectively, in a more and less contextual way. Given this characterization, the researcher must take care to emulate, in some way, the experimental realism needed for participants to understand. For instance, situations where there are differences in power as seen between managers and employees, junior and senior auditors, and between auditors and managers are easily captured in an experimental scenario, as participants receive instructions about which role they should take.

However, in the experimental game, this relationship needs to be emulated so that players see that there is a hierarchy between them, as seen in Schuhmacher et al. (2021). When talking about the effects of the leader's example in organizations that have social objectives on the followers' contributions regarding the achievement of a common goal, the authors randomly select the participants who take the role of leader or follower, but give freedom to leaders so that they can choose their level of contribution towards this achievement, while randomizing what type of leader contribution (low vs. high) followers will be exposed to. This experimental design choice helps to emulate the hierarchical differences of each position.

3. EXPERIMENTAL SCENARIOS OR VIGNETTES IN PRACTICE: HIGHLIGHTS REGARDING DEVELOPMENT

As in any experimental design, the main objective of experimental scenarios is to test causality relationships between variables (Shadish et al., 2002). That is, the researcher, when manipulating an independent variable *vs.* in order to eliminate possible confounding factors, establishes a causal relationship between this variable and a dependent variable y (Lonati et al., 2018).

The first point to highlight is that, both in the case of scenarios and in all experimental designs, it must be ensured that the only change in the experiment is manipulation of the independent variables (along with participant randomization), in order to ensure its internal validity (Shadish et al., 2002). In practical terms, for the experimental scenario, this means that the researcher must plan the part of the scenario that remains constant under all conditions and manipulate, whether between-participants or intra-participants, only the independent variables (Aguinis & Bradley, 2014; Kadous & Zhou, 2018).

For instance, Bonner et al. (2014) use a series of experiments with between-participants (in which each participant is part of only one condition) and intra-participants (in which participants are part of more than one condition) to assess the effects of mental accounting (Thaler, 1999): a) in the choice of aggregation/disaggregation of earnings statement items by managers; and b) in company assessment by investors in view of the aggregation/disaggregation of items.

Assuming the role of chief executive officer (CEO), participants assess whether to aggregate or disaggregate the items according to the conditions existing in experiment one (condition of small gains vs. condition of large losses), in experiment two (condition of large gains vs. condition of small losses), and experiment three (condition of high transparency vs. condition of extremely high transparency). In turn, in experiment four, participants take the role of investor responsible for choosing between two companies that submit their results in an aggregated or disaggregated way (intra-participants) to carry out their investments. Additionally, participant-investors receive manipulations about the financial results of companies (between-participants).

In short, with regard to differences in scenarios and manipulations, Bonner et al. (2014) keep the scenarios introduced to the participants constant, in which all information remains the same, regardless of the condition to which the participant is allocated, changing only the information relevant to each intra-participants and inter-participants condition. Finally, it is up to the researcher to use a previously validated experimental scenario, slightly modify an already validated scenario or create their own scenario that fits their research problem (remembering to validate them through pre-tests), however, compliance with the first point to highlight must be ensured so that the study's internal validity is not put into question.

The second point to highlight is about the participants used in the experiment, especially in the case of experimental scenarios. There is a very recurrent discussion about who should be the participants in experimental studies (Ashton & Kramer, 2006; Trottier & Gordon, 2018). It is usual for students, for instance, to be used by researchers for experimental research, including experimental scenarios (Ashton & Kramer, 2006). The reasons for the choice are clear: the availability of students is greater than that of auditors, accountants, or managers and their responsiveness to testing the theories under study tends to avoid confounding elements in the research (such as various experiences that lead to different interpretations, for instance) (Libby et al., 2002; Mortensen et al., 2012).

There is a certain consensus about the contexts in which the use of student-participants is possible, such as for the investigation of behavioral responses (decision making), which focus on responses concerning the situation introduced (be it an experimental scenario or an experimental game) (Mortensen et al., 2012). Also, studies such as Houghton and Hronsky (1993) empirically demonstrate the cognitive comparability between groups of students and accounting practitioners. Additionally, Liyanarachchi and Milne (2005) also found similar results between student and investor behavior regarding the pattern of resource allocation in short- and long-term study investments. Along the same line, Elliott et al. (2007) find results that assert the use of MBA students as a good proxy for individual investors when task complexity is adequate for the students' knowledge. Finally, Trottier and Gordon (2018) also find results demonstrating that students make similar decisions to managers regarding asset impairment assessments. However, some care and concerns need to be taken when using students indiscriminately in research. Thus, it is necessary for the researcher to assess how much knowledge is needed for correct interpretation and evaluation of the experimental scenario (Libby et al., 2002; Liyanarachchi, 2007).

Kadous et al. (2019) used three experiments to investigate the effects of intrinsic motivation, leadership emphasis on goals, and ambiguity in the audit question with regard to the audit team member's ability to address

their concerns ('speak up') about audit matters, more specifically, issues that could hinder an effective audit. The authors adopt various samples for the three experiments, assessing the participant's necessary knowledge for each study. In the first experiment, to assess the effects of intrinsic motivation (internal *vs*. external *vs*. control), the authors used (senior) undergraduate and graduate students who had at least one year of experience as audit interns and who had mostly accepted jobs in audit firms. These students served as a proxy for staff auditors, since they had the necessary knowledge to perform the same tasks as these auditors.

Note that the authors use a cognitive variable (intrinsic motivation) to investigate a research question that depends on the participant's formal knowledge of these audit matters. In this domain, participant sample is enough to ensure the comprehensibility of the case at stake, since all of them had already performed similar tasks.

Aguiar & Suave (2021) address managers' preferences for aggregating information in internal reporting of results (gains vs. losses) and the purpose of the report (review of goals vs. performance assessment). The authors explain that because they use the mental accounting theory (Thaler, 1999) to explain the relationships between variables, which consists of a cognitive theory, there is no need for a sophisticated sample (with practitioners), in line with recommendations by Libby et al. (2002).

Additionally, it is necessary to assess other points that may influence participant choice, such as knowledge acquired through experience (Mortensen et al., 2012). Experiments two and three in Kadous et al. (2019) use staff auditors from an auditing firm, given the need for participants to have experience with senior auditors as supervisors in order to emulate participant behavior when changing leadership emphasis (focused on intrinsic goals vs. focused on extrinsic goals). Likewise, elements such as age, culture, and gender should also be observed in choosing participants when there is an emphasis on the part of the researcher on these elements (Trottier & Gordon, 2018). Finally, it is understood that the choice to use students or practitioners in the research must be driven by the research question and by the theory that explains the phenomenon analyzed by the researcher.

The third point to highlight is the concern with the social desirability bias (Eckerd et al., 2021; Lonati et al., 2018). The social desirability bias is aligned with the human tendency to want to appear more altruistic, more intelligent, more honest, among other adjectives, it can also be a characteristic of the individual's personality, in which there is a need for social approval, or a response to a theme/topic, based on the most specific research question (Randall & Fernandes, 1991). Despite the differences in the sources of desirability, the literature points out similar ways of dealing with bias.

Although not exclusive to accounting research, researchers in the field should be concerned about this bias given the design of many studies that assess decision-making in ethical/unethical and honest/dishonest contexts (Chung & Monroe, 2003). In practical terms, the social desirability bias can affect how a participant sees him/herself and, consequently, how he/she responds to his/her behavior or intent to behave (Wason et al., 2002). That is, it is necessary for the researcher to look for ways to mitigate the problem so that he/she can ensure that his/her findings are not affected by participant bias.

The possibilities to deal with the problem are not mutually exclusive and can be used simultaneously. The first possible measure to reduce the social desirability bias is to assure research participant anonymity. Research shows that the level of social desirability bias is inversely proportional to research anonymity (Nederhof, 1985). Therefore, the researcher must, ensure the participant that his/her answers will be anonymized, that participant identification is impossible, and that the case is fictitious or hypothetical. This care provides the participant with greater assurance regarding the lack of relationship between the scenario and personal situations or actual consequences (Randall & Fernandes, 1991). Some researchers also argue about the possibility of replacing a detailed description of the study purpose with a rather generic alternative (Larson, 2019), for instance, explaining that the study purpose is to inquire about decision-making in organizations. The researcher, however, must be careful not to deliberately deceive the participant since this is, at the very least, frowned upon by research ethics committees (Eckerd et al., 2021).

Another possibility to tackle social desirability bias involves the researcher's decision regarding the building of the experimental scenario in first person (direct way) or third person (indirect way) from the participant's viewpoint. Research shows that individuals tend to believe they are more ethical than their peers (Randall & Fernandes, 1991). This makes self-report questions (in first person) subject to questioning with regard to data reliability, since the participant may be subject to social desirability bias. That is, the participant may believe, when answering, that he/she might have more ethical and/or honest attitudes when comparing his/her attitudes in a similar actual situation.

Thus, it is usual for many researchers to write the experimental scenario and ask the participant about his/her behavior indirectly, specifically in third person

(Fisher, 1993; Wason et al., 2002). For instance, instead of formulating the scenario by asking the participant to take the role of manager, as in Johnson et al. (2012), the researcher can ask participants to imagine a manager in the same situation. Some researchers, when creating the scenario and asking the participant in first person, also add a second measurement in the dependent variable by asking the participant about how he/she imagines that other people in the situation would behave, as seen in Murphy et al. (2019).

Another way to mitigate the possibility of social desirability bias is to bring actual consequences to the participant's decision (Lonati et al., 2018). For instance, in Church et al. (2019), the participant takes the role of a division manager who needs to make a decision about investments related to aspects of corporate social responsibility (CSR). Faced with the manipulation of the investment measurement base (financial *vs*. non-financial) and the measurement of the participant's personal standard regarding CSR (supporter *vs*. non-supporter), the authors capture the participant's behavior by making him/her suffer an actual economic consequence according to his/her investment. In the scenario, the manager has discretionary funds that can be spent ($ 0 to $ 1,000) to plant trees. In order to operationalize actual economic consequences, the authors established to the participants that 1% of the value of investments for planting trees would be donated to a non-governmental organization (NGO) devoted to this purpose and that 1% of the entire amount that was not invested in planting would be given to the participant at the end of the experiment.

Another possible way to check the social desirability bias is to measure this effect through some scale. This happens in Johnson et al. (2012), who used a reduced version of the Marlowe-Crowne scale (Crowne & Marlowe, 1960). Using the social desirability scale is also in line with the possibility that desirability is a respondent's characteristic and not necessarily a bias regarding the researcher's theme (Randall & Fernandes, 1991). In practice, measurement by scale allows using the result as a control variable for the researcher's findings.

Finally, given the various ways of measuring and controlling social desirability bias, it is important that the researcher focuses on minimizing the bias and assesses the best method to be used in the face of his/her own research question, in order to avoid problems with construct validity since the measurement of the variable was incorrect.

The fourth point to highlight is the choice between inquiring about a participant's behavior or intent to behave. Unlike experimental games, which assess the participant's behavior in the face of interventions, the experimental scenario often assesses intents to behave (Wason et al., 2002). This means that the researcher asks the participant what his/her intent is, given the experiment interventions, to engage in some type of behavior. As seen in Austin et al. (2021), which investigated the participant's intent to invest a hypothetical amount between US$ 0 to US$ 10,000 in stocks of the company introduced.

It is worth grasping that most experimental scenarios use intent measurement given the impossibility of measuring actual consequences for the participant (Eckerd et al., 2021). Also, the Brazilian research ethics code does not allow participants to be paid for their participation in any type of study (including experiments in social sciences/applied social sciences) as it is based on Resolution No. 466/2012 of the Ministry of Health. This resolution makes it more difficult to carry out studies in which participants receive any remuneration for participation or a variable remuneration dependent on their behavior in the experiment. Thus, the use of intent as a measure for the dependent variable can be a good alternative.

The fifth point to highlight is on where to look for the desired participant pool for the research. That is, who are the participants who will respond to the survey. The choice of the participant pool is linked to a series of subsequent assessments such as the possibility that the researcher manage the experiment personally or to do it online, and if he/she needs participant contacts or if he/she uses crowdsourcing platforms to apply the experiment (e.g. *MTurk*, *Prolific.ac*, *LinkedIn*, etc.). In short, the researcher must evaluate the availability and costs for carrying out the research.

So, given the possibility of using undergraduate or graduate students, as seen in the second point to highlight, it is usual for researchers to conduct their experiments in laboratories or in the classroom. Experimental laboratories for the areas of applied social, social, and human research are quite different from the laboratories used for research in biological areas, with the necessary instruments, lab coats, and materials. In practice, the experimental laboratory for applied social sciences is a room with several computers, which can be accessed by the researcher and which have a distance that allows the researcher to be able to observe participants while participating in the experiment and so that the participants themselves do not 'peek' at the computer next to them (usually by dividing participants into individual cubicles).

The use of students also allows the researcher to apply the experiment in the classroom, with undergraduate, M.A. or MBA students. These experiments are often done having the experimental scenario printed on paper. Again,

it is worth ensuring that the participants are distanced so that there is no contact or conversation during the study. In this case, the researcher also remains in the room during the experiment. Both experiments, in the experimental laboratory and in the classroom, manage to ensure greater internal validity given the researcher's control while it takes place, preventing participants from talking to each other or from ' peeking' what the colleague is answering (Bloomfield et al., 2016; Lourenço, 2019).

That said, a major aspect of research with experimental scenarios is that, regardless of where the experiment is being conducted, whether in the physical laboratory space, in the classroom, or online, every experimental scenarios is regarded as a laboratory experiment (Bloomfield et al., 2016). Although researchers such as Floyd and List (2016) classify any experiment that does not involve a sample of students as a field experiment or an artefactual experiment, experiment classification involves how the context was defined and not its sample (Bloomfield et al., 2016). Thus, any experimental scenario, which has been developed by a researcher, is classified as laboratory. Another aspect that differentiates the laboratory experiment from the field one is participant awareness regarding their participation in the experiment. Unlike the former, which requires an informed consent form before the study onset, the second, usually, does not rely on participant awareness regarding their participation (Bloomfield et al., 2016; Lourenço, 2019).

If the researcher needs expert participants, the experiment conducted in the classroom and laboratory becomes more complicated due to the limited time available for these professionals. Thus, the most common thing is that the researcher searches on the social media or crowdsourcing platforms professionals who have characteristics that fit their research needs. Studies like Brandon et al. (2014) cite social media platforms such as *Twitter*, *Facebook*, and *LinkedIn* as means of distribution for the experimental study, however, by the same logic, the search for participants who fit the researcher's specifications can also take place on the same social media platforms. *LinkedIn*, for instance, may be a useful tool to look for professional participants, with the caveat that there is a considerable demand for time to collect contacts, contact them, and ensure that they participate in the research. Also, the researcher must trust the participants to respond carefully and personally to the survey, losing part of his/her control over the experiment and, consequently, part of the internal validity.

Additionally, it is usual for researchers to use crowdsourcing platforms like *Amazon Mechanical Turk* (*MTurk*) and *Prolific.ac* to apply their research (Aguinis & Ramani, 2021). In practice, thousands of people have the possibility to register on the platforms, providing personal information at registration, such as: age, gender, education, profession, nationality. And a series of more specific information, such as: if you work with more people, if you work in home office, in addition to other information. With this, the researcher can select, before carrying out the study, the specific pool to which he/she wishes to make his/her research available. The number of people registered on the platforms brings benefits such as ease of access and speed of data collection (Aguinis & Ramani, 2021). However, these benefits have a considerable cost, since the payment is due to the number of participants and can be in dollars, in the case of *MTurk*, or in pounds sterling, in the case of *Prolific.ac*.

Another negative point is data reliability. In the case of *MTurk*, participants may respond to the survey inattentively, may lie about demographic details, may not have sufficient fluency in English to grasp the experimental scenario, and may have had contact with scenarios similar to the one introduced previously (Aguinis & Ramani, 2021). The same issues are also usual with *Prolific*. However, because it consists of a rather research-oriented platform and due to the smaller volume of studies available on a daily basis, it is less likely that participants use the platform as a means of obtaining any remuneration.

Peer et al. (2017) assess the differences between the platforms *Amazon MTurk*, *CloudResearch*, and *Prolific*. The researchers found results that support the superiority of *Prolific* and *CloudResearch* over *MTurk* in terms of data quality, better response rate, and less dishonesty.

In practice, the researcher can use some assumptions when choosing participants on these platforms to ensure higher quality responses. Regarding the issue of false identity or fluency problems in English, the researcher can prepare questions that ensure participant filtering, and later invite only approved participants for a specific study (Aguinis & Ramani, 2021).

Another possibility is using the same questions for participant validation, eliminating from the study the participants who do not get the questions right, in a kind of confirmation check regarding the information provided by the participant. Also, within the platforms there are some indices that may be valid for researchers, such as participant approval rate, which both for *MTurk*

and for *Prolific*, are filtering tools for determining study participants.

Finally, it is worth seeing that there is a trade-off with regard to the researcher's level of control when using a participant pool from a crowdsourcing platform when compared to the possibility of carrying out an experiment in a physical laboratory. However, the previous literature shows acceptance of the use of participants from platform like *Amazon MTurk* if data reliability is ensured.

Table 1 brings together the highlights mentioned in this topic.

Table 1
Summary of highlights regarding the building of experimental scenarios

Highlights	Decisions about highlights in experimental scenarios
Manipulations	When building the experimental scenario, researchers must take care that the only change in the various treatments (conditions) consists in manipulations.
Participants	Participants must be selected according to the researchers' needs, ensuring that they have the necessary knowledge to grasp the scenario.
Social desirability bias	Social desirability bias is usual in experimental scenarios that deal primarily with ethical issues. Measures such as building the scenario in third person, measuring social desirability through a scale, or bringing actual consequences to the participants' decision help to mitigate the effects of bias.
Behavior/action or intent to behave	It is usual for experimental scenarios to inquire both about the intent to behave and about a specific behavior.
Participant pool	The participants pool can be selected from a number of possibilities such as: undergraduate or graduate students, social media platforms (e.g. *LinkedIn*), or crowdsourcing platforms (e.g. *Amazon MTurk, Prolific*). The choice of where to look for the participants pool should be based on the need for specific knowledge and the costs associated with each choice.

Source: *Prepared by the author.*

4. CONCLUSION AND POSSIBILITIES FOR FURTHER RESEARCH

The main objective of this article is to systematize relevant points in the building of experimental scenarios as a research method for Brazilian accounting researchers. In the study, the main differences between the use of experimental scenarios and experimental games were explored, such as the level of instrument contextualization, the focus of research, and the choices relevant for participants to understand when lacking contextualization (usual to experimental games). Next, the highlights regarding the building of the experimental scenario were discussed, such as: the introduction of manipulations, participant choice, concerns with social desirability bias, measurement of behavior or intent to behave as a dependent variable, and, finally where to search for the participant pool. All these highlights must be assessed by researchers according to the theme of their studies and their availability of resources.

Themes usual to the financial accounting literature, such as the level of disclosure, form of presentation, and characteristics of information disclosed by companies, are extensively explored in the Brazilian national literature through archival studies. In the international literature, the same themes also tend to be explored through experimental scenarios to ascertain the effects of these choices on investors' decision-making, as seen in Asay et al. (2017) and Elliott et al. (2015). In this sense, the use of experimental studies is suggested as a way to complement and triangulate the results (Bloomfield et al., 2016) that were previously found in studies with Brazilian companies.

Themes usual to the management accounting literature, such as the influence of various management controls – whether in the form of packages or observed individually – which are often explored through surveys, can be complemented by experiments that investigate the effects of combining these controls, as proposed in Choi (2020). Also, themes usual to the audit literature, such as the relationship between clients and auditors, auditors' judgments and audit quality, which have also been addressed in the Brazilian national literature through archival studies, surveys, and interviews, can be explored through experiments. as seen in Kadous and Zhou (2019).

Finally, this study has limitations regarding the highlights, which were not exhaustively presented, and regarding the previous literature introduced in

the examples, which were occasional choices made by the author. It is key that the reader seeks a research method that is the most coherent with his/her research question, however, advancement of the investigation of theories and the measurement of causality are major parts of the research and can bring great progress to the Brazilian accounting research. Additionally, the researcher interested in building an experimental scenario has a vast international literature, which can help through examples and discussions about the method. In this way, it is understood that the objective of this study is also to encourage the search for more granular information that ensure the building of the appropriate experimental scenario for each research question.

REFERENCES

Aguiar, A. B. (2017). Pesquisa experimental em contabilidade: propósito, desenho e execução. *Advances in Scientific and Applied Accounting*, *10*(2), 224-244. https://doi.org/10.14392/asaa.2017100206

Aguiar, A. B. (2021). Is value statement an effective informal control for stimulating pro-environmental behaviors? *Revista de Contabilidade e Finanças*, *32*(86), 193-206. https://doi.org/10.1590/1808-057x202011500

Aguiar, A. B., & Suave, R. (2021). Managers' presentation preferences for aggregation: The role of internal reporting outcome, purpose, and Machiavellianism. *Spanish Journal of Finance and Accounting*, *51*(1), 21-49. https://doi.org/10.1080/02102412.2021.1952517

Aguinis, H., & Bradley, K. J. (2014). Best practice recommendations for designing and implementing experimental vignette methodology studies. *Organizational Research Methods*, *17*(4), 351-371. https://doi.org/10.1177/1094428114547952

Aguinis, H., & Ramani, R. S. (2021). MTurk research: Review and recommendations. *Journal of Management*, *47*(4), 823-837. https://doi.org/10.1177/0149206320969787

Arnold, V., & Triki, A. (2018). Use of student and online participants in behavioural accounting research. In T. Libby, & L. Thorne (Eds.), *The Routledge companion to behavioural accounting research* (pp. 287-299). https://doi.org/10.4324/9781315710129-20

Asay, H. S., Elliott, W. B., & Rennekamp, K. (2017). Disclosure readability and the sensitivity of investors' valuation judgments to outside information. *The Accounting Review*, *92*(4), 1-25. https://doi.org/10.2308/accr-51570

Ashton, R. H., & Kramer, S. S. (2006). Students as surrogates in behavioral accounting research: Some evidence. *Journal of Accounting Research*, *18*(1), 1-15. https://doi.org/10.2307/2490389

Austin, C. R., Bobek, D. D., & Harris, L. L. (2021). Does information about gender pay matter to investors? An experimental investigation. *Accounting, Organizations and Society*, *90*, 101193. https://doi.org/10.1016/j.aos.2020.101193

Bloomfield, R., Nelson, M. W., & Soltes, E. (2016). Gathering data for archival, field, survey, and experimental accounting research. *Journal of Accountig Research*, *54*(2), 341-395. https://doi.org/10.1111/1475-679X.12104

Bonner, S. E., Clor-Proell, S. M., & Koonce, L. (2014). Mental accounting and disaggregation based on the sign and relative magnitude of income statement items. *The Accounting Review* *89*(6), 2087-2114. https://doi.org/10.2308/accr-50838

Brandon, D. M., Long, J. H., Loraas, T. M., Mueller, J. M., & Vansant, B. (2014). Online instrument delivery and participant recruitment services: Emerging opportunities for behavioral accounting research. *Behavioral Research in Accounting*, *26*(1), 1-23.

Choi, J. W. (2020). Studying "and": A perspective on studying the interdependence between management control practices. *Accounting, Organizations and Society*, *86*, 101188. https://doi.org/10.1016/j.aos.2020.101188

Chung, J., & Monroe, G. S. (2003). Exploring social desirability bias. *Journal of Business Ethics*, *44*, 291-302.

Church, B. K., Jiang, W., Kuang, X. J., & Vitalis, A. (2019). A dollar for a tree or a tree for a dollar? The behavioral effects of measurement basis on managers' CSR investment decision. *The Accounting Review*, *94*(5), 117-137.

Collett, J. L., & Childs, E. (2011). Minding the gap: Meaning, affect, and the potential shortcomings of vignettes. *Social Science Research*, *40*(2), 513-522. https://doi.org/10.1016/j.ssresearch.2010.08.008

Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, *24*(4), 349-354. https://doi.org/10.1037/h0047358

Eckerd, S., DuHadway, S., Bendoly, E., Carter, C. R., & Kaufmann, L. (2021). On making experimental design choices: Discussions on the use and challenges of demand effects, incentives, deception, samples, and vignettes. *Journal of Operations Management*, *67*(2), 261-275. https://doi.org/10.1002/joom.1128

Elliott, W. B., Hodge, F. D., & Kennedy, J. J. (2007). Are M.B.A. students a good proxy for nonprofessional investors ? *The Accounting Review*, *82*(1), 139-168. https://doi.org/10.2308/accr.2007.82.1.139

Elliott, W. B., Rennekamp, K. M., & White, B. J. (2015). Does concrete language in disclosures increase willingness to invest ? *Review of Accounting Studies*, *20*, 839-865. https://doi.org/10.1007/s11142-014-9315-6

Fisher, R. J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, *20*(2), 303-315. https://doi.org/10.1086/209351

Floyd, E., & List, J. A. (2016). Using field experiments in accounting and finance. *Journal of Accounting Research*, *54*(2), 437-475. https://doi.org/10.1111/1475-679X.12113

Frezatti, F., Aguiar, A., Wanderley, C., & Malagueño, R. (2015). A pesquisa em contabilidade gerencial no Brasil: Desenvolvimento, dificuldades e oportunidades. *Revista*

Universo Contábil, *11*(1), 47-68. https://doi.org/10.4270/ruc.2015103

Frezatti, F., Bido, D. S., Cruz, A. P. C., & Machado, M. J. C. (2015). A Estrutura de artefatos de controle gerencial no processo de inovação: Existe associação com o perfil estratégico? *Brazilian Business Review*, *19*, 129-156. https://doi.org/10.15728/bbr.2015.12.1.6

Grasser, R. A., Majerczyk, M., Staehle, M., & Sydney, U. (2021). The benefits of deliberative involvement in the design of incomplete feedback systems. *Contemporary Accounting*, *38*(3), 2351-2375. https://doi.org/10.1111/1911-3846.12675

Gravetter, F. J., & Forzano, L. A. B. (2012). *Research methods for the behavioral sciences* (4th ed.). Cengage Learning.

Ham, C., Lang, M., Seybert, N., & Wang, S. (2017). CFO narcissism and financial reporting quality. *Journal of Accounting Research*, *55*(5), 1089-1135. https://doi.org/10.1111/1475-679X.12176

Hartmann, F. G. H., & Maas, V. S. (2010). Why business unit controllers create budget slack: Involvement in management, social pressure, and Machiavellianism. *Behavioral Research in Accounting*, *22*(2), 27-49. https://doi.org/10.2308/bria.2010.22.2.27

Haynes, C. M., & Kachelmeier, S. J. (1998). The effects of accounting contexts on accounting decisions: A synthesis of cognitive and economic perspectives in accounting experimentation. *Journal of Accounting Literature*, *17*, 97-136.

Hesford, J. W., Lee, S. H. S., Van der Stede, W. A., & Young, S. M. (2006). Management accounting: A bibliographic study. *Handbooks of Management Accounting Research*, *1*, 3-26. https://doi.org/10.1016/S1751-3243(06)01001-7

Houghton, K. A., & Hronsky, J. J. F. (1993). The sharing of meaning between accounting students and members of the accounting profession. *Accounting & Finance*, *33*(2), 131-147. https://doi.org/10.1111/j.1467-629X.1993.tb00324.x

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm managerial behaviour, agency costs and ownership structure. *Journal of Financial Economics*, *3*, 305-360. http://uclafinance.typepad.com/main/files/jensen_76.pdf

Johnson, E. N., Fleischman, G. M., Valentine, S., & Walker, K. B. (2012). Managers' ethical evaluations of earnings management and its consequences. *Contemporary Accounting Research*, *29*(3), 910-927. https://doi.org/10.1111/j.1911-3846.2011.01135.x

Kadous, K., Proell, C. A., Rich, J., & Zhou, Y. (2019). It goes without saying: The effects of intrinsic motivational orientation, leadership emphasis of intrinsic goals, and audit issue ambiguity on speaking up. *Contemporary Accounting Research*, *36*(4), 2113-2141. https://doi.org/10.1111/1911-3846.12500

Kadous, K., & Zhou, Y. (2018). Maximizing the contribution of JDM-style experiments in accounting. In T. Libby & L. Thorne (Eds.), *The Routledge companion to behavioural accounting research* (pp. 175-192). https://doi.org/10.4324/9781315710129-13

Kadous, K., & Zhou, Y. D. (2019). How does intrinsic motivation improve auditor judgment in complex audit tasks? *Contemporary Accounting Research*, *36*(1), 108-131. https://doi.org/10.1111/1911-3846.12431

Kinney, W. R. (2019). The Kinney three paragraphs (and more) for accounting Ph.D. students. *Accounting Horizons*, *33*(4), 1-14. https://doi.org/10.2308/acch-52451

Larson, R. B. (2019). Controlling social desirability bias. *International Journal of Market Research*, *61*(5), 534-547.

Libby, R., Bloomfield, R., & Nelson, M. W. (2002). Experimental research in financial accounting. *Accounting, Organizations and Society*, *27*, 775-810.

Liyanarachchi, G. A. (2007). Feasibility of using student subjects in accounting experiments: A review. *Pacific Accounting Review*, *19*(1), 47-67. https://doi.org/10.1108/01140580710754647

Liyanarachchi, G. A., & Milne, M. J. (2005). Comparing the investment decisions of accounting practitioners and students: An empirical study on the adequacy of student surrogates. *Accounting Forum*, *29*(2), 121-135. https://doi.org/10.1016/j.accfor.2004.05.001

Lonati, S., Quiroga, B. F., Zehnder, C., & Antonakis, J. (2018). On doing relevant and rigorous experiments: Review and recommendations. *Journal of Operations Management*, *64*, 19-40. https://doi.org/10.1016/j.jom.2018.10.003

Lourenço, S. M. (2019). Field experiments in managerial accounting research. *Foundations and Trends® in Accounting*, *14*(1), 1-72. https://doi.org/10.1561/1400000059

Lucas, J. W. (2003). Theory-testing, generalization, and the problem of external validity. *Sociological Theory*, *21*(3), 236-253. https://doi.org/10.1111/1467-9558.00187

Mortensen, T., Fisher, R., & Wines, G. (2012). Students as surrogates for practicing accountants: Further evidence. *Accounting Forum*, *36*(4), 251-265. https://doi.org/10.1016/j.accfor.2012.06.003

Murphy, P. R., Wynes, M., Hahn, T. A., & Devine, P. G. (2019). Why are people honest? Internal and external motivations to report honestly. *Contemporary Accounting Research*, *53*(9), 1689-1699. https://doi.org/10.1017/CBO9781107415324.004

Nash Jr, J. F. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences*, *36*(1), 48-49.

Nash Jr., J. F. (1951). Non-cooperative games. *Annals of Mathematics*, 286-295.

Nederhof, A. J. (1985). Methods of coping with social desirability bias: A review. *European Journal of Social Psychology*, *15*, 263-280.

Onderstal, S. (2014). *Economics of organizations and markets*. Pearson.

Osborne, M. J. (2003). *An introduction to game theory*. Oxford University Press.

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the *Turk*: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153-163. https://doi.org/10.1016/j.jesp.2017.01.006

Randall, D. M., & Fernandes, M. F. (1991). The social desirability response bias in ethics research. *Journal of Business Ethics*, *10*(11), 805-817.

Santos, L. P. G., Cerqueira, A. J. F., & Carvalho, C. V. O., Júnior. (2021). Uma análise experimental do efeito da manutenção de registros sobre a reciprocidade direta. *Revista Contabilidade & Finanças*, *32*(86), 359-375. https://doi.org/10.1590/1808-057x202010800

Schuhmacher, K., Towry, K. L., & Zureich, J. (2021). Leading by example in socially driven organizations: The effect of transparent leader compensation contracts on following. *The Accounting Review*, *97*(3), 373-393. https://doi.org/10.2308/tar-2019-0227

Shadish, W. R., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Cengage Learning.

Suave, R., Altoé, S. M. L., & Ferreira, M. M. (2021). Pesquisas experimentais aplicadas à educação contábil: Panorama atual e oportunidades no cenário brasileiro. *Revista Contemporânea de Contabilidade*, *18*(47), 155-176. https://doi org/10.5007/2175-8069.2021.e77682

Swieringa, R. J., & Weick, K. E. (1982). An assessment of laboratory experiments in accounting. *Journal of Accounting Research*, *20*, 56-101.

Tayler, W. B., & Bloomfield, R. (2011). Norms, Conformity and controls. *Journal of Accountig Research*, *46*(2), 1-23. https://doi.org/10.1111/j.l475-679X.2011.00398.x

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, *12*, 183-206.

Trevino, L. K. (1992). Experimental approaches to studying ethical-unethical behavior in organizations. *Business Ethics Quarterly*, *2*(2), 121-136.

Trottier, K., & Gordon, I. M. (2018). Students as surrogates for managers: Evidence from a replicated experiment. *Canadian Journal of Administrative Sciences*, *35*(1), 146-161. https://doi.org/10.1002/cjas.1377

Von Neumann, J., & Morgenstern, O. (2004). *Theory of games and economic behavior* (6th ed.). Princeton University Press.

Wason, K. D., Polonsky, M. J., & Hyman, M. R. (2002). Designing vignette research in marketing. *Australasian Marketing Journal*, *10*(3), 41-58. https://doi.org/10.1016/S1441-3582(02)70157-2