**WILEY** | Journal of Organizational Behavior

# Thought experiments: Review and recommendations

Herman Aguinis[1] ⓘ | Jose R. Beltran[2] ⓘ | Estelle E. Archibold[3] ⓘ |
Esther L. Jean[4] ⓘ | Darryl B. Rice[5] ⓘ

[1]Department of Management, The George Washington University, Washington, D.C., USA

[2]Department of Management, Iowa State University, Ames, Iowa, USA

[3]Department of Management and Organization, The Pennsylvania State University, State College, Pennsylvania, USA

[4]Department of Management, University of Texas at Arlington, Arlington, Texas, USA

[5]Department of Management, Miami University, Oxford, Ohio, USA

**Correspondence**
Herman Aguinis, Department of Management, The George Washington University, Washington, D.C., USA.
Email: haguinis@gwu.edu

**Summary**

Thought experiments have been used as an effective methodological approach to advance theory in numerous scientific fields. However, they are underutilized in organizational behavior (OB) and adjacent fields. Accordingly, we conducted a comprehensive and multidisciplinary literature review of thought experiments that entailed 174 sources in economics, psychology, marketing, medicine, sociology, finance, and other fields. We used insights from this literature review to define and describe the unique nature of thought experiments and offer a taxonomy of four main types based on a theory's development stage (i.e., early vs. late) and a study's theoretical goal (i.e., confirmation vs. disconfirmation). We also provide a decision-making tree useful for evaluating whether conducting a thought experiment is beneficial for a particular research situation and which of the four types is most likely to produce a meaningful contribution. Then, we offer best-practice recommendations for conducting thought experiments that address how to plan, execute, report results, and discuss implications. In addition, we demonstrate the potential of thought experiments by using the best-practice recommendations to design and conduct a thought experiment in the domain of workplace allyship. Finally, we offer suggestions for future substantive research that would benefit from thought experiment methodology (i.e., diversity, equity, and inclusion; leadership; performance; selection and recruitment; teams; and turnover). Overall, our article offers a comprehensive review and recommendations that we hope will be a catalyst for using thought experiments to advance theory in OB and related fields.

**KEYWORDS**
methodology, research design, theory advancement, thought experiments

# 1 | INTRODUCTION

Thought experiments are judgments about what would happen if an imagined scenario were real (Gendler, 1998, 2010; Kornberger & Mantere, 2020). Thought experiments are also commonly referred to as thought trials (Dietrich & Haider, 2015), imaginary illustrations (Lennox, 1991), and metatheorizing (Carr & Zanetti, 1999). As noted by Folger and Turillo (1999), "thought experiments illustrate one way that theorizing can bridge the gap between the abstract and the concrete" (p. 742). Thought experiments have effectively created novel insights and advanced theory in numerous scientific fields for decades (e.g., Brown, 2011; Sorensen, 2017). For example, thought experiments have been used to reconceptualize how artificial intelligence can change information systems (Davies, 1989), deconstruct constructivism in philosophy (Kruger, 2002), understand a plausible strategy for determining when modifications are desirable regarding tropical cyclone landfall predictions (Klima & Morgan, 2012), and challenge the conventional profits-versus-people tension in economics (Hatherly et al., 2020). Overall, by presenting hypothesized alternate explanations, extending extant theory to include new contexts, and providing counterexamples for prevailing theories, fields such as economics, public and international policy, physics, ethics, and many others have utilized thought experiments to produce novel and useful theoretical insights.

Unfortunately, except for a few notable exceptions (e.g., Cornelissen & Durand, 2012; Folger & Turillo, 1999), the use of thought experiments is virtually absent from research in organizational behavior (OB) and adjacent fields such as human resource management, industrial and organizational psychology, entrepreneurship, and strategy. We believe there are at least two reasons for the underutilization of thought experiments. First, thought experiments are not usually taught in research methods courses in doctoral programs in management and related fields (e.g., Tett et al., 2013). Thus, OB researchers may not have sufficient opportunities to acquire the competencies needed to conduct thought experiments. Second, most OB researchers may not know what thought experiments are and their benefits. Relatedly, on a more pragmatic note, most researchers may not be aware that thought experiments are cost-efficient compared with more traditional types of experiments and can be conducted for little to no financial cost.

The purpose of our article is to demonstrate the usefulness of thought experiments to advance theory in OB and adjacent fields such as human resource management, industrial and organizational psychology, entrepreneurship, and strategy. First, we describe procedures we implemented in conducting a comprehensive and multidisciplinary literature review. Second, we define and describe the unique nature of thought experiments. Third, we offer a taxonomy of four main types based on a theory's development stage (i.e., early vs. late) and a study's theoretical goal (i.e., confirmation vs. disconfirmation). Fourth, based on our review, we offer best-practice recommendations for deciding whether to conduct a thought experiment and which type, and how to plan and execute as well as report results and discuss implications. Fifth, to not only tell but also show the potential of thought experiments, we implement the best-practice recommendations to design and conduct a thought experiment in the domain of workplace allyship. Finally, we offer suggestions for future substantive research that would benefit from thought experiment methodology in the illustrative areas of diversity, equity and inclusion (DEI), leadership, performance, selection and recruitment, teams, and turnover. Overall, our article offers best-practice recommendations based on a comprehensive review that we hope will be a catalyst for using thought experiments to advance theory in OB and related fields.

# 2 | LITERATURE REVIEW

We used the Business Source Complete (i.e., 3,427 business sources) and PsychINFO (i.e., 2,275 behavioral and social science sources) databases to conduct our review. These two databases include academic journals, trade publications, book chapters, dissertations, and magazines. Our search covered the period from January 1926 through April 2021. Given the different terminology used to refer to thought experiments across fields, we used the following list of keywords: thought experiment, imagined scenario, imaginary experiments, gedankenexperiment, metatheorizing, thought trials, imaginary cases, and trolley problem.

Our initial search resulted in 999 records, 580 from Business Source Complete, and 419 from PsychINFO. Many records were identified by both databases because of their overlap in journal coverage. Therefore, we first eliminated duplicates, which resulted in 772 unique records. Then, we eliminated records that, although they had included at least one of our keywords, did not actually describe a thought experiment. For example, some included at least one of our keywords, but the article described a computer simulation or empirical experiment involving data. Finally, we also eliminated sources that merely referred to thought experiments as a method that could be used for future research without actually conducting an experiment or offering design or implementation recommendations about thought experiments. In the end, our review relied on 174 sources: 146 journal articles, 12 book chapters, eight law reviews, four magazine articles, one book, one dissertation, one essay, and one online content item. In the interest of transparency, the list of the 174 sources is included in the Supporting Information.

Our review revealed that thought experiments have been used to address a variety of topics across several fields including economics (Mankiw, 2013), psychology (Fischer, 2009; Holt, 1999; Nanay, 2015), marketing (Christian, 1963; Van Bockhaven & Matthyssens, 2021), medicine (Burton et al., 2006), sociology (Miller, 2006; Turner, 2017), finance (Vasileiou, 2021), and strategy (Bankins & Waterhouse, 2019), among others. Regarding the field of OB, the list of sources in the Supporting Information shows that only 14 articles address OB-related topics (e.g., Bankins & Formosa, 2020; Cornelissen & Durand, 2012; Folger & Turillo, 1999). To put this meager number in perspective, consider that several dozen journals publish OB research and *Journal of Organizational Behavior* and *Journal of Applied Psychology* alone have published more than 20,000 articles since their inception. Our finding that only 14 articles describe thought experiments provides empirical evidence about their very low rate of use, particularly

compared with other much more popular types of research designs and data collection procedures (e.g., Aguinis et al., 2009; Scandura & Williams, 2000).

## 3 | UNIQUE DEFINING CHARACTERISTICS OF THOUGHT EXPERIMENTS

Thought experiments take place in a unique setting: the laboratory of the mind (Brown, 2011). As such, thought experiments are uniquely distinct from more traditional types of experiments, including those involving hypothetical scenarios such as vignette studies and Monte Carlo simulations. Specifically, thought experiments are unique regarding the (1) nature of independent and dependent variables, (2) number of independent variables that can be manipulated, (3) type of and sample size, and (4) ethical and legal considerations associated with implementing the experiment.

First, the independent and dependent variables are imagined (Hatherly et al., 2020; Kuhn, 1964; McAllister, 1996) and not empirically manipulated or measured. For example, in more traditional field and lab experiments (including vignette studies), researchers include empirical manipulations of independent variables (Aguinis & Bradley, 2014; Cook et al., 2002; Eden, 2017; Highhouse, 2009). However, thought experiments do not include such empirical manipulations of independent variables. Instead, in thought experiments researchers manipulate independent variables in imagined scenarios. For instance, Hong (2012) expanded the understanding of how knowledge management theories are adopted by mentally manipulating the characteristics of routines that may influence this process.

Second, thought experiments are unique because of their ability to include a virtually unlimited number of independent variables. In contrast, traditional empirically-based experiments are limited due to sample size constraints (Ferguson & Ketchen, 1999) or participant limitations such as fatigue, response acquiescence, and social desirability (Schwab, 2005). Relatedly, traditional experiments are also limited by logistical and resource constraints such as the cost and time involved in empirically manipulating too many independent variables. Because thought experiments are conducted in the researcher's mind, they are uninhibited by such resource limitations other than a researcher's imagination.

Another key characteristic distinguishing thought experiments from traditional types of experiments is the study's sample. Traditional OB experiments involve human participants, which require a significant time commitment through recruiting and compensation costs (Aguinis et al., 2021). In addition, in traditional experiments there is a need to obtain a sample size that is large enough to allow for data analyses at adequate levels of statistical power. None of these requirements are of concern for thought experiments because, again, they take place within the mind of a researcher, and therefore do not require human participants.

Finally, traditional experiments in OB and related fields require ethical considerations due to the involvement of people as research participants. Understandably, traditional experiments require appropriate approval from an institutional review board (IRB) that evaluates potential risks. Because thought experiments are imagined, they are not subject to the same ethical and legal considerations as traditional experiments requiring IRB approval. This feature is particularly attractive for researchers interested in investigating socially sensitive phenomena (Barnes et al., 2018) (e.g., racial discrimination, sexual harassment, workplace bullying) or other invasive inquiries, for which traditional experiments pose a substantial risk to participants and therefore require understandably complex IRB approval processes—or are not possible at all (Boser, 2007; Liberale & Kovach, 2017).

## 4 | TAXONOMY OF THOUGHT EXPERIMENTS

In this section, we present a taxonomy of thought experiments. We classify thought experiments along two dimensions: (a) Theory's development stage: Early versus late (Colquitt & Zapata-Phelan, 2007) and (b) study's primary theoretical purpose: Theory disconfirmation versus theory confirmation (Brown, 1986; Popper, 1959). This taxonomy is useful for two purposes. First, it is useful for understanding the different types of thought experiments conducted in the past and their distinct value-added contributions. Second, this taxonomy can be used to choose which type of thought experiment would be most appropriate for a particular situation and particular theory-based goals. Accordingly, this taxonomy is also useful for making design choices in future research involving thought experiments, as we describe in more detail later in our article.

As a preview, Figure 1 includes a summary of the four main types of thought experiments. Next, we describe each type in detail and also provide examples to illustrate each.

### 4.1 | Type I: Early theory stage and theory confirmation purpose

Type I involves a situation where a researcher wants to build theory and hopes to confirm it rather than disconfirm it. A prominent example of this type of thought experiment is Newton's cannonball (Newton, 1728). Early into his work on the theory of gravity and how it affects orbiting celestial bodies, Newton sought to confirm his ideas using a thought experiment. Newton imagined cannons firing cannonballs atop higher and higher mountains with increasing velocity. He surmised through this thought experiment that, eventually, the cannonballs would orbit the earth, confirming his nascent theoretical propositions on the nature of gravity.

### 4.2 | Type II: Early theory stage and theory disconfirmation purpose

Type II involves a thought experiment wherein a counterfactual or paradoxical situation is created in the early stages of theory

**FIGURE 1** Taxonomy of thought experiments and illustrations based on a theory's development stage (i.e., early vs. late) and a study's theoretical goal (i.e., confirmation vs. disconfirmation)
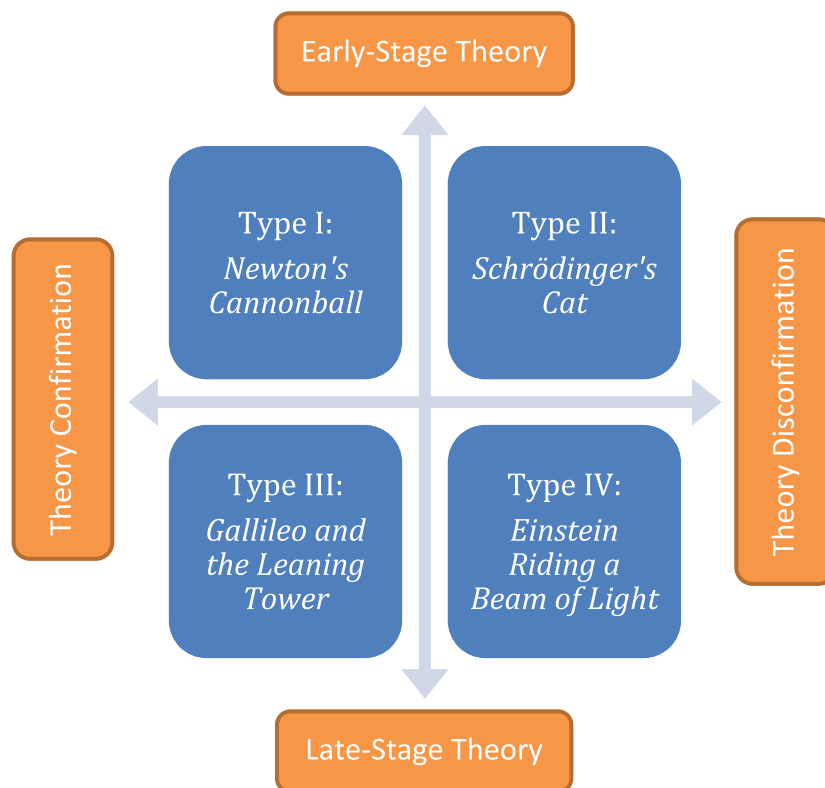


development. Specifically, this type of thought experiment evaluates a nascent theory by reimagining an alternate explanation that must also be plausible. Schrödinger's cat (Schrödinger, 1935) is an exemplar of this type of thought experiment. The fate of a (hypothesized) cat hinges on a subatomic event that may or may not occur according to the Copenhagen interpretation of quantum mechanics, which implies that the cat would be simultaneously alive and dead. However, someone observing the cat in the box would observe it to be alive or dead. Thus, Schrödinger conducted a thought experiment in which the theory he sought to question breaks down or results in a paradoxical event, thus disconfirming this early-stage theory and forcing theoretical refinement in the process of theory building.

## 4.3 | Type III: Late theory stage and theory confirmation purpose

A Type III thought experiment involves attempting to confirm a more developed theory. An example is Galileo and the Leaning Tower of Pisa in which he tested the more developed Aristotelian theory of gravity. According to Moody (1951), this thought experiment took place in the year 1589. Galileo is said to have imagined dropping two bodies from the top of the Tower of Pisa to show that the times of their fall would be equal, although the objects differed regarding their weight. The story is that this thought experiment evolved into Galileo actually climbing to the top of the tower and dropping the objects. According to Moody, this thought experiment is particularly noteworthy because it is paradigmatic in "the interpretation of the history of

ideas" (p. 164). According to other theories, particularly Aristotle's theory of gravity, objects of varying mass were surmised to fall at different rates. Galileo's thought experiment helped to confirm his own already-developed theoretical propositions.

## 4.4 | Type IV: Late theory stage and theory disconfirmation purpose

A Type IV thought experiment creates imaginary paradoxical situations in an attempt to disconfirm a more developed theory. For example, Einstein famously imagined himself simultaneously riding a beam of light and observing that same beam of light from a distance. By implementing a thought experiment in which he imagined what both versions of himself would witness, he created a paradoxical situation given the parameters of the prevalent emission theory of light. Thus, he could disconfirm extant theory by using what he termed a *Gedankenexperiment* (i.e., "thought experiment" in German). Eventually, he developed his own theory of the nature of light and gravity.

## 5 | THOUGHT EXPERIMENTS: BEST-PRACTICE RECOMMENDATIONS

In this section, we provide recommendations on how to decide whether to conduct a thought experiment and which type, and how to plan, execute, report results, and discuss implications of a thought experiment. Following best practices in conducting a methodological

literature review, we created our recommendations as described by Aguinis et al. (2023). Specifically, the process involved a critical review of the 174 sources included in the Supporting Information. Our goal was not to conduct a systematic content or text analysis. Rather, we extracted recommendations from the 174 sources based on supporting evidence about their appropriateness across sources (i.e., triangulation of supporting evidence).

We organize our best-practice recommendations as follows. First, for anyone potentially interested in using this methodological approach, the initial step in the process involves deciding whether conducting a thought experiment may be helpful for a given research situation and goals. Second, assuming the first step results in a decision to conduct a thought experiment, we describe how to plan it (i.e., theory considerations). Third, we offer recommendations on how to execute a thought experiment (i.e., research design considerations). Fourth, we provide recommendations on how to report results. Finally, we offer suggestions on how to describe implications. As a preview of the material that follows, Table 1 includes a summary of our recommendations for each of these four stages.

## 5.1 | Deciding whether to conduct a thought experiment and which type

Although we believe that thought experiments have great potential for advancing theory in OB and other fields, we are certainly not

**TABLE 1** Summary of best-practice recommendations for deciding whether to conduct a thought experiment (and which type), planning and executing it, reporting results, and discussing implications

| Stage | Recommendation |
|---|---|
| 1. Deciding whether to conduct a thought experiment and which type | • Use the decision tree in Figure 2 to decide whether a thought experiment (TE) is useful and appropriate<br>• Determine the most appropriate type of TE based on the taxonomy summarized in Figure 1 |
| 2. Planning a thought experiment: Theory considerations | • Establish theoretical domain by reviewing key constructs, variables, and any relevant background information to create a mind model and set the scene for the experiment [1, 26, 33]<br>• Target a specific theoretical proposition or assumption to be challenged or affirmed [5, 19]<br>• Summarize theoretical assumptions to set the experimental boundaries of the TE [14]<br>• Illustrate the dimensions of the theory used in the mind model by providing diagrams, tables, and figures [1, 33] |
| 3. Executing a thought experiment: Research design considerations | • Be a raconteur to theorize based on abstraction (use a narrative storylike structure) [9, 10]<br>• Choose an existing TE [11, 33], alter a seminal TE to suit particular needs [8, 16, 18, 20, 22, 24, 31], or create a novel TE [7, 13]<br>• Use disciplined imagination (consistent internal and external logic) and specify why the proposed boundaries offered in the TE make sense [5, 14, 19, 27, 32, 33]<br>• Consider using multiple TEs to provide contrasting scenarios that test the theory [4, 16]<br>• Create *what if* counterfactual scenarios to tie abstract theory to understandable situations based on events from the present, distant past, or imagined in the distant future [17, 29, 30] |
| 4. Reporting results of a thought experiment | • Report the type of TE based on the taxonomy summarized in Figure 1<br>• Specify the usefulness of the thought experiment given the situational context of the theory [26, 33]<br>• Stipulate the meaning of the concepts used in the thought experiment, especially<br>• boundary conditions [2, 12]<br>• Include a figure or diagram to help illustrate the TE in the reader's mind [1, 3, 23]<br>• Describe what the expected outcome(s) should be based on the proposed boundaries of the TE [5, 19]<br>• Bolster the results of the TE by combining it with additional evidence, archival, and qualitative data [15] |
| 5. Discussing implications of a thought experiment | • Present the proposed shift in thinking or perspective based on the TE [19, 25]<br>• Discuss obscure truths or biases illuminated by the TE and offer alternative explanations [14, 16]<br>• Discuss interdisciplinary linkages and the value of implications to other disciplines [6, 21, 28] |

*Note*: Sources used to derive best-practice recommendations: [1]Botha (2019), [2]Bozeman and Feeney (2007), [3]Brown (2011), [4]Burge (1979), [5]Caste (1992), [6]Darnell (2001), [7]Elias and Gallagher (2014), [8]Emmerich and Gordjin (2018), [9]Fisher (2020), [10]Folger and Turillo (1999), [11]Harre and Wang (1999), [12]Hatherly et al. (2020), [13]Haukioja (2020), [14]Hong (2012), [15]Kadvany (2010), [16]Lachenicht (1993), [17]Leicester (2012), [18]Lucas (2003), [19]Mankiw (2013), [20]Maziarz (2017), [21]McDonald et al. (2017), [22]Nanay (2015), [23]Nothhaft and Stensson (2019), [24]Otero-Iglesias and Weissenegger (2020), [25]Raverty (2007), [26]Reichstein (2019), [27]Rinsley (1980), [28]Smith (2007), [29]Stögbauer and Komlos (2004), [30]Tateo and Valsiner (2015), [31]Vasileiou (2021), [32]Weick (1989), [33]Wempe (2008).

advocating their use in all situations and contexts. Accordingly, the first step involves deciding the appropriateness of conducting a thought experiment based on the particular research situation and goals. Figure 2 summarizes a decision tree to determine whether conducting a thought experiment is potentially useful for making theory advancements and, if so, what type of thought experiment is most appropriate based on the taxonomy we described earlier and summarized in Figure 1.

## 5.2 | Decision Point #1: Is there a need to confirm or disconfirm theory?

The first decision point involves clarifying whether there is a need to confirm or disconfirm theory (Wempe, 2008). In other words, is there a theoretical proposition or assumption to be confirmed or disconfirmed (Brun, 2017; Jalal & Ramachandran, 2017)? On the other hand, if the goal is primarily to conduct exploratory research, it is unlikely that using thought experiments will result in substantial theory advancements.

## 5.3 | Decision Point #2: Can an imagined scenario model the theory?

The second decision point asks whether the researcher can conceive a scenario that succinctly and effectively models the theory to be

tested. Scenarios should have a storylike narrative that brings the thought experiment to life in a manner that helps the reader follow the researcher's line of reasoning for their confirmation or disconfirmation assertion (Folger & Turillo, 1999). In this regard, the examples we described to illustrate each of the four types in our taxonomy are useful exemplars that can be used as templates for future thought experiments. For example, Einstein effectively modeled complex concepts of physics that non-experts are able to understand. Likewise, Newton modelled the highly abstract concept of gravity. The researcher, in other words, must be a *raconteur* to theorize and "bridge the gap between the abstract and the concrete" (Folger & Turillo, 1999, p. 742). Thought experiments can also arise from different situations, including real-life experiences and situations from the distant past or present (e.g., Fisher, 2020; Leicester, 2012; Stögbauer & Komlos, 2004) or completely new situations (Elias & Gallagher, 2014; Haukioja, 2020). In sum, the second decision point refers to whether and how the theory can be modeled in the thought experiment.

## 5.4 | Decision Point #3: Is the theory well developed?

The third decision point is about the theory's development stage. In other words, whether the theory is in the early (i.e., less developed) or late (i.e., more developed) stage. For example, returning to the exemplar thought experiments included in Figure 1, Newton's cannonball
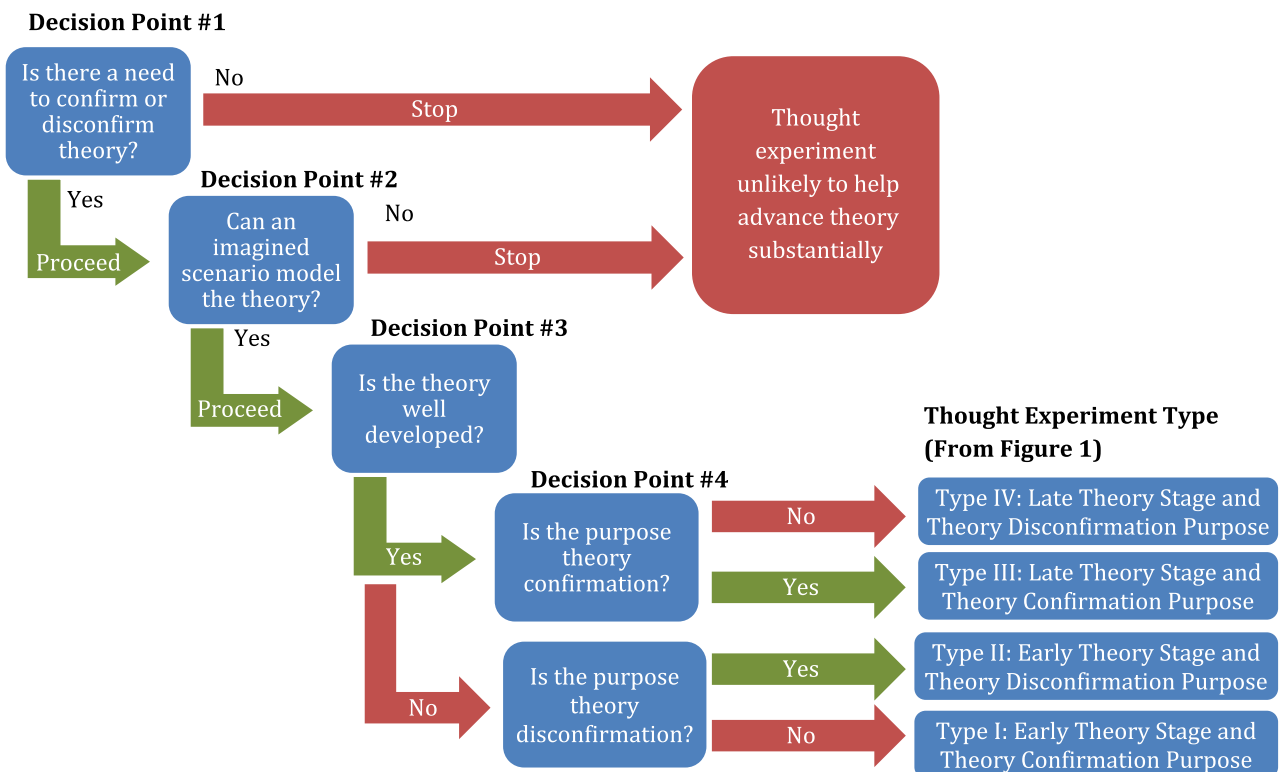


**FIGURE 2**  Deciding whether to implement a thought experiment and which type: decision tree

and Schrödinger's cat addressed early-stage theories, whereas Galileo and the Leaning Tower and Einstein riding a beam of light addressed late-stage theories.

## 5.5 | Decision Point #4: Is the purpose theory confirmation or disconfirmation?

The fourth decision point asks whether the goal is theory confirmation or disconfirmation. For example, DeMartino et al. (2016) confirmed theory in their thought experiment on fair trade, whereas Bozeman and Feeney (2007) disconfirmed their focal theory on mentoring. From our exemplar experiments, Galileo and the Leaning Tower exemplifies a scenario where the purpose is theory confirmation (of a well-developed theory), leading to a Type III experiment. Schrödinger's cat exemplifies a scenario where the purpose is theory disconfirmation (of an early-stage theory), leading to a Type II experiment. Thus, whereas Decision Point 1 should be answered with a simple "yes" or "no," Decision Point 4 should be answered explicitly as "theory confirmation" or "theory disconfirmation."

## 5.6 | Planning a thought experiment: Theory considerations

Once a decision has been made to conduct a thought experiment, the next step is planning it. Creating an engaging and convincing "mind model" for the experiment is crucial to facilitate a better understanding of complex abstractions (Botha, 2019). This is done by understanding and relaying the theory being tested, including its propositions, dimensions, and assumptions. Thus, planning a thought experiment requires identifying existing theory (either nascent or well developed) and deciding to test the theory by confirming or disconfirming it. This can be achieved by specifying a point of dissension or agreement (e.g., a debate in the literature or a commonly misunderstood relationship between similar concepts) that justifies the need for a thought experiment. Whereas early-stage theoretical testing relies on sparse or disparate studies, later stage theoretical testing relies on more abundant literature and better-developed and established ideas.

First, "setting the scene" (Reichstein, 2019, p. 57) requires conjuring a "mind model" constructed by reviewing key constructs, variables, and any relevant background information (Botha, 2019; Wempe, 2008). Of note, given sufficient knowledge of a theory, researchers can begin mentally experimenting without having to conduct additional background research. For example, Einstein and Newton did not explicitly set out to do a thought experiment; however, they had sufficient knowledge to begin experimenting within their minds. Einstein performed many *Gedankenexperimente* in his idle time, which led to the development of his thought experiment involving riding a beam of light.

However, when performing a *Gedankenexperiment*, Einstein still explicitly made the theory tested explicit, as was the case with his beam of light experiment addressing how the emission theory of light

was unfeasible. Caste (1992) exemplified this best practice in his own thought experiment. Specifically, Caste clearly communicated to the reader that he explicitly sought to disconfirm the *employee productivity proposition* as a valid justification for mandatory drug testing. Caste noted that the employee productivity proposition "states that since the employer has purchased the employee's time, the employer has a proprietary right to ensure that the time purchased is used as efficiently and productively as possible" (p. 301). He conducted a thought experiment to demonstrate how the belief in this proposition as a valid justification for mandatory drug testing also commits one logically to morally repugnant consequences (e.g., employers having a proprietary right to mandate the use of a pain-inducing drug to increase employee efficiency).

The theoretical assumptions that form the boundaries of the experiment should be explicit. For example, Hong (2012) used an organized table to quickly summarize the assumptions of the key concepts relating to the experiment. To facilitate both creating an effective mental model and addressing assumptions, diagrams or figures can also help set the scene visually. As an illustration, Botha (2019) used several diagrams to explain the mind model of the experiment. Alternatively, Wempe (2008) used two tables to ground the reader in the assumptions, logic, and domain characteristics of the thought experiment as it relates to different theories under comparison.

## 5.7 | Executing a thought experiment: research design considerations

The execution stage involves conducting the thought experiment in the laboratory of the mind with a well-defined model developed in the planning stage. Foremost, thought experiments should seek to test theory. Folger and Turillo (1999) recommended that researchers be *raconteurs*, which involves "thinking in narrative, storylike terms provides a route to theorizing via abstraction" (p. 754). This recommendation is highlighted by Jabri and Pounder's (2001) argument that "narratives express the richness and diversity of human experiences and thus challenge simplistic analyses of managerial issues ... Narrative is the expression of actual human experience" (p. 682).

The narrative can take shape either by using an existing thought experiment (Harre & Wang, 1999; Lucas, 2003; Wempe, 2008), altering one to suit the particular needs of a theory (Emmerich & Gordjin, 2018; Lucas, 2003; Otero-Iglesias & Weissenegger, 2020), or creating an entirely novel one (Elias & Gallagher, 2014). Researchers should also consider adopting thought experiments from completely different fields (Nanay, 2015; Vasileiou, 2021), using multiple thought experiments (Lachenicht, 1993), basing a thought experiment on a real-life experience or situation (Fisher, 2020), and using metaphors (Wempe, 2008). For example, McMullen (2018) examined entrepreneurial ecosystems as organizational hybrids akin to biological hybrids.

While thought experiments offer a great deal of freedom compared with other types of experiments, researchers should use "disciplined imagination." As explained by Weick (1989), "'discipline' in theorizing comes from consistent application of selection criteria to

trial-and-error thinking, and the 'imagination' in theorizing comes from deliberate diversity introduced into the problem statements, thought trials, and selection criteria that comprise that thinking" (p. 516). As such, disciplined imagination means that an underlying internal and external logic should make practical sense and can be conveyed to a reader (Caste, 1992; Hong, 2012; Mankiw, 2013; Wempe, 2008). Therefore, even the most fantastical thought experiments (e.g., Einstein riding a beam of light) are explained and grounded in natural and explainable phenomena that effectively convey abstract ideas.

To give a thought experiment more breadth, researchers may also decide to include multiple thought experiments (Burge, 1979; Lachenicht, 1993) or create new *what if* counterfactual scenarios to narratively apply and explain abstract theoretical mechanisms (Leicester, 2012; Stögbauer & Komlos, 2004). For example, when considering behavioral solutions to the AIDS crisis, Lachenicht (1993) conducted multiple thought experiments, with slight variations, to contrast ideas. Another useful approach for including several thought experiments is to consider adding *what-if* counterfactual scenarios to ground abstract theory narratively. In an effort to demonstrate the effects of fiscal policy, Stögbauer and Komlos (2004) used the Nazi seizure of power to create a counterfactual thought experiment.

## 5.8 | Reporting results of a thought experiment

To report results of thought experiments, researchers must guide readers through the experiment and justify the choices made. First, the type of thought experiment used, as summarized in the taxonomy in Figure 1, should be detailed to orient the reader as to the nature of the thought experiment (early vs. late stage, confirmation vs. disconfirmation).

Second, it is essential to justify why the method was used, given the situational context of the theory. For example, Reichstein (2019), in examining whether prisoners have the right to die, explained how: "Few European jurisdictions are close to implementing a 'right to die' for citizens, and they are even further away from a 'right to die for prisoners'. But that does not mean we should not engage with the idea. When we engage in utopian thinking, considering what our society could and should look like, we can come across and highlight shortcomings and problems for the status quo" (p. 56). Given the situational context, utopian thought was used and justified as there were no alternatives to consider that particular situation.

Readers should be given a clear description of all key concepts given the abstract nature of thought experiments, particularly for boundary conditions (Botha, 2019; Caste, 1992; Hong, 2012; Nothhaft & Stensson, 2019). As an illustration, Bozeman and Feeney (2007) explained that "often the concepts presented are suggestive, identifying the attributes of mentoring rather than stipulating the meaning of the concept itself and, in particular, its boundary conditions" (p. 721). They went on to cite several researchers who failed to define their focal construct. Hatherly et al. (2020) defined the central concept of the thought experiment to ensure clarity in how they defined stakeholders.

To more fully immerse the reader in the experiment and help detail what occurred in the imagination of the researcher, researchers should also consider including a figure or diagram to help ground the thought experiment and bring it to life (Botha, 2019; Brown, 2011; Nothhaft & Stensson, 2019). Figures can help facilitate understanding as well as provide evidence for the explanations for each of the questions raised in the imagined scenario (Botha, 2019; Brown, 2011). In other words, given the proposed specifications of the thought experiment, it should explain what the outcome is given proposed boundaries (Caste, 1992; Mankiw, 2013).

Finally, while thought experiments do not create empirical data, data can be used to inform thought experiments. Stated differently, results can be supplemented and enhanced by connecting the thought experiment to historical evidence or archival data (Kadvany, 2010) or including qualitative data (Van Bockhaven & Matthyssens, 2017). To this effect, part of the results reporting process includes incorporating lived experiences to enhance realism and fully immerse readers into the narrative scenario (Fisher, 2020).

## 5.9 | Discussing implications of a thought experiment

Discussing implications begins with presenting the proposed shift in thinking or perspective resulting from the thought experiment. For example, Mankiw (2013) explained the difference between the original position before the thought experiment and, if the policy in the thought experiment was enacted, potential outcomes. Raverty (2007) used the results of the thought experiment to consider how the focal construct, gender, affects an overall monastic framework. In addition, the results of the thought experiment can be integrated or reconciled with other theories to create a new theory or framework (Liu, 2017; Sarmiento & Shukla, 2011; Wolfe, 2000).

While many groundbreaking and paradigm-shifting results have originated from thought experiments, "Thought experiments do not always reveal deep truths ... there is always the possibility that the thought experiment might reveal a truth which would otherwise have remained obscure" (Lachenicht, 1993, p. 15). For example, Hong (2012) used a thought experiment, not to create a new grand theory on international trade, but simply to raise key questions of how processes are adapted to foreign environments.

Finally, discussing implications also includes a consideration of interdisciplinary linkages. As an example, Darnell (2001) discussed how Sapir's work made connections between psychiatry and cultural anthropology and how, metaphorically, the two disciplines represented the two sides of a coin: "Cultural anthropologists, psychiatrists, sociologists, and so on should pool the insights deriving from their particular standpoints so as to arrive at a full picture of the whole of human existence. Multiplicity of standpoints could compensate to some extent for the impossibility of objectivity by any human observer" (p. 19). As another illustration, Smith (2007) applied the results of the shift in perception of the thought experiment to the "impact on the practical dimensions of the work of a manager" (p. 478).

Next, to show and not just tell that our recommendations are actionable, practically doable, and useful, and that thought experiments have great potential for making theory advancements, we describe an original thought experiment in the area of allyship. In doing so, we will implement our own recommendations in the same sequence in which they are summarized in Table 1.

# 6 | ALLYSHIP MY WAY OR YOUR WAY: A THOUGHT EXPERIMENT

Allyship refers to members of advantaged groups engaging in committed action to improve the treatment and status of members of a disadvantaged group (Droogendyk et al., 2016). Ostrove and Brown (2018) summarized this stream of research as follows: "Although dominant group allies have been increasingly studied by social psychologists interested in positive intergroup relations and the promotion of social justice, most of the existing research focuses on self-identified allies or dominant group individuals who are engaging in social justice activities" (p. 195).

Consider the following theoretical propositions based on extant theory: dominant group allies should (a) exhibit characteristics of both affirmation and informed action and (b) demonstrate "supportive contact", or demonstrate care toward the beneficiary and oppose workplace inequality (Droogendyk et al., 2016, p. 318; Ostrove & Brown, 2018, p. 201). Accordingly, some researchers have placed the value of allyship on advantaged group members who are interested in positive intergroup relationships and social justice (Goodman, 2011). For example, researchers have explained the roles and responsibilities of allies (Spanierman & Smith, 2017) and identified the challenges of being a White ally (Sue, 2017). Nonetheless, this is a point of theoretical tension because there is a different perspective that the value of allyship is determined by the intended beneficiaries (Ashforth et al., 2016; Erskine & Bilimoria, 2019). We therefore conducted a thought experiment to clarify *who* should determine the value of allyship: A well-intentioned ally or the intended beneficiary. Next, we describe how we implemented our recommendations in each stage of our thought experiment.

## 6.1 | Deciding whether to conduct a thought experiment and which type

The first step in the process involves deciding whether a thought experiment may be an appropriate and useful methodological approach for making a substantial theory advancement and, if yes, which type would be best. To make this determination, we followed the decision-making tree in Figure 2.

## 6.2 | Decision Point #1: Is there a need to confirm or disconfirm theory?

In our illustrative experiment, there is a need to confirm or disconfirm the theory regarding the understanding of allyship. Our rationale is based on the theoretical tension regarding who determines the value of allyship. Confirming or disconfirming this aspect of allyship research via a thought experiment should also provide insights regarding the difference between effective allyship and performative allyship. Performative allyship is "where well-meaning people with power and privilege show interest in becoming an ally but do not engage in the ongoing emotional labor, self-reflection, continuous education, courage, commitment, and exchange of power inherent in true allyship" (Erskine & Bilimoria, 2019, p. 329).

## 6.3 | Decision Point #2: Can an imagined scenario model the theory?

A storylike narrative scenario can be imagined wherein the theory is tested efficiently and effectively. As such, our answer is "yes." Specifically, a situation wherein allyship theory is tested in a relatable and understandable situation. For example, imagine an allyship situation in academia where a minority female faces a situation placing her majority male supervisor in an allyship role.

## 6.4 | Decision Point #3: Is the theory well developed?

There is a dearth of literature on workplace allyship in OB. However, it has much deeper roots in other fields in which it has studied extensively (e.g., social sciences: Case, 2012; Droogendyk et al., 2016; Ostrove & Brown, 2018). Thus, while the theory is being applied in a relatively new OB context, the theoretical propositions are considered from later stage theory. Consequently, our answer is "yes," theory in this domain is well-developed.

## 6.5 | Decision Point #4: Is the purpose theory confirmation or disconfirmation?

The purpose of our thought experiment is to confirm theory. Specifically, we seek to confirm the allyship proposition that argues that the value of allyship is determined by the intended beneficiary (Erskine & Bilimoria, 2019). Based on the decision tree summarized in Figure 1, a Type III thought experiment is the most appropriate.

Next, we implemented recommendations about planning, executing, reporting, and discussing implications as summarized in Table 1.

## 6.6 | Planning a thought experiment: Theory considerations

We sought to confirm extant allyship theory to support key characteristics and assumptions in our narrative. Thus, our research question inquires who should determine the value of allyship. Given the inherent racial and gender implications, along with power dynamics in allyship research, we contextualize our theory in an imagined scenario including how White male allies should respond to Black, Indigenous, and People of Color (BIPOC) women colleagues' requests for support in the workplace. We also inquire about the consequences for dismissing specific requests for support.

We also propose that the practice of allyship in organizations emphasizes acknowledging the value of self-determination wherein advantaged others may act to remedy organizational harms that violate the self-determined action of disadvantaged persons. Specifically, in cases in which White male allies can intervene to remedy harmful or marginalizing organizational action (Warren et al., 2020), we explore what is considered effective allyship and propose that effective allyship is supportive action (in this case, by White male allies) that affirms the self-determination (Gagné & Deci, 2005) of the beneficiary. In sum, our thought experiment involved implementing recommendations for the planning stage summarized in Table 1 such as establishing the theoretical domain by reviewing key constructs, variables, and relevant information to create a mind model and set the scene for the experiment, and targeting specific theoretical propositions.

## 6.7 | Executing a thought experiment: Research design considerations

Our goal is to confirm theoretical allyship propositions (Ostrove & Brown, 2018) in cases in which White male allies intervene to remedy harmful or marginalizing organizational action. Our thought experiment is parsimonious and allows for the application of domain-specific theories. By following recommendations for the execution stage summarized in Table 1, we created the following:

> Dr. Beauford Manigold, Chair of the Finance Department at a top-tier private research university, has been asked to meet with Dr. Dianne Winfrey, a junior faculty member recruited three years ago as a part of a university-wide faculty diversity and inclusion initiative. Dr. Manigold is a White male recruited to the Department 10 years ago from a very prestigious university and has quickly risen through the ranks into informal and formal leadership in his department, university, and field. During the initiative, Dr. Winfrey, a young African American woman, was the only junior faculty member recruited to the university's School of Business. Dr. Winfrey is also the first African American/Latina woman to be recruited into a tenure track-position in the department. She recently had her third-year review and received an additional year of reprieve to improve her dossier after appealing (with the Dean) a recommendation from the review committee not to renew her contract. Specifically, the committee cited her lack of progress on reputable publications in the field and not being a visible part of the department's academic community. Dr. Manigold feels compelled to support Dr. Winfrey, as he played a significant role in recruiting her. He invites Dr. Winfrey to an informal meeting to discuss her email formally responding to the department's review process and her appeal. In her email, she lists three concerns that she hopes to discuss during their meeting. Her concerns include: (1) the lack of community in the department and her inability to connect with colleagues, (2) although she has not met "expected" benchmarks regarding publications, her scholarship is mainly qualitative, which she argues requires more time in the research and review process, and (3) she has been overwhelmed with the service requests for diversity, equity, and inclusion (DEI) initiatives and feels pressure from colleagues and "higher-ups" to serve on time-consuming committees. Paradoxically, Dr. Winfrey's request from Dr. Manigold, the Department Chair, is to lobby on her behalf within the department and externally to renegotiate her benchmark requirements for tenure and dissuade others from requesting her participation in service work related to DEI efforts. However, Dr. Manigold believes that asking for either on her behalf would compromise her standing in the university and shed a negative light on the department that wants to uphold rigorous requirements for faculty tenure. For the paradox to exist, several assumptions are rejected: Dr. Manigold has experience supporting BIPOC women colleagues in this context, Dr. Winfrey has already acted on her own behalf, Dr. Winfrey knows which colleagues in the department voted against her tenure, and finally that Dr. Winfrey has substantial knowledge of the organizational systems of the university. Explicitly stating rejected assumptions allows the thought experimenter to conduct the exercise within certain boundary conditions.

Again, implementing our recommendations, we created Figure 3 to show visually and more clearly the options available to Dr. Manigold as a dominant group member. Those options include using his own judgment, consulting and collaborating with Dr. Winfrey on an alternative solution, or enacting Dr. Winfrey's requests as she desires.
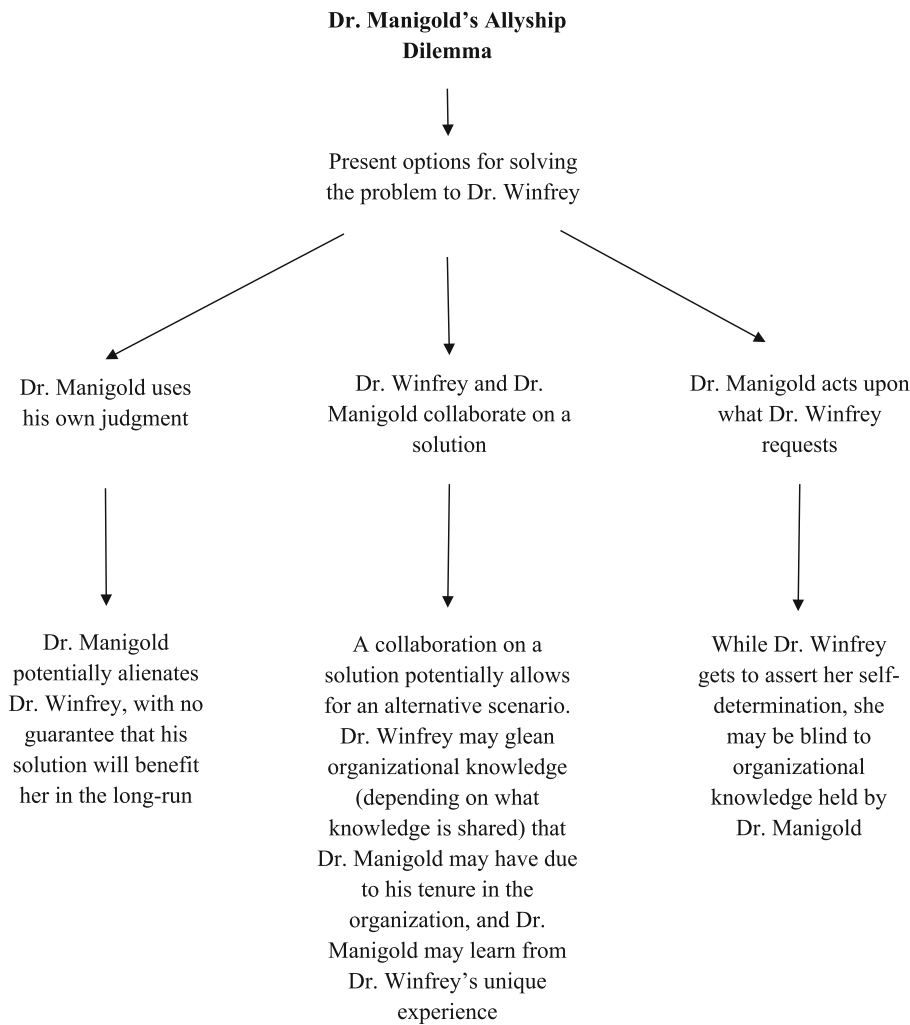
**Dr. Manigold's Allyship Dilemma**

**FIGURE 3** Visual representation of allyship thought experiment scenario options

Present options for solving the problem to Dr. Winfrey

Dr. Manigold uses his own judgment

Dr. Winfrey and Dr. Manigold collaborate on a solution

Dr. Manigold acts upon what Dr. Winfrey requests

Dr. Manigold potentially alienates Dr. Winfrey, with no guarantee that his solution will benefit her in the long-run

A collaboration on a solution potentially allows for an alternative scenario. Dr. Winfrey may glean organizational knowledge (depending on what knowledge is shared) that Dr. Manigold may have due to his tenure in the organization, and Dr. Manigold may learn from Dr. Winfrey's unique experience

While Dr. Winfrey gets to assert her self-determination, she may be blind to organizational knowledge held by Dr. Manigold

## 6.8 | Reporting results of a thought experiment

Figure 3 also relies on relevant theoretical vantages of our thought experiment, namely, self-determination theory (Deci & Ryan, 1985; Gagné & Deci, 2005) and mentoring theory (Ragins & Kram, 2007). Subsequently, these two theoretical perspectives guide how we report the outcomes of Dr. Manigold's dilemma of being an ally to Dr. Winfrey.

Ostrove and Brown (2018) proposed that dominant group allies should affirm the identity of the beneficiary, consider the beneficiary's expressed needs, and demonstrate supportive action. Hence, we advance that the best course of action for Dr. Manigold is to first acknowledge Dr. Winfrey's marginalized identity in her profession and learn from her experiences. In addition, we determined that Dr. Manigold must share his institutional knowledge (i.e., mentor Dr. Winfrey, see Blake-Beard et al., 2006; Ragins & Kram, 2007), collaborate on proposed solutions, and advocate on her behalf within the Department and University.

Alternatively, suppose that Dr. Manigold does not lobby in the manner as Dr. Winfrey wished to protect Dr. Winfrey from potential harmful consequences. In this scenario, Dr. Manigold seems to have legitimate concerns about unintended backfiring effects. However,

the argument can be made that Dr. Winfrey would not be considered an effective ally to Dr. Manigold, as his intended action may not fully consider what Dr. Winfrey believes will benefit her most. Thus, Dr. Manigold (by not fully considering Dr. Winfrey's requests) may violate her self-determination and thwart her psychological need for intrinsic motivation and consequent internalization of strategies that may lead to her benefit (Gagné & Deci, 2005).

It is also important to acknowledge that in some respects, Dr. Manigold potentially has more knowledge of organizational factors given his tenure that may ultimately prevent Dr. Winfrey from succeeding if she approaches her concerns in the fashion she proposes. The opportunity to proactively mentor Dr. Winfrey allows Dr. Manigold to engender a sustaining positive relationship and enhance knowledge and skills that Dr. Winfrey may enact in future situations (Blake-Beard et al., 2006; Blake-Beard et al., 2017; Ragins & Kram, 2007). Given these factors, Dr. Manigold's approach may also include an offer to share his institutional knowledge of the systems within which they are operating. Depending on the knowledge shared, Dr. Winfrey may feel supported and empowered in a manner that is consistent with Dr. Winfrey's opportunity to enact self-determination, or Dr. Winfrey may feel as if she is being asked to assimilate and appease others due to "the system."

## 6.9 | Discussing implications of a thought experiment

Our thought experiment confirms key theoretical propositions and advances theory by invoking self-determination and mentoring theories to ascertain Dr. Manigold's best course of action. To this end, our thought experiment also resolved a point of theoretical tension in the allyship literature and demonstrated that *it should be the intended beneficiaries who determine the value of allyship*. In doing so, the difference between effective allyship and performative allyship becomes clearer as a result of our thought experiment. As noted by Erskine and Bilimoria (2019), "performative allies are driven by the need for validation and may intellectually understand the issues at hand, yet not sacrifice their personal or professional capital to challenge or transform systems that they benefit from, even unwittingly" (p. 329). Dr. Manigold may understand Dr. Winfrey's situation yet be slow to fully challenge the current system (i.e., pre-existing institutional criteria for tenure). Thus, for Dr. Manigold to be an effective ally (not a performative ally), he must challenge the current systems to advocate for Dr. Winfrey.

Our thought experiment has implications for future research as well. As opposed to predominantly focusing on the role of self-identified allies from dominant groups, researchers can shift their focus to the role of intended beneficiaries. This shift in perspective should reveal insights regarding how effective allyship looks like to women, BIPOC, LGBTQIA+ employees, or members of any other marginalized group. Thus, future research could also examine allyship in a demographic-specific context as opposed to combining all "minoritized groups." This approach would increase our understanding of the differences and similarities of various intended beneficiaries of allyship.

Furthermore, of critical significance is the exchange of organizational and institutional knowledge garnered via different experiential vantages. Our own thought experiment illustrates how we were able to uncover obscure truths. In the context of allyship, there are allies and intended beneficiaries. Thus, it is an exchange relationship and hopefully one that is mutually beneficial. In our thought experiment, there is an opportunity for Dr. Winfrey and Dr. Manigold to benefit from sharing knowledge about their different vantage points. While it is critical to uphold Dr. Winfrey's self-determination, it should be noted that her self-determination may be amplified if she has more organizational knowledge. This type of collaboration can result in effective problem-solving, which has been demonstrated in other research domains (Guest & King, 2004; Pfeffer, 1981).

Next, we discuss the applicability of thought experiments as an untapped methodological approach in several OB areas. As illustrations, we focus on just six.

## 7 | ILLUSTRATIONS OF OB THEORIES AND DOMAINS THAT CAN BENEFIT FROM THOUGHT EXPERIMENTS

A 2008 special issue of *Journal of Organizational Behavior* included articles offering recommendations on what we should do, as a field, to move forward. For example, Edwards (2008) drew particular attention to the need to overcome methodological barriers to advance theory, particularly the overreliance and overemphasis on the methods researchers learn at the beginning of their careers. To remedy this challenge, Edwards recommended a shift in attitudes to overcome methodological barriers by adopting new methods that "see new ways to answer theoretical questions and perhaps identify questions that would not have occurred to us otherwise" (p. 484). Relatedly, Lefkowitz (2008) advocated for proactive methodological changes to overly "scientistic" (p. 443) analysis due to eschewing humanistic values as they are viewed as antithetical to objective science. We believe that thought experiments are particularly suited to answer both of these calls. Many OB domains could benefit from thought experiments given their need to either confirm or disconfirm theories. We describe just six and Table 2 offers a summary of the material that follows.

First, diversity resides within a sizeable nomological network of theories, including social psychology, social identity, and categorization theories (Roberson et al., 2017). Diversity and inclusion has been studied in terms of sex and race composition (Blau, 1977; Kanter, 1977), diversity climate (Mor Barak, 2022), effects on firm performance (Sacco & Schmitt, 2005), and the importance of the diversity context (Joshi & Roh, 2009). Moving forward, Roberson et al. (2017) offered several recommendations to advance diversity theory, including employing diverse methodologies, broadening diversity contexts, and integrating different theoretical perspectives.

**TABLE 2** Selected organizational behavior theories and domains that would benefit from thought experiments

| Theories and domains | Illustrative research questions to be answered using thought experiments |
|---|---|
| Diversity and inclusion | Given a large nomological network and myriad theories from multiple fields, how can employing new methods, broadening the construct, and integrating different theoretical perspectives spur new theoretical insight? |
| Leadership | How can the dense leadership construct landscape be better understood with fewer related constructs? Is it possible to develop a more parsimonious full-range model of leadership? |
| Performance | How can objective and subjective measures of performance be combined to form a more accurate picture of the performance construct? |
| Selection and recruitment | How can context-sensitive results be applied more broadly, and if so, what variables are responsible? Can complicated individual-difference-outcome relations be modeled more parsimoniously? |
| Teams | How can the dynamic and emergent nature of teams be modeled while considering the multi-level units of analysis? |
| Turnover | Can post-turnover implications for both organizations *and* employees be conceptualized? How can turnover studies be expanded to better capture contextual issues? |

Thought experiments are not only well-suited for addressing each of these recommendations individually but also uniquely suited to address them simultaneously.

Leadership theories can directly and quickly benefit from thought experiments to address the issue of construct proliferation (Shaffer et al., 2016). Specifically, despite the myriad leadership constructs, scholars can offer little practical advice to practitioners regarding their use and implementation. The most popular constructs (charismatic-transformational and leader-member exchange, or LMX) suffer from serious deficiencies (e.g., Gottfredson & Aguinis, 2017). Both constructs offer similar advice; good leadership is associated with good outcomes (and vice-versa). Thus the dominant paradigms of charismatic-transformational and LMX are less than ideal. Some argue that scrapping them altogether may be a good idea (Van Knippenberg & Sitkin, 2013). Others advocate fixing them (Antonakis et al., 2016), recommend moving closer back to "square one" (Gottfredson et al., 2020; Judge et al., 2004), or call for creating a new full-range theory of leadership (Anderson & Sun, 2017). Thought experiments offer a method of quickly, easily, and cheaply addressing these thorny issues—at least in part.

Performance is a critical construct in organizational behavior, human resources, and strategy and, given its prominence, can be bolstered using thought experiments. Namely, numerous theories conceptualize performance as an individual, firm, or multilevel construct (DeNisi & Murphy, 2017; Koopmans et al., 2011; Viswesvaran & Ones, 2000; Welbourne et al., 1998). Despite the importance of the construct and the various conceptualizations, there has yet to be a simple but not simplistic, parsimonious framework for understanding performance. Each of the conceptualizations focuses on different levels of performance and leads to paradoxical results depending on a researcher's preferred framework. Thought experiments offer opportunities to look at performance from different levels of abstraction to gain insight into a potentially unifying theoretical framework.

Thought experiments offer a unique method to consider how a large scope of multiple factors interact simultaneously, particularly how business, legal, and societal challenges impact the "supreme problem" of selection and recruitment (Hall, 1917; Ployhart et al., 2017). Individual difference-outcome relationships remain a concern for OB and human resource management researchers (Aguinis et al., 2005; Le et al., 2011). Additionally, the context in which personnel research is carried out is particularly sensitive, and results are difficult to generalize (Ployhart et al., 2017). Thought experiments offer a potential avenue to address an unlimited number of variables and, more importantly, change variables and contexts almost instantly within the laboratory of the mind. Thus, they offer a practical way to consider multiple factors under different circumstances to provide a barometer for which directions might prove fruitful for study.

Interest in teams has shifted from strictly focusing on individuals within teams to focus on the team itself and networks of teams (Mathieu et al., 2017). Team study emphasizes developing proper indices for measuring team effectiveness (Ilgen, 1999) at different levels of analysis, such as the team level (Hackman & Morris, 1975) and the individual level (O'Reilly, 1977). Teams research is thus a complex domain involving organizational structure and culture, leadership, and multi-team systems, and future directions point to the importance of considering its dynamic temporal nature (Cronin et al., 2011). However, acquiring large-scale, multilevel, and longitudinal data on teams takes considerable time and resources (Bradley & Aguinis, 2022). Thought experiments offer a relatively faster method to consider these types of questions. A thought experiment can also act as a starting point to decide the research direction before proceeding with a large-scale study.

Finally, turnover has been one of the most dominant and persistently studied areas in OB; however, there has been a significant change in how it has been studied since the turn of the century (Hom et al., 2017). Notably, more attention and focus have been given to staying rather than leaving (Mitchell et al., 2001). Hom et al. (2017) suggested that, given this new direction in research, there are several ways to advance turnover scholarship by theorizing about the change in turnover antecedents and consequences, investigating post-turnover implications for both employees and organizations, and expanding turnover studies to better capture context. Given the difficulties in studying individuals who have left the organization, thought experiments can be used to help theorize about why individuals leave by addressing these post-turnover concerns for both organizations and individuals, as well as considering those same concerns in different contexts.

# 8 | CONCLUDING REMARKS

Thought experiments have been responsible for significant breakthroughs in the history of science. However, despite the instrumental role they have played in improving our understanding of numerous phenomena, they have been used very rarely in OB and related fields including human resource management, industrial and organizational psychology, entrepreneurship, and strategy. As described by the many examples throughout our article, as well as our own substantive thought experiment in the domain of allyship, this underutilized methodological approach has many advantages and can result in meaningful theory advancements. Clearly, thought experiments are not a methodological silver bullet that will answer all questions and resolve all theoretical dilemmas. However, given their cost effectiveness and untapped benefits, OB researchers have little, if anything, to lose by conducting a thought experiment. We hope our best-practice recommendations will be useful and inspire future research using thought experiments in OB and related fields.

## ENDNOTE

[1] Studies preceded by an asterisk were used to generate recommendations in Table 1.

## DATA AVAILABILITY STATEMENT

The list of articles upon which we based our review is included in the online supplement.

## ORCID

*Herman Aguinis* https://orcid.org/0000-0002-3485-9484
*Jose R. Beltran* https://orcid.org/0000-0002-3886-8142
*Estelle E. Archibold* https://orcid.org/0000-0001-6528-6722
*Esther L. Jean* https://orcid.org/0000-0003-1860-0365
*Darryl B. Rice* https://orcid.org/0000-0002-5601-5821

## REFERENCES[1]

Aguinis, H., Beaty, J. C., Boik, R. J., & Pierce, C. A. (2005). Effect size and power in assessing moderating effects of categorical variables using multiple regression: A 30-year review. *Journal of Applied Psychology*, *90*(1), 94–107. https://doi.org/10.1037/0021-9010.90.1.94

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., Pierce, C. A., Bosco, F. A., & Muslin, I. S. (2009). First decade of organizational research methods: Trends in design, measurement, and data-analysis topics. *Organizational Research Methods*, *12*(1), 69–112. https://doi.org/10.1177/1094428108322641

Aguinis, H., Ramani, R. S., & Alabduljader, N. (2023). Best-practice recommendations for producers, evaluators, and users of methodological literature reviews. *Organizational Research Methods*, *26*(1), 46–76. https://doi.org/10.1177/1094428120943281

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

Anderson, M. H., & Sun, P. Y. (2017). Reviewing leadership styles: Overlaps and the need for a new 'full-range' theory. *International Journal of Management Reviews*, *19*(1), 76–96. https://doi.org/10.1111/ijmr.12082

Antonakis, J., Bastardoz, N., Jacquart, P., & Shamir, B. (2016). Charisma: An ill-defined and ill-measured gift. *Annual Review of Organizational Psychology and Organizational Behavior*, *3*, 293–319. https://doi.org/10.1146/annurev-orgpsych-041015-062305

Ashforth, B. E., Schinoff, B. S., & Rogers, K. M. (2016). "I identify with her," "I identify with him": Unpacking the dynamics of personal identification in organizations. *The Academy of Management Review*, *41*(1), 28–60. https://doi.org/10.5465/amr.2014.0033

Bankins, S., & Formosa, P. (2020). When AI meets PC: Exploring the implications of workplace social robots and a human-robot psychological contract. *European Journal of Work and Organizational Psychology*, *29*(2), 215–229. https://doi.org/10.1080/1359432X.2019.1620328

Bankins, S., & Waterhouse, J. (2019). Organizational identity, image, and reputation: Examining the influence on perceptions of employer attractiveness in public sector organizations. *International Journal of Public Administration*, *42*(3), 218–229. https://doi.org/10.1080/01900692.2018.1423572

Barnes, C. M., Dang, C. T., Leavitt, K., Guarana, C. L., & Uhlmann, E. L. (2018). Archival data in micro-organizational research: A toolkit for moving to a broader set of topics. *Journal of Management*, *44*(4), 1453–1478. https://doi.org/10.1177/0149206315604188

Blake-Beard, S., Halem, J., Archibold, E. E., Boncoeur, D. O., Burton, A. R., & Kumar, P. (2017). Mentoring relationships of professional Indian women: Extending the borders of our understanding at the intersection of gender and culture. In A. J. Murrell & S. Blake-Beard (Eds.), *Mentoring diverse leaders* (pp. 25–43). Routledge. https://doi.org/10.4324/9781315747569

Blake-Beard, S., Murrell, A. J., & Thomas, D. A. (2006). *Unfinished business: The impact of race on understanding mentoring relationships*. Division of Research, Harvard Business School.

Blau, P. M. (1977). *Inequality and heterogeneity*. Free Press.

Boser, S. (2007). Power, ethics, and the IRB: Dissonance over human participant review of participatory research. *Qualitative Inquiry*, *13*(8), 1060–1074. https://doi.org/10.1177/1077800407308220

*Botha, A. P. (2019). A mind model for intelligent machine innovation using future thinking principles. *Journal of Manufacturing Technology Management*, *30*(8), 1250–1264.

*Bozeman, B., & Feeney, M. K. (2007). Toward a useful theory of mentoring: A conceptual analysis and critique. *Administration and Society*, *39*(6), 719–739. https://doi.org/10.1177/0095399707304119

Bradley, K. J., & Aguinis, H. (2022). Team performance: Nature and antecedents of non-normal distributions. *Organization Science*. https://doi.org/10.1287/orsc.2022.1619

*Brown, J. R. (1986). Thought experiments since the scientific revolution. *International Studies in the Philosophy of Science*, *1*(1), 1–15. https://doi.org/10.1080/02698598608573279

Brown, J. R. (2011). *The laboratory of the mind*. Routledge. https://doi.org/10.4324/9780203979150

Brun, G. (2017). Thought experiments in ethics. In T. S. Michael, F. Yiftach, & J. R. Brown (Eds.), *The Routledge companion to thought experiments* (pp. 195–210). Routledge.

*Burge, T. (1979). Individualism and the mental. *Midwest Studies in Philosophy*, *4*, 73–121. https://doi.org/10.1111/j.1475-4975.1979.tb00374.x

Burton, A., Altman, D. G., Royston, P., & Holder, R. L. (2006). The design of simulation studies in medical statistics. *Statistics in Medicine*, *25*(24), 4279–4292. https://doi.org/10.1002/sim.2673

Carr, A., & Zanetti, L. A. (1999). Metatheorizing the dialectic of self and other: The psychodynamics in work organizations. *American Behavioral Scientist*, *43*(2), 324–345. https://doi.org/10.1177/00027649921955281

Case, K. A. (2012). Discovering the privilege of whiteness: White women's reflections on anti-racist identity and ally behavior. *Journal of Social Issues*, *68*(1), 78–96.

*Caste, N. J. (1992). Drug testing and productivity. *Journal of Business Ethics*, *11*(4), 301–306.

Christian, R. C. (1963). Marketing experimentation: Forgotten tool. *Journal of Marketing*, *27*(4), 75–76. https://doi.org/10.2307/1248652

Colquitt, J. A., & Zapata-Phelan, C. P. (2007). Trends in theory building and theory testing: A five-decade study of the Academy of Management Journal. *Academy of Management Journal*, *50*(6), 1281–1303. https://doi.org/10.5465/AMJ.2007.28165855

Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin. https://doi.org/10.1002/pam.10129

Cornelissen, J., & Durand, R. (2012). More than just novelty: Conceptual blending and causality. *Academy of Management Review*, *37*(1), 152–154. https://doi.org/10.5465/amr.2011.0129

Cronin, M. A., Weingart, L. R., & Todorova, G. (2011). Dynamics in groups: Are we there yet? *The Academy of Management Annals*, *5*(1), 571–612. https://doi.org/10.1080/19416520.2011.590297

*Darnell, R. (2001). Commentary on why cultural anthropology needs the psychiatrist: Edward Sapir's thought experiment in the interdisciplines of cultural anthropology and psychiatry. *Psychiatry*, *64*(1), 17–22. https://doi.org/10.1521/psyc.64.1.17.18234

Davies, S. P. (1989). The validity of Searle's gedankenexperiment: Rule-based AI and connectionism. *Cognitive Systems*, *2*(3), 291–300.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. Plenum. https://doi.org/10.1016/0092-6566(85)90023-6

DeMartino, G. F., Moyer, J. D., & Watkins, K. M. (2016). Achieving fair trade through a social tariff regime: A policy thought experiment. *Cambridge Journal of Economics*, *40*(1), 69–92. https://doi.org/10.1093/CJE%2FBEU066

DeNisi, A. S., & Murphy, K. R. (2017). Performance appraisal and performance management: 100 years of progress? *Journal of Applied Psychology*, *102*(3), 421–433. https://doi.org/10.1037/apl0000085

Dietrich, A., & Haider, H. (2015). Human creativity, evolutionary algorithms, and predictive representations: The mechanics of thought trials. *Psychonomic Bulletin & Review*, 22(4), 897–915.

Droogendyk, L., Wright, S. C., Lubensky, M., & Louis, W. R. (2016). Acting in solidarity: Cross-group contact between disadvantaged group members and advantaged group allies. *Journal of Social Issues*, 72(2), 315–334. https://doi.org/10.1111/josi.12168

Eden, D. (2017). Field experiments in organizations. *Annual Review of Organizational Psychology and Organizational Behavior*, 4, 91–122. https://doi.org/10.1146/annurev-orgpsych-041015-062400

Edwards, J. R. (2008). To prosper, organizational psychology should … overcome methodological barriers to progress. *Journal of Organizational Behavior*, 29(4), 469–491. https://doi.org/10.1002/job.529

*Elias, J. Z., & Gallagher, S. (2014). Word as object: A view of language at hand. *Journal of Cognition and Culture*, 14(5), 373–384. https://doi.org/10.1163/15685373-12342132

*Emmerich, N., & Gordjin, B. (2018). A morally permissible moral mistake? Reinterpreting a thought experiment as proof of concept. *Journal of Bioethical Inquiry*, 15(2), 269–278.

Erskine, S. E., & Bilimoria, D. (2019). White allyship of Afro-Diasporic women in the workplace: A transformative strategy for organizational change. *Journal of Leadership and Organizational Studies*, 26(3), 319–338. https://doi.org/10.1177/1548051819848993

Ferguson, T. D., & Ketchen, D. J. Jr. (1999). Organizational configurations and performance: The role of statistical power in extant research. *Strategic Management Journal*, 20(4), 385–395. https://doi.org/10.1002/(SICI)1097-0266(199904)20:4<385::AID-SMJ24>3.0.CO;2-X

Fischer, I. (2009). Friend or foe: Subjective expected relative similarity as a determinant of cooperation. *Journal of Experimental Psychology: General*, 138(3), 341–350. https://doi.org/10.1037/a0016073

*Fisher, D. J. (2020). A psychoanalyst serves on a jury. *Canadian Journal of Psychoanalysis*, 28(2), 277–298.

*Folger, R., & Turillo, C. J. (1999). Theorizing as the thickness of thin abstraction. *Academy of Management Review*, 24(4), 742–758. https://doi.org/10.2307/259352

Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26(4), 331–362. https://doi.org/10.1002/job.322

Gendler, T. S. (1998). Galileo and the indispensability of scientific thought experiment. *The British Journal for the Philosophy of Science*, 49(3), 397–424. https://doi.org/10.1093/bjps/49.3.397

Gendler, T. S. (2010). *Intuition, imagination, and philosophical methodology*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199589760.001.0001

Goodman, D. J. (2011). *Promoting diversity and social justice: Educating people fromprivileged groups*. Routledge.

Gottfredson, R. K., & Aguinis, H. (2017). Leadership behaviors and follower performance: Deductive and inductive examination of theoretical rationales and underlying mechanisms. *Journal of Organizational Behavior*, 38(4), 558–591. https://doi.org/10.1002/job.2152

Gottfredson, R. K., Wright, S. L., & Heaphy, E. D. (2020). A critique of the Leader-Member Exchange construct: Back to square one. *The Leadership Quarterly*, 31(6), Article 101385. https://doi.org/10.1016/j.leaqua.2020.101385

Guest, D., & King, Z. (2004). Power, innovation and problem-solving: The personnel managers' three steps to heaven? *Journal of Management Studies*, 41(3), 401–423.

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. *Advances in Experimental Social Psychology*, 8, 45–99.

Hall, G. S. (1917). Practical relations between psychology and the war. *Journal of Applied Psychology*, 1(1), 9–16. https://doi.org/10.1037/h0070238

*Harre, R., & Wang, H. T. (1999). Setting up a real 'Chinese Room': An empirical replication of a famous thought experiment. *Journal of Experimental & Theoretical Artificial Intelligence*, 11(2), 153–154. https://doi.org/10.1080/095281399146517

*Hatherly, D., Mitchell, R. K., Mitchell, J. R., & Lee, J. H. (2020). Reimagining profits and stakeholder capital to address tensions among stakeholders. *Business & Society*, 59(2), 322–350. https://doi.org/10.1177/0007650317745637

Haukioja, J. (2020). Semantic burden-shifting and temporal externalism. *Inquiry*, 63(9–10), 919–929. https://doi.org/10.1080/0020174X.2020.1805704

Highhouse, S. (2009). Designing experiments that generalize. *Organizational Research Methods*, 12(3), 554–566. https://doi.org/10.1177/1094428107300396

Holt, L. (1999). Rationality is still hard work: Some further notes on the disruptive effects of deliberation. *Philosophical Psychology*, 12(2), 215–219. https://doi.org/10.1080/095150899105882

Hom, P. W., Lee, T. W., Shaw, J. D., & Hausknecht, J. P. (2017). One hundred years of employee turnover theory and research. *Journal of Applied Psychology*, 102(3), 530–545. https://doi.org/10.1037/apl0000103

*Hong, J. F. (2012). Glocalizing Nonaka's knowledge creation model: Issues and challenges. *Management Learning*, 43(2), 199–215. https://doi.org/10.1177/1350507611428853

Ilgen, D. R. (1999). Teams embedded in organizations: Some implications. *American Psychologist*, 54(2), 129–139. https://doi.org/10.1037/0003-066X.54.2.129

Jabri, M., & Pounder, J. S. (2001). The management of change: A narrative perspective on management development. *Journal of Management Development*, 20(8), 682–691. https://doi.org/10.1108/02621710110401400

Jalal, B., & Ramachandran, V. S. (2017). The twin vantage point paradox: A thought experiment. *The Journal of Mind and Behavior*, 38(2), 111–118.

Joshi, A., & Roh, H. (2009). The role of context in work team diversity research: A meta-analytic review. *Academy of Management Journal*, 52(3), 599–627. https://doi.org/10.5465/AMJ.2009.41331491

Judge, T. A., Piccolo, R. F., & Ilies, R. (2004). The forgotten ones? The validity of consideration and initiating structure in leadership research. *Journal of Applied Psychology*, 89(1), 36–51. https://doi.org/10.1037/0021-9010.89.1.36

*Kadvany, J. (2010). Indistinguishable from magic: Computation is cognitive technology. *Minds and Machines*, 20(1), 119–143. https://doi.org/10.1007/s11023-010-9185-z

Kanter, R. (1977). *Men and women of the organization*. Basic Books.

Klima, K., & Morgan, M. G. (2012). Thoughts on whether government should steer a tropical cyclone if it could. *Journal of Risk Research*, 15(8), 1013–1020. https://doi.org/10.1080/13669877.2012.686054

Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., Schaufeli, W. B., de Vet, H. C. W., & van der Beek, A. J. (2011). Conceptual frameworks of individual work performance: A systematic review. *Journal of Occupational and Environmental Medicine*, 53(8), 856–866. https://doi.org/10.1097/JOM.0b013e318226a763

Kornberger, M., & Mantere, S. (2020). Thought experiments and philosophy in organizational research. *Organization Theory*, 1(3), 1–19. https://doi.org/10.1177/2631787720942524

Kruger, A. (2002). Counseling and philosophy: A personal existential view. *American Journal of Pastoral Counseling*, 6(1), 51–62. https://doi.org/10.1300/J062v06n01_04

Kuhn, T. (1964). A function for thought experiments. In T. S. Kuhn (Ed.), *The essential tension: Selected studies in scientific tradition and change* (pp. 240–265). University of Chicago Press.

Lachenicht, L. G. (1993). A skeptical argument concerning the value of a behavioural solution for AIDS. *South Africa Journal of Psychology*, 23(1), 15–20. https://doi.org/10.1177/008124639302300103

Le, H., Oh, I.-S., Robbins, S. B., Ilies, R., Holland, E., & Westrick, P. (2011). Too much of a good thing: Curvilinear relationships between personality traits and job performance. *Journal of Applied Psychology*, 96(1), 113–133. https://doi.org/10.1037/a0021016

Lefkowitz, J. (2008). To prosper, organizational psychology should … expand the values of organizational psychology to match the quality of its ethics. *Journal of Organizational Behavior*, 29(4), 439–453. https://doi.org/10.1002/job.527

*Leicester, J. (2012). Counterfactuals, belief, and inquiry by thought experiment. *The Journal of Mind and Behavior*, 33(3–4), 195–203.

Lennox, J. G. (1991). Darwinian thought experiments: A function for just-so stories. In T. Horowitz & G. J. Massey (Eds.), *Thought experiments in science and philosophy* (pp. 223–245). Rowman & Littlefield Publishers. https://doi.org/10.2307/2108510

Liberale, A. P., & Kovach, J. V. (2017). Reducing the time for IRB reviews: A case study. *Journal of Research Administration*, 48(2), 37–50.

Liu, J. H. (2017). Neo-Confucian epistemology and Chinese philosophy: Practical postulates for actioning psychology as a human science. *Asian Journal of Social Psychology*, 20(2), 137–149. https://doi.org/10.1111/ajsp.12168

*Lucas, R. E. Jr. (2003). Macroeconomic priorities. *American Economic Review*, 93(1), 1–14. https://doi.org/10.1257/000282803321455133

*Mankiw, N. G. (2013). Defending the one percent. *Journal of Economic Perspectives*, 27(3), 21–34. https://doi.org/10.1257/jep.27.3.21

Mathieu, J. E., Hollenbeck, J. R., van Knippenberg, D., & Ilgen, D. R. (2017). A century of work teams in the Journal of Applied Psychology. *Journal of Applied Psychology*, 102(3), 452–467. https://doi.org/10.1037/apl0000128

*Maziarz, M. (2017). Ethics, uncertainty, and macroeconomics. *Ethics in Economic Life*, 20(4), 51–63.

McAllister, J. W. (1996). The evidential significance of thought experiment in science. *Studies in History and Philosophy of Science Part A*, 27(2), 233–250. https://doi.org/10.1016/0039-3681(95)00044-5

*McDonald, R. L., Chilton, S. M., Jones-Lee, M. W., & Metcalf, H. R. T. (2017). Evidence of variable discount rates and non-standard discounting in mortality risk valuation. *Journal of Environmental Economics and Management*, 82, 152–167. https://doi.org/10.1016/J.JEEM.2016.11.005

McMullen, J. S. (2018). Organizational hybrids as biological hybrids: Insights for research on the relationship between social enterprise and the entrepreneurial ecosystem. *Journal of Business Venturing*, 33(5), 575–590. https://doi.org/10.1016/j.jbusvent.2018.06.001

Miller, J. H. (2006). Virtual automobility: Two ways to get a life. *The Sociological Review*, 54(Suppl. 1), 193–207. https://doi.org/10.1111/j.1467-954X.2006.00644.x

Mitchell, T. R., Holtom, B. C., Lee, T. W., Sablynski, C. J., & Erez, M. (2001). Why people stay: Using job embeddedness to predict voluntary turnover. *Academy of Management Journal*, 44(6), 1102–1121. https://doi.org/10.2307/3069391

Moody, E. A. (1951). Galileo and Avempace: The dynamics of the Leaning Tower experiment (I). *Journal of the History of Ideas*, 12(2), 163–193.

Mor Barak, M. E. (2022). *Managing diversity: Toward a globally inclusive workplace* (5th ed.). Sage.

*Nanay, B. (2015). Cognitive penetration and the gallery of indiscernibles. *Frontiers in Psychology*, 5, 1527. https://doi.org/10.3389/fpsyg.2014.01527

Newton, I. (1728). *A treatise of the system of the world*. Courier Corporation.

*Nothhaft, H., & Stensson, H. (2019). Explaining the measurement and evaluation stasis: A thought experiment and a note on functional stupidity. *Journal of Communication Management*, 23(3), 213–227. https://doi.org/10.1108/JCOM-12-2018-0135

O'Reilly, C. A. (1977). Supervisors and peers as information sources, group supportiveness, and individual decision-making performance. *Journal of Applied Psychology*, 62(5), 632–635. https://doi.org/10.1037/0021-9010.62.5.632

Ostrove, J. M., & Brown, K. T. (2018). Are allies who we think they are?: A comparative analysis. *Journal of Applied Social Psychology*, 48(4), 195–204. https://doi.org/10.1111/jasp.12502

*Otero-Iglesias, M., & Weissenegger, M. (2020). Motivations, security threats and geopolitical implications of Chinese investment in the EU energy sector: The case of CDP Reti. *European Journal of International Relations*, 26(2), 594–620. https://doi.org/10.1177/1354066119871350

Pfeffer, J. (1981). Understanding the role of power in decision making. *Classics of Organization Theory*, 3, 404–423.

Ployhart, R. E., Schmitt, N., & Tippins, N. T. (2017). Solving the Supreme Problem: 100 years of selection and recruitment at the Journal of Applied Psychology. *Journal of Applied Psychology*, 102(3), 291–304. https://doi.org/10.1037/apl0000081

Popper, K. (1959). On the use and misuse of imaginary experiments, especially in quantum theory. In K. Popper (Ed.), *The logic of scientific discovery* (pp. 442–456). Routledge.

Ragins, B. R., & Kram, K. E. (2007). *The handbook of mentoring at work: Theory, research, and practice*. Sage Publications.

*Raverty, A. (2007). Are we monks, or are we men? The monastic masculine gender model according to the rule of Benedict. *The Journal of Men's Studies*, 14(3), 269–291. https://doi.org/10.3149/jms.1403.269

*Reichstein, A. (2019). A right to die for prisoners? *International Journal of Prisoner Health*, 16(1), 56–66. https://doi.org/10.1108/ijph-07-2019-0036

*Rinsley, D. B. (1980). A thought experiment in psychiatric genetics. *Bulletin of the Menninger Clinic*, 44(6), 628–638.

Roberson, Q., Ryan, A. M., & Ragins, B. R. (2017). The evolution and future of diversity at work. *Journal of Applied Psychology*, 102(3), 483–499. https://doi.org/10.1037/apl0000161

Sacco, J. M., & Schmitt, N. (2005). A dynamic multi-level model of demographic diversity and misfit effects. *Journal of Applied Psychology*, 90(2), 203–231. https://doi.org/10.1037/0021-9010.90.2.203

Sarmiento, R., & Shukla, V. (2011). Zero-sum and frontier trade-offs: An investigation on compromises and compatibilities amongst manufacturing capabilities. *International Journal of Production Research*, 49(7), 2001–2017. https://doi.org/10.1080/00207540903555544

Scandura, T. A., & Williams, E. A. (2000). Research methodology in management: Current practices, trends, and implications for future research. *Academy of Management Journal*, 43, 1248–1264. https://doi.org/10.5465/1556348

Schrödinger, E. (1935). Die gegenwärtige situation in der quantenmechanik. *Naturwissenschaften*, 23(49), 823–828. https://doi.org/10.1007/BF01491914

Schwab, D. P. (2005). *Research methods for organizational studies*. Psychology Press.

Shaffer, J. A., DeGeest, D., & Li, A. (2016). Tackling the problem of construct proliferation: A guide to assessing the discriminant validity of conceptually related constructs. *Organizational Research Methods*, 19(1), 80–110. https://doi.org/10.1177/1094428115598239

*Smith, C. (2007). Working from the inside out: Management and leadership through the lens of the perennial wisdom tradition. *Journal of Management Development*, 26(5), 475–483. https://doi.org/10.1108/02621710710748301

Sorensen, R. (1998). *Thought experiments*. Oxford University Press. https://doi.org/10.1093/019512913X.001.0001

Spanierman, L. B., & Smith, L. (2017). Roles and responsibilities of White allies: Implications for research, teaching, and practice. *The Counseling Psychologist*, 45(5), 606–617. https://doi.org/10.1177/0011000017717712

*Stögbauer, C., & Komlos, J. (2004). Averting the Nazi seizure of power: A counterfactual thought experiment. *European Review of Economic History*, 8(2), 173–199. https://doi.org/10.1017/S1361491604001145

Sue, D. W. (2017). The challenges of becoming a White ally. *The Counseling Psychologist*, 45(5), 706–716. https://doi.org/10.1177/0011000017719323

*Tateo, L., & Valsiner, J. (2015). Time breath of psychological theories: A meta-theoretical focus. *Review of General Psychology*, 19(3), 357–364. https://doi.org/10.1037/gpr0000035

Tett, R., Walser, B., Brown, C., Simonet, D., & Tonidandel, S. (2013). The 2011 SIOP graduate program benchmarking survey part 3: Curriculum and competencies. *The Industrial-Organizational Psychologist*, 50(4), 69–90.

Turner, J. H. (1991). Developing cumulative and practical knowledge through metatheorizing. *Sociological Perspectives*, 34(3), 249–268. https://doi.org/10.2307/1389510

Van Bockhaven, W., & Matthyssens, P. (2017). Mobilizing a network to develop a field: Enriching the business actor's mobilization analysis toolkit. *Industrial Marketing Management*, 67, 70–87. https://doi.org/10.1016/j.indmarman.2017.08.001

van Knippenberg, D., & Sitkin, S. B. (2013). A critical assessment of charismatic—Transformational leadership research: Back to the drawing board? *The Academy of Management Annals*, 7(1), 1–60. https://doi.org/10.1080/19416520.2013.759433

*Vasileiou, E. (2021). Are markets efficient? A quantum mechanics view. *Journal of Behavioral Finance*, 22(2), 214–220. https://doi.org/10.1080/15427560.2020.1772260

Viswesvaran, C., & Ones, D. S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment*, 8(4), 216–226. https://doi.org/10.1111/1468-2389.00151

Warren, M. A., Bordoloi, S., & Warren, M. T. (2020). Good for the goose and good for the gander: Examining positive psychological benefits of male allyship for men and women. *Psychology of Men & Masculinities*, 22(4), 723–731. https://doi.org/10.1037/men0000355

*Weick, K. E. (1989). Theory construction as disciplined imagination. *Academy of Management Review*, 14(4), 516–531. https://doi.org/10.2307/258556

Welbourne, T. M., Johnson, D. E., & Erez, A. (1998). The role-based performance scale: Validity analysis of a theory-based measure. *Academy of Management Journal*, 41(5), 540–555. https://doi.org/10.2307/256941

*Wempe, B. (2008). Contractarian business ethics: Credentials and design criteria. *Organization Studies*, 29(10), 1337–1355. https://doi.org/10.1177/0170840608093546

Wolfe, B. E. (2000). Toward an integrative theoretical basis for training psychotherapists. *Journal of Psychotherapy Integration*, 10(3), 233–246. https://doi.org/10.1023/A:1009492728103

## AUTHOR BIOGRAPHIES

**Herman Aguinis** is the Avram Distinguished Scholar and Chair of the Department of Management at The George Washington University School of Business. His research addresses behavioral science and data science, he served as President of the Academy of Management (2021–2022), and every year since 2018, Web of Science has ranked him among the world's 100 most impactful researchers in Economics and Business.

**Jose R. Beltran** is a PhD Candidate at the Ivy College of Business at Iowa State University. His main research interests include strategic leadership, multilevel measurement of performance, bibliometrics, and CEO personality. He has also published on the operationalization of psychological constructs using artificial intelligence.

**Estelle E. Archibold** is a Postdoctoral Scholar at The Pennsylvania State University in the Department of Management and Organizations, Smeal College of Business. Estelle's interests include research in conflict and cooperation. She applies embodiment theory to her study of these phenomena and uses qualitative as well as mixed methods in her research.

**Esther Lamarre Jean** is a researcher at Google. Her research focuses on identity, in particular employee experiences of identity construction occurring in the workplace. Further research interests include leadership, well-being, and research methods.

**Darryl B. Rice** is an Assistant Professor of Management in the Farmer School of Business at Miami University. His research areas are organizational justice, behavioral ethics, and diversity, equity, and inclusion. His work has been published in outlets such as *Journal of Management Studies*, *Human Relations*, *Journal of Organizational Behavior*, *Journal of Business Ethics*, and *Journal of Business and Psychology*.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.