Ehrig, Timo; Schmidt, Jens

Theory-based learning and experimentation: How strategists can systematically generate knowledge at the edge between the known and the unknown

STRATEGIC MANAGEMENT JOURNAL WILEY

Theory-based learning and experimentation: How strategists can systematically generate knowledge at the edge between the known and the unknown

Timo Ehrig[1] | Jens Schmidt[2] ⓘ

[1]Department of Strategy and Innovation, Copenhagen Business School, Frederiksberg, Denmark

[2]Department of Industrial Engineering and Management, Aalto University, Aalto, Finland

Correspondence
Jens Schmidt, Department of Industrial Engineering and Management, Aalto University, 02150 Espoo, Finland.
Email: jens.schmidt@aalto.fi

Abstract

Research Summary: We present a framework for theory-based learning and experimentation under uncertainty. Strategists' assumptions about how an envisioned future can be reached are likely incomplete and possibly wrong, for instance, if critical contingencies have been overlooked. We explain how strategists can learn from thinking about and testing necessary conditions for an envisioned future to materialize. By logically linking assumptions to consequences our framework allows drawing inferences from experiments with testable assumptions about elements of a strategy that cannot be tested without major investments. Our framework contains the first formal model of learning from arguments in the strategy literature. By using our framework, strategists can maintain focus on an envisioned future while at the same time systematically seeking out reasons and evidence for why they are wrong.

Managerial Summary: We develop a framework that helps strategists to learn and understand what it takes to reach ambitious goals when there is substantial uncertainty. We ask strategists to formulate their assumptions as a theory: what needs to be true

for their goal to materialize. Our framework enables strategists to scrutinize and improve their assumptions by raising objections against their theory and by pointing them to critical experiments to learn whether their assumptions hold. Using our results, strategists can in particular identify overlooked critical contingencies. Overall, we suggest how strategists should revise their beliefs about what it takes to be successful in the light of evidence and arguments for and against their strategy.

KEYWORDS

belief revision, experimentation, learning, managerial theories, uncertainty

1 | INTRODUCTION

> A good scientist pushes to the edge of knowledge and then reaches beyond, forming a conjecture—a hypothesis—about how things work in that unknown territory. [...] In the same way, a good business strategy deals with the edge between the known and the unknown (Rumelt, 2011, p. 242f).

Strategists and entrepreneurs must go beyond current knowledge to pursue profitable opportunities (Knight, 1921). Knowledge about the future is incomplete and many contingencies are unforeseen (Simon, 1962). Consequently, scholars have explored possibilities how strategists can generate knowledge on the edge between the known and the unknown. One stream of the literature views knowledge generation under uncertainty as a process of experimentation. Implementing a new business strategy can be seen as an experiment (Pillai, Goldfarb, & Kirsch, 2020), and strategists who follow scientific principles and systematically develop hypotheses and run experiments to test their assumptions make better decisions than strategists who do not (Camuffo, Cordova, Gambardella, & Spina, 2020; Ries, 2011). A second, related stream of the literature proposes that strategists can formulate and use theories to envision a possible future and make decisions on their basis (Felin & Zenger, 2009, 2017; Gavetti & Menon, 2016). Here it has been suggested that theories are particularly valuable when they are falsifiable, and that theories provide focus for strategy implementation.

In the context of the former stream of literature, scholars have started to address the question how strategists should experiment (Agrawal, Gans, & Stern, 2021), while at the same time emphasizing that experimentation can never completely eliminate uncertainty about a strategy before irreversible investments are made (Gans, Stern, & Wu, 2019). Here it is often assumed that strategists already know the set of experiments they could possibly run and that they have already formed beliefs—albeit probabilistic or ambiguous—about the mapping of the results of the experiments to the value of an idea and appropriate ways

to implement it.[1] The latter stream of literature, however, argues that strategists' initial challenge is a different one: they need to first of all mentally construct a representation or map of what they see as an opportunity. In particular, theories encapsulate a point of view about an uncertain future and what is needed to be successful (Felin & Zenger, 2017; Schmidt, 2015). They guide experimentation and thus guide learning about assumptions (Zellweger & Zenger, 2021). But, as we argue in this article, theories also relate what is observable today to consequences in the future and thus allow for making inferences from early stage experiments about assumptions that cannot be tested without major investments. Yet if theories are mentally constructed by strategists, it follows that they are likely incomplete and possibly wrong (Knight, 1921; Simon, 1962). Thus, theories can also misguide strategists by pointing them to experiments that are not informative about the success of their idea, or by wrongly suggesting that strategists should go forward with major investments. When strategists construct theories, their challenge is to identify which contingencies even matter. There might also be arguments why one or more of their assumptions are wrong, which could force them to revise their theory (Gardenfors, 1988; Spohn, 2012). This raises the question: How can theories be useful for guiding experimentation and decision-making toward an envisioned future? In particular, how can theories be improved before large or even irreversible investments are made on their basis?

To answer these questions, we propose that theories can be systematically evaluated and improved by learning about and testing the assumptions that the theories encapsulate. Because theories may be incomplete or wrong, and some assumptions cannot be tested without making major investments, we propose that strategists should exploit an asymmetry in learning theories, which has also been pointed out by Popper (1959) in the philosophy of science: strategists can learn that they are wrong (i.e., that their point of view and the underlying set of assumptions cannot be a basis for the creation of an envisioned future) whereas they cannot learn that they are ultimately right, and consequently when contingencies are not fully knowable theories should be considered as being tentative and subject to further possible revision.

Concretely, we develop a normative framework based on modeling a strategist's theory as a set of premises that imply a "conjecture," which is a belief that formulates a future possible state of the world that is associated with success (Prahalad & Hamel, 1994; Rindova & Martins, 2021). For example, the conjecture that it is possible for Tesla Motors as a manufacturer of electric vehicles to become one of the major players in the high-volume segment of the car market has been central to the company's strategy at least since 2006. A theory is then a strategist's *subjective* formulation of the necessary and jointly sufficient conditions under which the conjecture will be true. We call such conditions "premises," which encapsulate what strategists believe to be the major contingencies and how they map on the conjecture. For example, Tesla's strategists hypothesized that the successful development of a battery that provides enough power for long-range rides is a necessary condition for the conjecture about success in the market for electric vehicles. Other contingencies (e.g., a change in the oil price) may be considered to be less relevant for the conjecture and thus *not* be included in the

---

[1]In the formal literature on the topic, experimentation is conceptualized as updating prior beliefs using Bayes' rule after receiving a (potentially noisy) signal from a fixed space of possible signals (e.g., Agrawal et al., 2021). However, the space of possible signals is assumed to be known to the strategist and—importantly—unchanging, that is strategists face a signal extraction problem rather than the problem of evaluating and possibly changing the mapping from possible signals to outcomes, which is the challenge highlighted in our article.

theory. Using our framework strategists can then identify testable premises, where the results of such tests allow them to make inferences about assumptions that are not testable without making major investments, thus mitigating the "paradox of entrepreneurship" noted by Gans et al. (2019). For example, from experiments with batteries in prototypes Tesla may infer whether—given its other assumptions—the plan to become a major player in the high-volume segment is feasible.

The challenge is, of course, that strategists' assumptions and how they relate to each other (and thus what would be valid inferences) may be wrong. For instance, Tesla's strategists may have been correct that solving the battery problem is necessary and the oil price is irrelevant for the success of Tesla's overall strategy, but they could have overlooked other necessary conditions for success, for instance related to charging infrastructure. To account for this possibility, in our framework strategists may learn from an argument formulated as an "objection," which is the statement that a belief in their theory is wrong and which is backed by a "counter-theory": a theory that explains why that belief is wrong. For example, Tesla's strategists realized that a key premise for their conjecture about success was that the battery of electric vehicles must be large enough to supply sufficient driving range. An argument against the conjecture would be that a large battery will also make electric vehicles so costly to make that they will be more expensive than comparable gasoline-power cars and customers are not willing to pay higher prices for electric vehicles. Our framework helps strategists learn from such arguments against the conjecture or one of the premises even if they are rejected because they can expose hidden premises: To reject the argument and continue believing in demand for electric vehicles in the volume segments, Tesla's strategists must now also believe that the company will be successful *despite* having higher prices due to large batteries. Specifically, the belief that customers are willing to pay higher prices for electric vehicles now becomes a premise: a critical assumption on which the strategist should consequently focus attention (Ocasio, 1997) and possibly learn about through experimentation (Camuffo et al., 2020).

As basis for our framework, we formulate two axioms. First, strategists should be open to revising their assumptions when they learn that they are wrong. This means that strategists learn from counter-theories, integrate new assumptions into their theories but also give up beliefs if they turn out to be wrong. It also means that while strategists' beliefs are subjective, these beliefs are constrained by the requirement that strategists logically deduce what follows from what they already believe, and that they resolve logical contradictions if they appear. Second, strategists should order their beliefs in terms of how willing they are to question them. In particular, strategists should be as willing to question a belief (call it C) as they are willing to question its *weakest premise*, that is, the strategist's belief about the weakest necessary condition for the belief that, together with their other beliefs, forms a sufficient condition for C. Using our framework, we then derive three formal results about learning from counter-theories. First, we show under which conditions a strategist who is exposed to one or more counter-theories should continue to believe that the conjecture is true. Second, we show what and, in particular, which hidden premises the strategist learns from counter-theories. And third, we show when the strategist should focus experiments on such newly learned beliefs.

In the following, we develop our framework. While our core results about experimentation and learning are mathematically deduced, our exposition is verbal, using a stylized version of the case of Tesla's strategy from the perspective of 2008 as illustrative example. We then discuss the implications of our framework and how it relates to and complements other approaches for dealing with uncertainty, including learning by applying Bayes' rule (e.g., Camuffo et al., 2020;

Gans et al., 2019), recent work on strategy process and the problem-based view (Nickerson, Yen, & Mahoney, 2012; Nickerson & Zenger, 2004) as well as established practitioner approaches and frameworks (Ries, 2011; Schoemaker, 1993). We close by discussing limitations and providing guidance for future research.

2 | A FRAMEWORK FOR THEORY-BASED LEARNING AND EXPERIMENTATION

2.1 | Conjectures, theories, and premises

Our framework is based on the idea that managers can develop knowledge about an unknown future by formulating theories that encapsulate exclusive, firm-specific points of view and then testing and refining the underlying assumptions (Felin & Zenger, 2009, 2017; Gavetti & Menon, 2016; Schmidt, 2015). Nickerson and Argyres (2018) argue that the knowledge embodied in theories is an essential input into strategic problem formulation and thereby guides strategic decision-making and strategy implementation. While the process of formulating and testing theories will typically be a group-level process (Baer, Dirks, & Nickerson, 2013; Nickerson & Zenger, 2004), in order to focus on the underlying core mechanisms in the following we make the simplifying assumption of an individual-level process.

To illustrate out framework, we use a simplified and stylized version of a strategist's thinking about Tesla from the perspective of 2008 when Tesla had just introduced its first model, the Tesla Roadster. We use the example for the purpose of illustrating our framework and therefore claim neither descriptive accuracy nor completeness. However, as noted in the introduction, theories are likely incomplete and assumptions wrong. The value of our framework is precisely that it can overcome inaccuracy and incompleteness in theories.

After Elon Musk had become a major shareholder in Tesla, in 2006 he revealed his "secret masterplan" in a blog post that described his point of view about how it would be possible for Tesla to create sales for electric vehicles first in the high-end and later in the high-volume segments of the market. Concretely, his reasoning was:

> The overarching purpose of Tesla Motors (and the reason I am funding the company) is to help expedite the move from a mine-and-burn hydrocarbon economy toward a solar electric economy, which I believe to be the primary, but not exclusive, sustainable solution. Critical to making that happen is an electric car without compromises, which is why the Tesla Roadster is designed to beat a gasoline sports car like a Porsche or Ferrari in a head to head showdown. [...] The strategy of Tesla is to enter at the high end of the market, where customers are prepared to pay a premium, and then drive down market as fast as possible to higher unit volume and lower prices with each successive model (Musk, 2006).

Central to this reasoning was the idea that Tesla would be successful in the high-volume segments of the market if certain conditions were fulfilled. We call such a belief a *conjecture* about what will be true in the future (cf. Nickerson & Argyres, 2018): a future possible state of the world that is associated with success (what Prahalad & Hamel, 1994 have termed a "strategic intent"). The Merriam-Webster (2021) dictionary defines a conjecture as "inference formed without proof or sufficient evidence". A core part of this definition is that conjectures are

formed on the basis of incomplete knowledge and that they are based on an inference. Furthermore, Musk's above statement identifies the underlying assumptions from which an inference toward the conjecture is made. Concretely, these assumptions include the belief that Tesla's high-end car is perceived to be superior to traditional cars with internal combustion engine ("ICE-powered cars") and that achieving success in the high-end segment leads to success for electric vehicles in the volume segments. Missing from this theory are issues such as charging infrastructure, costs or competition. As we will show further below, such overlooked issues can be included when strategists learn hidden premises from arguments why their theories are incomplete or wrong.

In our framework, theories contain the assumptions behind a conjecture in the form of connected beliefs. A belief can be any elementary statement (e.g., *Customers perceive Tesla's high-end electric car to be equal or superior to ICE-powered high-end cars by the established brands like Porsche, Ferrari, Mercedes, and Audi*) or a sentence that relates statements to each other using standard logic (e.g., *If customers perceive Tesla's high-end electric car to be equal or superior to ICE-powered high-end cars by the established brands like Porsche, Ferrari, Mercedes, and Audi, Tesla will sell as many or more high-end electric cars as the number of ICE-powered high-end cars sold by these established brands*). Combining these allows for drawing logical conclusions (e.g., to continue the example, a strategist who holds the two aforementioned beliefs will conclude that *Tesla will sell as many or more high-end electric cars as the number of ICE-powered high-end cars sold by these established brands*). For depicting beliefs and connections between beliefs we use standard logic notation.[2] In our notation, capital letters denote elementary statement. Both elementary statements and sentences that relate statements to each other can be either true or false.

Theories are based on drawing logical conclusions from premises, which are the strategist's assumptions behind a conjecture: the conditions that are—from the perspective of a strategist—necessary and jointly sufficient for a particular future to materialize. We call a belief A of the strategist a premise for another belief Z of the strategist, if the strategist thinks that A is a necessary condition for Z and that A, together with his other beliefs B, C and so on, forms a sufficient condition for Z. More generally, each belief Z that logically follows from other beliefs is associated with a set of premises (all the necessary conditions that together form a sufficient condition for the belief Z). We call this set of premises a *theory* for the belief Z (see Table 1 for a definition of the key terms). Importantly, if a strategist believes that all premises are true, then he must also believe that the conjecture is true as it logically follows from the premises.

To learn on the basis of theories requires strategists to reason using deductive logic. The use of logic allows them to formulate their assumptions and enables them to think with precision about what needs to be true so that, according to their thinking, an envisioned future will materialize. It also allows them to identify testable assumptions and then draw inferences from these to beliefs that are not testable ex ante or that are only testable at prohibitively high costs.

To be useful for theory-based learning and experimentation, beliefs have to be formulated in a way that they are at least in principle verifiable. In the Tesla case, we therefore, for example, refer to a specific segment and competing car manufacturers in formulating the focal

---

[2]In standard logic, logical operators are used to connect statements. In this article we use the following ones: ¬ denotes negation (¬A means that A is false); & denotes logical conjunction (A&B is true only if both A and B are true) and ⇒ denotes logical implication (if A is true then B is true). In our Figures, the beliefs depicted in the boxes are either atomic statements or beliefs that relate statements. We use a verbal exposition for the ease of presentation, but explicitly mention the expressed logical connections between statements to add precision to the example.

TABLE 1 Definition and explanation of key terms

| Term	Dictionary definition	Explanation
Belief	Something that is accepted, considered to be true, or held as an opinion	Both premises and conjectures are beliefs. The strategist may have other beliefs relevant for thinking about a focal conjecture that are not premises or conjectures.
Conjecture	An inference formed without proof or sufficient evidence	A conjecture logically follows from a theory.
Theory	A hypothesis assumed for the sake of argument or investigation	A theory is a set of premises. The conjecture is believed to be true if and only if all the premises are believed to be true.
Premise	A proposition antecedently supposed or proved as a basis of argument or inference	A belief is a premise for a conjecture if it is a necessary condition for the conjecture and, together with other beliefs, forms a sufficient condition for the conjecture.
Objection	A reason or argument presented in opposition	An objection is a belief that the conjecture or a premise of the theory for the conjecture is false.
Counter-theory	See above definition of theory	A theory for an objection: The set of premises that imply an objection.
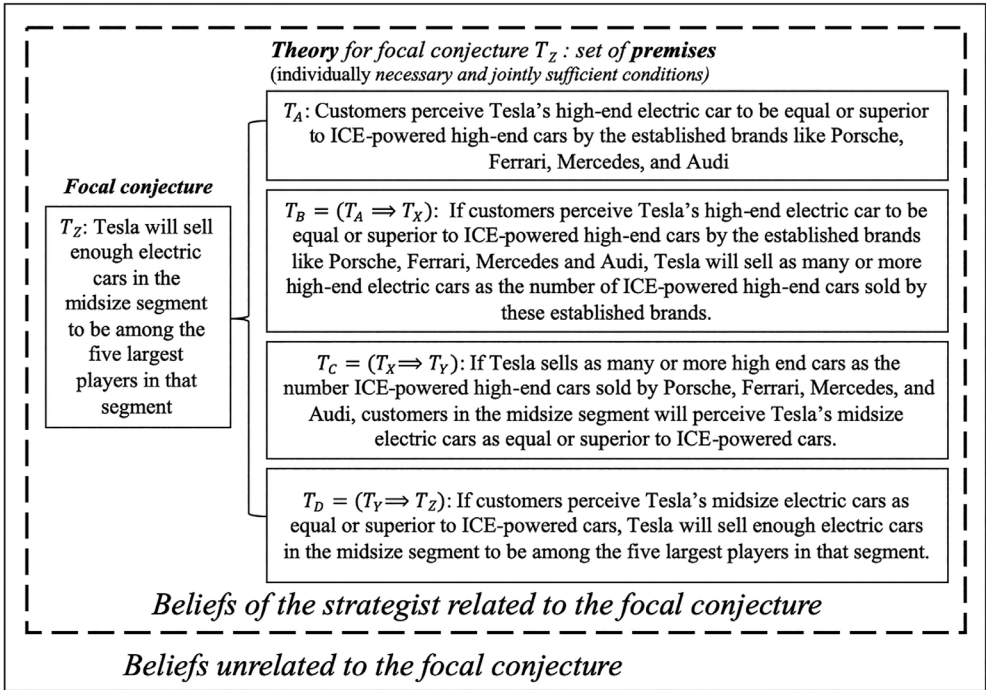
FIGURE 1 Key concepts (illustrated with the Tesla example)

conjecture: *Tesla will sell enough electric cars in the midsize segment to be among the five largest players in that segment.*[3] Figure 1 depicts Tesla's theory for T_Z, the focal conjecture in the example (beliefs of the form X refer to beliefs in general whereas beliefs of the form T_X refer to our Tesla example): T_A to T_D are necessary and jointly sufficient conditions for T_Z, as T_Z would no longer follow if any of these beliefs were removed (Figure 1 also clarifies how the key concepts of belief, conjecture, theory and premise relate to each other). If the strategist believes T_A, T_B, T_C, and T_D, he also believes T_Z because it follows logically from the premises, which thus together form a sufficient condition for T_Z. As the strategist has no other reasons to believe T_Z, the beliefs T_A, T_B, T_C, and T_D are the respective necessary conditions for T_Z and, therefore, by our definition are the premises of T_Z.

Throughout this article, and in particular in the examples depicted in the Figures, we assume that the strategist's theory provides an *exclusive*, firm-specific point of view on the situation (Felin & Zenger, 2017).[4] Under uncertainty, the strategist cannot be sure that reaching the conjecture is possible at all, and thus it makes sense to first of all ponder about the question what would be necessary to reach it. As is evident from the Tesla example in Figure 1, the theory formulated there is specific to Tesla. Exclusivity means that the strategist has no reasons to believe the conjecture other than the ones that are explicitly stated, and that he thus believes that the conjecture is true *if and only if* all premises are believed to be true (this is illustrated by the curly brackets in the figures).[5] More importantly, exclusivity also means that all beliefs that are premises for Z are *strategic beliefs*, because without any of them the strategist would not believe that Z were possible. Strategic beliefs are those beliefs the strategist considers critical for reaching the conjecture, and therefore the strategist should pay close attention to them and, if feasible, test them.

Note that the strategist might also have other beliefs that are relevant for thinking about the conjecture but which are not premises, as they are not connected to the focal conjecture (e.g., the belief that past attempts to successfully commercialize electric vehicles in the volume segment of the market have not been successful). Furthermore, note that implications are also beliefs in our framework. For example, in Figure 1, T_D (which is the belief that $T_Y \Rightarrow T_Z$) is *not* mathematically deduced from the other beliefs. Instead it is an expression of the belief of the strategist that relates customers' perception of Tesla's midsize cars relative to ICE-powered cars to sales for them relative to competitors. But in addition to the expression of such beliefs, we *also* assume that the strategist forms logical consequences of his beliefs. That is, if the strategist believes that T_A, T_B, T_C, and T_D are true, the strategist will deduce that he also believes T_Z. Thus, even though the beliefs of the strategist are subjective, they are constrained by the requirements that logical consequences among the beliefs are formed and believed and that logical consistency among beliefs is maintained.

Taken together, the starting point for our analysis is a strategist who has an exclusive theory for a focal conjecture Z, which is a logically consistent set of beliefs. In addition, the strategist can also actively consider other beliefs to be relevant for the focal conjecture but which are not

---

[3]A further consideration in formulating theories is the relevant time horizon: It makes a difference whether Z is expected to materialize within 6 months or 10 years. Here we only note the importance of specifying time horizons for applying our framework in practice, but refrain from doing so in our example because it introduces complexity that interferes with clarity of exposition.

[4]In principle, a strategist could also consider multiple alternative explanations for the conjecture. We will address this point in more detail before the discussion.

[5]If the strategist would have other reasons to believe a conjecture, the stated detailed beliefs would not be necessary conditions for the conjecture from the perspective of the strategist.

premises for Z (see Figure 1). In the following, we are interested in how theory-based learning and experimentation affects whether or not the strategist continues to believe that Z is true after scrutinizing his theory through experiments and objections backed by counter-theories. We do not, however, address the question how the strategist acts upon his beliefs or how he continues learning (even if his theory has been refuted), but we will return to these issues in the discussion.

2.2 | Learning through contradiction and maintaining consistency

At the time of writing it seems that the conjecture about sales of electric vehicles in the midsize segment may actually materialize, though in 2008 this was still very unclear and a purely hypothetical scenario. In fact, it would have been entirely possible that Tesla's strategy would fail. The ultimate reasons for success or failure could not be known in 2008. Therefore, any theory about success is likely incomplete and at least partly wrong. More generally, due to the nature of uncertainty, the strategist cannot know if his beliefs about necessary and sufficient conditions and relations among beliefs are correct, and they are therefore subject to revision. By making his assumptions explicit, however, the strategist can learn that he is wrong.

Like scientists, strategists can systematically scrutinize their assumptions by formulating them as testable hypotheses and then engaging in experimentation (Zellweger & Zenger, 2021). Yet not all assumptions are testable without major investments (Gans et al., 2019). Theories help to overcome this problem as they provide a subjective mapping from assumptions that are testable today to consequences that are only knowable in the future. To learn on the basis of theories thus clearly requires more than formulating and testing assumptions. It also requires linking beliefs that can only be tested by making major investments to premises that are testable at a relatively low cost before the major investments are made. More generally, it requires exposing oneself to the possibility that one's theories are wrong or incomplete and thus in need of revision. While the outcome of an experiment can show that an assumption has been wrong, another possibility is that there is an argument that provides a reason why a premise is wrong. We explain this in more detail below, but the general point is that both evidence obtained through an experiment and arguments against the conjecture or one of the premises in the form of objections that are backed by counter-theories can create contradictions with the theory that must be resolved by revising one's theory (Gardenfors, 1988). Like in science, strategists can not only learn from experiments, but also from juxtaposing different theories that contradict each other.

To do so the strategist must detect any inconsistencies among his beliefs, such as when a new belief he forms contradicts one or more of his existing beliefs related to the conjecture. We thus assume in our framework that the strategist engages in a mental simulation and thinks through the logical consequences of any new beliefs (e.g., from counter-theories) combined with all existing beliefs. On the other hand, we allow strategists to hold incorrect beliefs, as long as they are consistent with other currently held beliefs related to the conjecture. In fact, the key point of our framework is that the strategist learns that he has been wrong by detecting and revising inconsistencies among different theories. This is of particular relevance in the process of pondering about arguments against one's theory that are given in the form of objections and counter-theories. This also means that the set of premises (what the strategist believes to be necessary and/or sufficient conditions for the conjecture) is not fixed but may change and,

therefore, the strategist's theories may change. Below we derive formal results about this theory change process.

In our framework, the basic unit of analysis is therefore the set of beliefs of a strategist related to the conjecture, and theories (as defined above) are subsets of beliefs of a strategist related to a conjecture. Moreover, in order to learn on the basis of theories the strategist also needs to apply logic (e.g., he needs to infer the consequences of a refuted assumption or a counter-argument). Of course, it would pose unrealistic demands on a strategist to ask for consistency among all of his beliefs. We, however, ask from a strategist what we typically also ask from a scientist: His beliefs related to a focal conjecture should be consistent.[6] This leads to the first axiom of our framework.[7]

> **Axiom 1. (consistency and consequences)**: *A strategist's beliefs related to a conjecture are logically consistent. Beliefs that are logical consequences of existing beliefs related to a conjecture are included in the strategist's beliefs related to the conjecture. If contradictions arise, they are resolved by giving up beliefs.*

Of course, it is always an option to formulate theories in a way that they cannot easily be contradicted, or to refrain from experiments that will disconfirm them. For instance, strategists often use blurry concepts or measurements in formulating beliefs, or avoid thinking about logical relations—necessary and sufficient conditions—among their beliefs. But such tendencies in no way invalidate our argument, quite to the contrary. Only when strategists impose discipline on their theory formulation and revision process will they be able to detect contradictions inside their theories, and they will also be able to identify experiments and integrate counter-theories into their theories, which arguably provides them an advantage over strategists who don't use such principles (cf. Camuffo et al., 2020). In addition, they should also be less likely to persist in pursuing strategies based on wrong assumptions.

2.3 | The weakest premise

When the strategist has formulated a theory for a conjecture, he can put it to a test by an experiment. The logic of such experimentation is to try to refute one of the premises. Here our definition of a theory implies an asymmetry when it comes to evaluating it: Because premises are, by definition, individually necessary and jointly sufficient for a conjecture, a conjecture is considered true *only if all* premises are considered true, but not anymore considered to be true if just one of its premises is shown to be wrong. Therefore, a theory can be efficiently scrutinized by singling out *one* of its premises for scrutiny through experimentation. If this premise is

---

[6]We explicitly allow the strategist to make mistakes in forming subjective beliefs about necessary and sufficient conditions among his beliefs that express beliefs about his task environment. For instance, T_D may or may not describe a necessary condition for the conjecture, and T_A, T_B, T_C, and T_D may or may not be jointly sufficient for T_Z. However, we *exclude* logical reasoning mistakes in forming logical consequences. Given explicitly articulated logical relations among beliefs, the strategist is assumed to correctly infer logical consequences that follow by deduction. For instance, as the strategist articulated in T_D that he believes $T_Y \Rightarrow T_Z$ (T_Y implies T_Z), the strategist must believe T_Z if he believes that T_Y is true.

[7]The assumptions behind Axiom 1—that strategists' representations are consistent and that they maintain consistency when faced with new evidence—are well-established in the formal strategy literature (e.g., Casadesus-Masanell & Zhu, 2013; Martignoni, Menon, & Siggelkow, 2016).

supported by the outcome of the experiment, the conjecture continues to be accepted (and another premise may be scrutinized). If it turns out to be false, the strategist learns that an assumption he believed to be necessary for the conjecture is in fact wrong and he thus does not anymore have sufficient reason to believe that the conjecture can be reached.

This leaves us with the question which premise should be questioned first or, in other words, how the strategist should focus where he experiments. To answer this question, we introduce the notion of belief strength, which we define as follows:

> **Definition (belief strength).** The strength of a belief is given by the strategist's unwillingness to question it.[8]

We find it useful to assume that the strategist orders all beliefs related to the conjecture with respect to his willingness to question them. Essentially, belief strength is a measure that helps the strategist to organize and focus his experimentation when his beliefs are tentative and at most only partially grounded in factual knowledge.[9]

The notion of belief strength is particularly useful because it allows solving a dilemma of the strategist: As the conjecture expresses the strategist's strategic intent (Hamel & Prahalad, 1989), he is in principle unwilling to question it. On the other hand, running experiments that may falsify premises means to put the conjecture to a test, and thus to be willing to question it! This paradox gets resolved in our framework by requiring that the willingness to question the conjecture must be the same as the willingness to question the weakest premise in the theory for the conjecture. Thinking about necessary and sufficient conditions for the conjecture thus forces the strategist to assign a strength to the conjecture that is equal to the weakest premise. The goal is, however, to find a theory for the conjecture in which every premise is

---

[8]Thus, a belief that is easily questioned is weak. To avoid confusion, note that mathematicians sometimes speak using a different convention. They speak of strong assumptions when they refer to assumptions that should be questioned. We use weak and strong in the sense that we speak in everyday language. For example, the statement "I strongly believe that Tesla will become the manufacturer that sells most cars worldwide" would express that one is less willing to question Tesla's leadership in sales than if one would make the statement "I only weakly believe that Tesla will become the manufacturer that sells most cars worldwide."

[9]Belief strengths and probabilities are important complements. Probabilities are very useful when they can be estimated from data. Strategists' theories, however, contain beliefs (such as T_C or T_D in the example) for which the probability cannot be estimated from data ex ante, as they express relations between events that will take place in the future. Moreover, sometimes such beliefs express relations between singular events. When probabilities cannot be estimated from data, strategists must thus formulate subjective conditional beliefs (how what is testable today relates to what will happen in the future). To formulate such subjective conditionals in a Bayesian fashion is problematic. First, a technical problem appears: Strategists would have to formulate probabilities for a number of conjunctions that grows exponentially in the number of elementary statements, as Bayesian updating over possible logical relations among elementary statements would not be defined if there would not be ex ante probabilities for all possible conjunctions, given a system of statements (see Harman, 1986, p. 25ff, for a conceptual, and Griffiths & Tenenbaum, 2009, for a more technical discussion of this problem). Second, we point out that untestable subjective conditional beliefs can be improved by listening to and integrating counter-theories. However, here a second technical problem appears: It would not be defined how systems of probabilistic conditional beliefs should be updated, given a counter-theory, as Bayesian updates are conceptually meaningful when priors are conditioned on facts, not on arguments about relations of statements raised in a debate. For these reasons belief revision and belief strengths are the more appropriate technical means to cope with revising ex ante beliefs about relations of beliefs (in particular, conditionals; beliefs that express logical relations). Conceptually, belief revision and belief strength are more appropriate, as here it is defined how strategists should cope with contradictions (and contradictions will appear when relations of statements are debated ex ante, before they can be tested), while Bayesian updates cannot cope with such contradictions (Bryan, Ryall, & Schipper, 2021).

stronger than any plausible objection so that the strategist can have a strong belief in the conjecture *even though* it is only as strong as its weakest premise.

To enable a systematic comparison and revision of beliefs, it is useful to impose the "weakest premise principle" to any belief: the strength of a belief that logically follows from a set of premises is equal to the strength of its weakest premise. In other words, the strategist should question any belief as much as he is willing to question the weakest premise of this belief. This principle is well established in cognitive science (e.g. Gardenfors, 1988) and philosophy (Spohn, 2012).

In the Tesla example, one can argue that among the premises for T_Z, the strategist is most willing to question T_A (which is thus the weakest premise) and—as a consequence—focus experimentation on scrutinizing T_A while tentatively considering T_B, T_C, and T_D as true. However, if such an experiment would show that T_A is wrong, a contradiction would arise and T_A would be given up (in line with Axiom 1). More generally, if contradictions arise among a strategist's beliefs, belief strength also implies willingness to give up: to resolve contradictions, the strategist rather gives up weaker than stronger beliefs.

Taken together, we assume that a strategist has a subjective ordering over his beliefs related to the conjecture in terms of how willing he is to question them. It describes the order in which the strategist would put his beliefs to a test through experiments or be willing to give up beliefs as a result of performing a mental simulation. However, any ordering needs to fulfill the principle that a belief is as strong as its weakest premise.[10] This is stated in our second axiom as follows:

> **Axiom 2. (belief strength and the weakest premise)**: *The strategist orders the beliefs considered relevant for thinking about a focal conjecture with respect to their strength. The strength of a belief (and in particular, a conjecture) is equal to the strength of its weakest premise (the weakest necessary condition for this belief that, together with other beliefs, forms a sufficient condition for this belief). If contradictions arise, only as few beliefs as necessary are given up, and weaker beliefs are given up first.*[11]

As Axiom 2 requires the strategist to only give up as few beliefs as necessary, by definition only the weakest premise and its logical repercussions are given up when the weakest premise turns out to be false. This has important implications for theory formulation and, in particular, testing through experimentation, as we show next.

2.4 | Sub-premises and focused experimentation

In the formulation of a theory each of the premises might themselves be the consequence of "sub-premises," which are the necessary and jointly sufficient conditions for the premise or, in other words, a theory for the respective premise. More generally, theories can have a recursive structure and we can therefore apply the definition of a premise recursively. According to our

---

[10]In the extreme, some beliefs are never questioned as they are always true (like$(A\&B) \Rightarrow A$), and logically false beliefs (like $A \Rightarrow \neg A$) are always questioned.

[11]If a contradiction could be resolved by giving up either of two beliefs that have equal strength and there are no weaker beliefs that could be given up to resolve the contradiction, both beliefs will be given up.
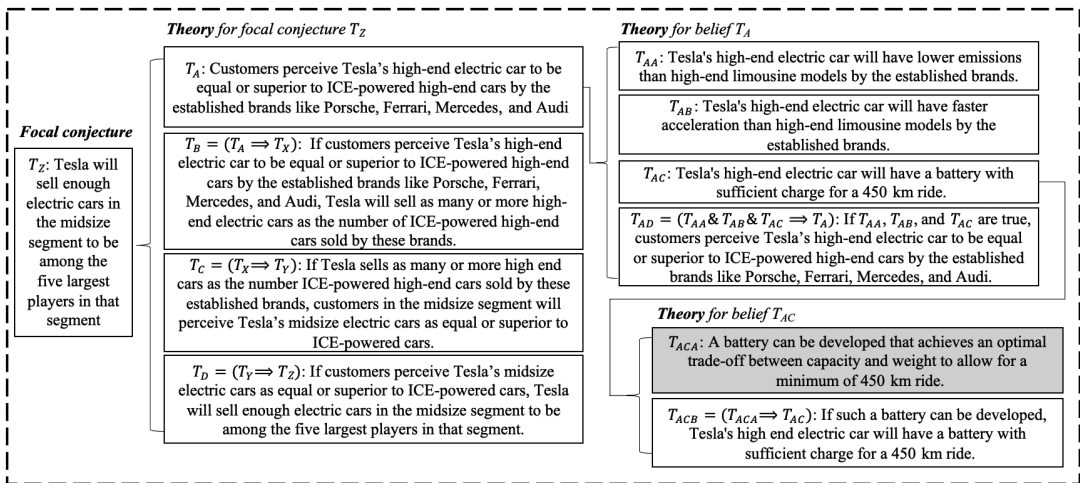
—WILEY�____

Theory for focal conjecture T_Z

Focal conjecture

T_Z: Tesla will sell enough electric cars in the midsize segment to be among the five largest players in that segment

T_A: Customers perceive Tesla's high-end electric car to be equal or superior to ICE-powered high-end cars by the established brands like Porsche, Ferrari, Mercedes, and Audi

$T_B = (T_A \Rightarrow T_X)$: If customers perceive Tesla's high-end electric car to be equal or superior to ICE-powered high-end cars by the established brands like Porsche, Ferrari, Mercedes, and Audi, Tesla will sell as many or more high-end electric cars as the number of ICE-powered high-end cars sold by these brands.

$T_C = (T_X \Rightarrow T_Y)$: If Tesla sells as many or more high end cars as the number ICE-powered high-end cars sold by these established brands, customers in the midsize segment will perceive Tesla's midsize electric cars as equal or superior to ICE-powered cars.

$T_D = (T_Y \Rightarrow T_Z)$: If customers perceive Tesla's midsize electric cars as equal or superior to ICE-powered cars, Tesla will sell enough electric cars in the midsize segment to be among the five largest players in that segment.

Theory for belief T_A

T_{AA}: Tesla's high-end electric car will have lower emissions than high-end limousine models by the established brands.

T_{AB}: Tesla's high-end electric car will have faster acceleration than high-end limousine models by the established brands.

T_{AC}: Tesla's high-end electric car will have a battery with sufficient charge for a 450 km ride.

$T_{AD} = (T_{AA} \& T_{AB} \& T_{AC} \Rightarrow T_A)$: If T_{AA}, T_{AB}, and T_{AC} are true, customers perceive Tesla's high-end electric car to be equal or superior to ICE-powered high-end cars by the established brands like Porsche, Ferrari, Mercedes, and Audi.

Theory for belief T_{AC}

T_{ACA}: A battery can be developed that achieves an optimal trade-off between capacity and weight to allow for a minimum of 450 km ride.

$T_{ACB} = (T_{ACA} \Rightarrow T_{AC})$: If such a battery can be developed, Tesla's high end electric car will have a battery with sufficient charge for a 450 km ride.

FIGURE 2 Tesla's theory (with examples of sub- and sub–sub-premises)

definition of a theory as an exclusive explanation all such sub-premises are also premises of the conjecture.

To illustrate the idea of sub-premises, see Figure 2. This figure includes not only the premises for the conjecture T_Z but also the sub-premises for the premise T_A: T_{AA}, T_{AB}, T_{AC}, and T_{AD}. These sub-premises constitute a theory for T_A: they are considered necessary and jointly sufficient for T_A. Specifically, they identify the three criteria that must be fulfilled for customers to perceive Tesla's electric vehicle to be superior to competitors in the high-end segment: emissions (T_{AA}), acceleration (T_{AB}) and driving range (T_{AC}), and T_{AD} states that these are indeed jointly sufficient. Taking recursion seriously, we can, for example, zoom further into the sub-premise concerning battery charge (T_{AC}) to identify the conditions under which a battery with the required characteristics can be developed (the two beliefs shown on the lower right of Figure 2 are premises for T_{AC}).

The notion of sub-premises is important when it comes to identifying beliefs to which the strategist should target his experimentation efforts. According to Axiom 2, the premise the strategist is most willing to question and thus give up is the weakest premise. We thus define a focused experiment as follows:

> **Definition (focused experimentation).** An experiment is called focused when the strategist scrutinizes the weakest premise of the focal conjecture.

The definition of a focused experiment in combination with the notion of sub-premises is useful because many premises cannot be tested ex ante, or doing so would be prohibitively costly. Therefore, if a premise is currently not testable strategists can identify testable sub-premises by considering what they believe are the necessary conditions for that premise. Ultimately, this process of identifying sub-premises should yield a weakest premise for the conjecture that is testable at low costs.

In the Tesla example, say the strategist considers the assumption about the ability to develop batteries that allow long-distance rides but are not too heavy (T_{ACA}) to be the weakest premise of the conjecture T_Z, as it was questionable whether that would be attainable given the state of battery technology as of 2008 (throughout the article, ex ante weakest premises of the

conjecture and objections are indicated in gray color in the figures). A focused experiment would therefore be to test T_{ACA}, for example, by building a prototype, which means that T_{ACA} is testable at a cost that is a fraction of the cost of implementing the entire strategy (e.g., investing in a factory and distribution). If it turns out to be false, the strategist would (by Axiom 1) also give up beliefs that are consequences of this assumption, namely that Tesla's high-end electric vehicles will have a battery with 450 km driving range (T_{AC}), that customers perceive Tesla's cars to be equal or superior to ICE-powered cars by established manufacturers (T_A), and that Tesla would be successful in the midsize segment (T_Z), but he would keep all other beliefs.

The Tesla example also shows how belief strength depends on *both* logical relations among beliefs *and* subjective attributions of plausibility. According to Axiom 2, beliefs "inherit" their strength from sub-premises, as each belief is as strong as its weakest premise. Thus, in our example the strength of T_Z, T_A, and T_{AC} is given by the strength of T_{ACA}. Moreover, note that by definition, if premises are grouped together, the group of premises is again a premise and its strength is equal to the weakest sub-premise. In our example, the belief "both T_{AB} and T_{AC} are true" (formally written as $T_{AB}\&T_{AC}$) is also a premise of the conjecture and its strength is equal to the strength of T_{ACA}.[12]

More generally, in our framework strategists can draw inferences about whether to believe in a conjecture from evidence about its weakest premise. Therefore, the fact that a premise is as strong as its weakest sub-premise allows for an "indirect" test of some beliefs that are currently untestable without making major investments through testing their weakest premise. The notion of belief strength thus also helps to overcome a key problem of learning under uncertainty, namely that some relevant beliefs cannot be evaluated based on observations or evidence from an experiment (Gans et al., 2019). In our framework, such beliefs are as strong as their weakest premise, that is as strong as the underlying necessary assumption the strategist is most willing to question. However, while testing weakest premises allows learning about the assumptions behind a conjecture through experimentation, as we stated earlier, it is of course also possible that the theory itself is wrong or incomplete, and thus strategists may draw wrong conclusions from testing premises. To overcome this problem, our framework also accounts for the possibility to question theories by raising objections against a conjecture or its premises, which we address next.

3 | COUNTER-THEORIES AND FOCUSED EXPERIMENTATION

3.1 | Objections, counter-theories, and contradictions

So far, we have argued that to learn whether the theory for the conjecture is wrong a strategist should first become clear about the underlying premises and then test the weakest premise by putting it to an experiment. However, this is only one step. The strategist can also ask himself whether his conjecture or any of the premises would be wrong. For example, in the Tesla case one could argue that the conjecture is wrong because Tesla has a competitive disadvantage compared with traditional car manufactures as it cannot produce cars at scale to be one of the

---

[12]This is because $T_{AB}\&T_{AC}$ is true only if T_{AC} is true, and the strength of T_{AC} is the same as its weakest premise T_{ACA}, whereas T_{AB} is stronger than T_{AC}. However, the strength of T_B or T_C *does not* have to be equal to the strength of T_{ACA}, as T_{ACA} is not a premise of these beliefs.

—WILEY

largest car manufacturers, or that the absence of an infrastructure for out-of-home charging could limit the attractiveness of electric vehicles. The strategist can take such arguments he hears from others seriously, or he can also actively work to find reasons why he is wrong.

When such an argument is made, an *objection* is raised. An objection denotes the belief that a premise of the theory for the conjecture, or the conjecture itself, is wrong. Objections provide a basis for learning if they are backed by a counter-theory: a theory that explains why the conjecture or one of its premises is wrong. The idea is analogous to the creation of counter-factuals, which are alternative ways in which the past may have unfolded. Counter-factuals are based on examining the possible consequences of assumptions that certain past events (which are known) would have been different and comparing these alternative consequences to the outcome that actually occurred (e.g., Durand & Vaara, 2009; Lewis, 1973; Wason, 1960). A counter-theory, on the other hand, takes as starting point a possible future event about which a theory exists that explains under which conditions it will occur. Thus, while a counter-factual can help pinpoint the past conditions that made a critical difference to a known event, a counter-theory can help identify the present conditions that make a critical difference to a possible future event. Importantly, to improve theories by learning counter-theories, the respective premises of the theory and the counter-theory must be comparable. In particular, statements about, for example, customers, markets and technologies must agree in units of measurement, so that contradictions between theory and counter-theory can be identified. We define objection and counter-theory as follows (see also Table 1):

> **Definition (objection and counter-theory).** An objection is a belief that Z or a premise of Z is false. A counter-theory is a theory (as defined above) that implies an objection as a conjecture.

From this definition it follows that a counter-theory will also have a weakest premise and that the objection is as strong as the weakest premise of the underlying counter-theory. When asking why the conjecture could be wrong, the strategist of course considers beliefs that are logically inconsistent with beliefs he considers relevant for the focal conjecture. A counter-theory, therefore, leads to a contradiction and therefore, when exposed to a counter-theory, a strategist will need to revise his beliefs (in line with Axioms 1 and 2). For any theory, there may be multiple counter-theories, as an objection may be raised against each of multiple specific premises of the theory. We use the Tesla example and formulate two counter-theories for illustrating our results. Hereby we focus on issues that are likely to have been raised in 2008, such as those related to charging infrastructure and batteries (the mechanisms we describe apply equally well to objections concerning other issues, such as competition or costs). For both we explain the underlying logic and what is the respective weakest premise.

The first of these is the objection that T_A false (see Figure 3). The underlying counter-theory raises the point that a *dense* out-of-home infrastructure is necessary for customers to perceive Tesla's electric vehicles to be superior to ICE-powered cars (an issue that is absent in Tesla's theory). It is formulated as a set of premises that jointly imply T_L, where T_L is equivalent to the belief that "T_A is false." The counter-theory argues that unless charging infrastructure is available within 5 miles from all homes in major cities, customers will find charging to be inconvenient (T_{LA}), that such charging infrastructure will not be built (T_{LB}), and that inconvenience caused by its absence means that customers will not perceive Tesla's high-end electric car to be equal or superior to ICE-powered high-end cars by the established brands (T_{LC}). Because a counter-theory is also based on premises that are necessary and jointly sufficient conditions for

the contradictory belief (here T_L), it also has a weakest premise (which, ideally, should be testable). In addition, the strength of the objection is equal to the strength of this weakest premise (according to Axiom 2). In our example, the weakest premise for the objection about the need for charging infrastructure is T_{LAB}: the belief that customers buying the Tesla Roadster consider charging to be more inconvenient than filling up gas unless there is an out-of-home charging infrastructure that is available within 5 miles. This premise is testable, for example, by observing the behavior of Roadster customers.

The second objection is about the fact that battery weight (and thus the range of electric cars) interacts negatively with acceleration (see Figure 4). The counter-theory expresses that it cannot be both true that Tesla's high end model will have acceleration faster than high-end combustion engine models (T_{AB}) *and* a battery with 450 km range (T_{AC}). The theory behind this objection expresses the argument that a car with a battery sufficiently large for a 450 km range will be heavier than comparable ICE-powered cars (T_{MB}) and thus be so heavy that the car cannot accelerate faster than these cars (T_{MA}). As such a large battery would indeed make Tesla's cars heavier than comparable ICE-powered cars, T_{MA} is the weakest premise of this counter-theory. This premise is testable, for instance, by computing the weight that the high-end model cannot exceed to sustain acceleration higher than comparable ICE-powered
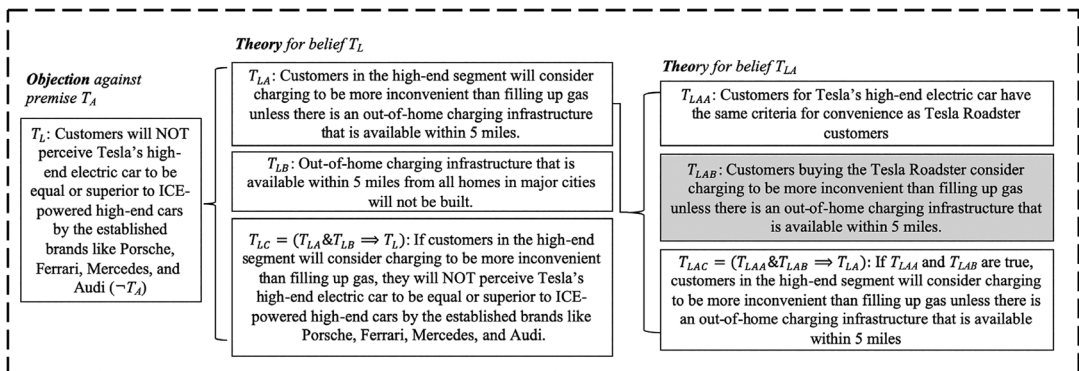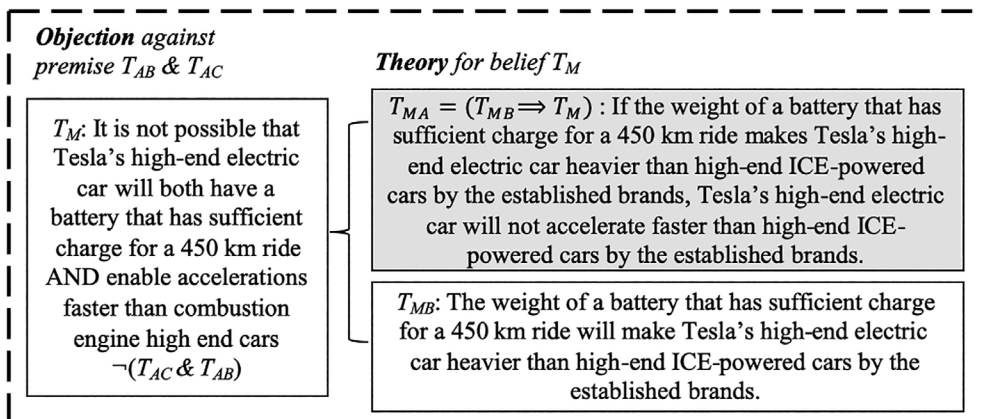

FIGURE 3 Counter-theory 1: Out-of-home infrastructure

FIGURE 4 Counter-theory 2: Battery weight and acceleration

—WILEY—

cars and then check whether a battery with a 450 km range can be built below the critical weight.

3.2 | Learning as integration of new beliefs

The question is now how counter-theories affect the strategist's beliefs regarding the conjecture. We can clarify this by conceptualizing a theory-based learning process by which counter-theories are integrated with the existing beliefs related to the conjecture. In our framework, learning then refers to a mental process by which new beliefs are added to the existing set of beliefs related to the conjecture, while at the same time contradictions are resolved as required by Axiom 1. This process involves considering both existing and new beliefs (here the counter-theories) while simultaneously thinking through implications given the logical relations between existing and new beliefs. Because logical relations also affect the relative strength of beliefs (according to Axiom 2), the relative strength of *all* beliefs may change as a result of this integration. Contradictions are resolved by removing beliefs (applying Axiom 1) based on the updated strength ordering (applying Axiom 2).

In the rest of this section, we derive formal results based on the following assumptions:

- The strategist follows Axioms 1 and 2.
- One well-defined theory for Z with weakest premise W is given.
- A set of $i \in \{1,...\}$ well-defined counter-theories with weakest premise E_i are given, where each counter-theory i implies that a premise for the conjecture, denoted S_i, in the theory is false and the counter-theory i does not create further contradictions with the theory.
- Counter-theories do not provide alternative explanations for why Z is true, or why a premise of the strategist's initial theory for Z is true, and they do not contradict each other.
- Theory and counter-theories have been formulated in a way that the relevant premises and sub-premises have been identified (in particular, an ex-ante testable weakest premise has been identified for the theory and for each of the counter-theories).
- The beliefs relevant for the conjecture are a result of integrating the theory and the counter-theories (no other beliefs are given).

Making these assumptions, we will derive two formal results about the mental process of learning by integrating counter-theories and resolving contradictions (Propositions 1 and 2) and one formal result about focused experimentation (Proposition 3).

3.3 | Maintaining the belief in the conjecture

The strategist wants to decide whether or not to continue investing with the aim of creating a world in which the conjecture is true. Our first result specifies the conditions under which a strategist who is exposed to one or more counter-theories will continue to believe that the conjecture is true. In brief, he should continue to believe in the conjecture if he considers the weakest premise of his theory to be stronger than the weakest premise of the strongest objection. More precisely, assume the strategist's beliefs contain a theory for the focal conjecture Z, which he thus believes to be true. The theory's weakest premise is denoted by W. The strategist now integrates one or more counter-theories into his beliefs that each contradict a premise in

the theory. Specifically, each counter-theory i implies that a premise in the theory S_i is false, with $i \in \{1, ...\}$. Each counter-theory thus contains as conjecture the belief S_i is false (formally $\neg S_i$, where $\neg$ denotes negation) and a set of premises for $\neg S_i$. In the example, the counter-theories contradict T_A and $T_{AB}\&T_{AC}$, which thus correspond to the respective S_i's in our notation.

The beliefs $\neg S_i$ are ordered by relative strength so that $\neg S_1$ is the strongest objection, $\neg S_2$ the second-strongest and so on. Each counter-theory may contain multiple premises, so denote the weakest premise of $\neg S_i$ by E_i (recall that by Axiom 2 the strength of $\neg S_i$ is equal to the strength of E_i). As $\neg S_1$ is the strongest objection, E_1 is stronger than any E_i with $i > 1$. The following proposition then follows from Axioms 1 and 2, stating the conditions under which the strategist—after hearing one or more counter-theories—continues to believe that the conjecture Z is true .

> **Proposition 1. (maintaining belief in the conjecture)**: A strategist follows Axiom 1 and 2, and his beliefs related to the conjecture Z imply that it is true. The strategist now learns one or more counter-theories related to thinking about the conjecture Z. Then, if E_1 is weaker than W, the strategist will continue to believe that Z is true.

Proof: See Appendix A.

To get an intuition for Proposition 1, see the left side of Figure 5. Per Axiom 2, it suffices to remove the weakest premises of each of the counter-theories to resolve all contradictions among theory and counter-theories. The strength of the strongest objection is given by its weakest premise E_1. If E_1 is weaker than the weakest premise of the theory W, then E_1 and not W is given up, as weaker beliefs are given up first to resolve contradictions (Axiom 2). As E_1 is stronger
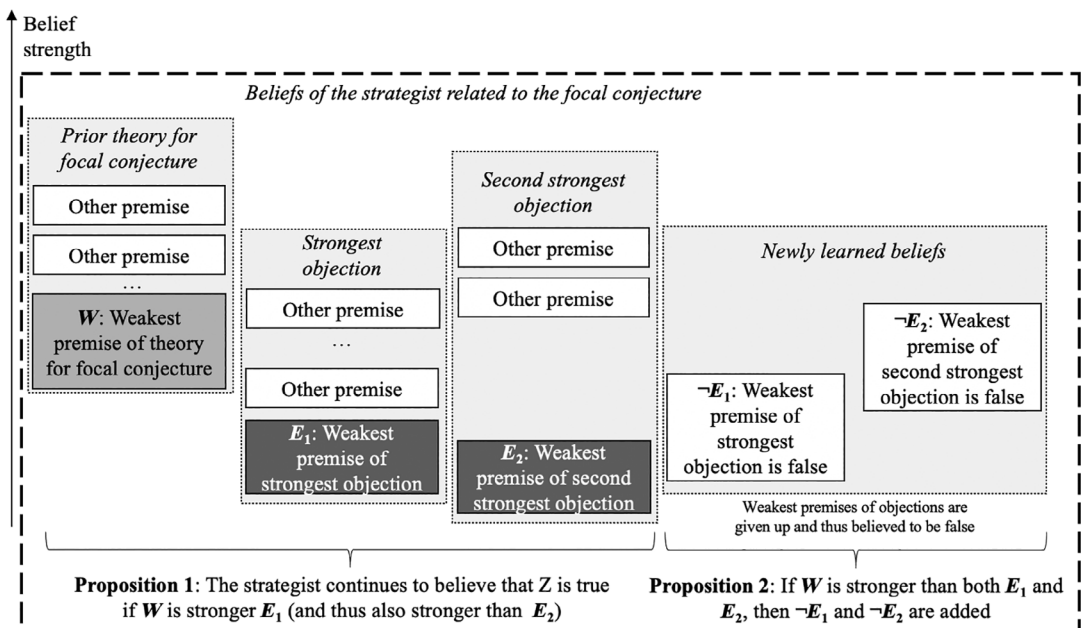

FIGURE 5 Illustration of Propositions 1 and 2

than E_2 (the weakest premise of the second strongest objection), W is also stronger than E_2 and thus, the contradiction caused by the second strongest objection is also resolved in favor of the theory: W is maintained and E_2 is given up. As W is stronger than the strongest objection, the weakest premises of all objections are discarded and W is maintained.

To illustrate Proposition 1 with our Tesla example, recall that the weakest premise of the theory was T_{ACA}: the assumption that it is feasible to develop a battery that allows for sufficient charge for a 450 km ride. In the notation of Proposition 1, T_{ACA} thus corresponds to W. According to Proposition 1, W has to be compared with E_1, the weakest premise of the strongest objection. To determine what E_1 is, we have to compare the weakest premises of the respective objections with each other, as the strength of objections is given by the strength of their weakest premises.

Say the strategist considers it more plausible that Tesla Roadster customers consider charging to be more inconvenient than filling up gas unless there is an out-of-home charging infrastructure that is available within 5 miles (T_{LAB}) than that the weight of batteries will prevent superior acceleration (T_{MA}) and are thus most concerned that T_{LAB} may actually be true. Thus, T_{LAB} corresponds to E_1 in the notation of the proposition. Thus, after hearing the counter-theories behind the two objections, in order to decide whether to (at least tentatively) continue to believe that Tesla will sell enough electric cars in the midsize segment to be among the five largest players in that segment (T_Z), the strategist need only consider whether or not he believes that T_{ACA} is stronger than T_{LAB}—the strength of W needs to be compared with the strength of E_1. Then, if W is considered to be stronger than E_1 according to Proposition 1, the strategist will tentatively reject both objections. That is, he believes that (given all his other assumptions) the absence of a dense charging infrastructure and the large weight of batteries will not prevent Tesla from becoming one of the large players.

3.4 | Hidden premises and learning from counter-theories

In this and the following subsection we assume that the strategist's mental comparison resulted in the conclusion that W is considered to be stronger than E_1. Thus, we analyze the case in which all objections were rejected and the strategist thus continues to believe that the conjecture is true. Recall that by giving up the weakest premises of objections, the strategist also rejects the objections. But even if all objections are rejected, the strategist may learn from counter-theories and thus his beliefs can change in the process of considering objections. This is because Axiom 2 requires that strategists resolve contradictions in a way that keeps as many beliefs as possible. In particular, the strategists also keeps all beliefs of a counter-theory, except for their weakest premises, as giving up the weakest premises suffices to resolve the contradictions with the theory. Because the strategist draws all logical consequences from his beliefs, this implies that he will also believe that the weakest premises of the counter-theories must be false. Importantly, also the strategist's *strategic* beliefs can change. That is, a newly learned belief can become a hidden *premise*: giving up this belief would force the strategist to also give up his belief in Z. We explain this in more detail as follows.

As the other beliefs from the counter-theories do not lead to any contradictions after the counter-theories' weakest premises have been removed, the beliefs related to the conjecture will include all these other beliefs. In addition, the strategist also learns that the weakest premise of each counter-theory *must be false*. This is because according to Axiom 1 the strategist will draw all logical conclusions from his beliefs, which now also include the other beliefs from the counter-theories except the respective weakest premises. Now recall that a counter-theory implies that a premise of the theory S_i is false. For each objection, the fact that it was rejected

thus means that the strategist still believes that S_i is true. Therefore, the belief that S_i is true together with the other, newly accepted beliefs from counter-theory i imply that E_i is false (if E_i would hypothetically be added back the strategist would again conclude that S_i is false and the contradiction would reappear).[13]

Learning that the weakest premise of a counter-theory is false can provide the strategist with important insights about contingencies that would otherwise have been overlooked. To illustrate, in the Tesla example the strongest objection (see Figure 3) is that T_A is false (T_A corresponds to S_1): customers will *not* perceive Tesla's high-end electric cars as superior. The weakest premise of the counter-theory is T_{LAB}: that Tesla Roadster customers find charging inconvenient if there is no dense out-of-home charging infrastructure (which thus corresponds to E_1). If the objection is rejected, T_{LAB} is given up whereas the strategist will believe that all other beliefs from the counter-theory are true. For example, the strategist now explicitly believes that convenient charging is necessary for high-end customers to perceive Tesla's cars as equal or superior (which is expressed in T_{LC}). Because he also continues to believe that T_A is true (as the objection was rejected), he now must also believe that T_{LAB} is false. The strategist thus concludes that while convenient charging is necessary (T_{LC} is true) the absence of a dense charging infrastructure in itself *does not* make charging inconvenient (T_{LAB} is false). Of course, the strategist does not believe that a charging infrastructure is not necessary. But he now believes that charging can be made convenient *even if* there is no dense charging infrastructure (perhaps through at-home charging and a smaller number of public charging stations in selected places).

But there is an additional, subtle consequence of considering but rejecting objections: The belief that E_i is false can (but not necessarily does) become a *hidden premise* for the conjecture Z. That is, it could become a necessary condition for the conjecture to be true, and if it were not true the conjecture could not be true given the strategists' other beliefs. Formally, the belief that E_i is false will become a premise for Z if the strategist, given his current belief strength ordering, cannot start believing E_i without also giving up the belief in the conjecture Z. The intuition behind this result is based on the following thought experiment: Consider that after having concluded that E_i is false the strategist obtains evidence that E_i is true after all, which also means that he has to consider it as very strong. This would lead to a contradiction, as the strategist would conclude that S_i is false. So the strategist would need to resolve a contradiction again. Whether in that case the strategist would give up his belief in Z as a result of the thought experiment that E_i is true depends on which belief would be given up: According to Axiom 2, weaker beliefs are given up first and only as few beliefs as necessary are given up to restore consistency. If none of the other beliefs from counter-theory i are weaker than S_i, then S_i and, by implication, Z would be given up. If, on the other hand, at least one of the other beliefs from counter-theory i is weaker than S_i, then the weakest of these would be given up and the strategist would continue to believe that S_i is true (and that Z is true) even if he would believe that E_i is true. Therefore, whether or not the belief that E_i is false becomes a premise for Z depends on the strength of the other beliefs from the counter-theory relative to S_i.[14]

---

[13]In logical terms, this result is a consequence of the modus tollens. As E_1 is a premise of the objection, removing this premise, but keeping all other beliefs of the counter-theory, implies that the strategist believes that $E_1 \Rightarrow \neg S_1$. $\neg S_1$, by definition, is a belief that contradicts the strategist's theory for the conjecture. But the strategist's theory is fully maintained! Therefore the strategist believes that $\neg S_1$ is false. From $E_1 \Rightarrow \neg S_1$ it then follows that E_1 must be false.

[14]More formally, if the strategist considers the belief that E_i is false to be weaker than S_i, he currently believes that E_i is false, but he is more willing to question this belief than the belief in S_i. Technically, this means that $E_i \Rightarrow \neg S_i$ is weaker than S_i. In such cases, the belief that E_i is false does not become a premise of Z. See proof of Proposition 2 in the Appendix.

In the Tesla example, this means that the belief that T_{LAB} is false becomes a premise if all other beliefs from the first counter-theory are stronger than T_A. For example, the belief that T_{LAB} is false will not become a premise if the strategist considers the belief that customers for Tesla's high-end electric cars have the same criteria for convenience as Tesla Roadster customers to be weaker than T_A. Before explaining in more detail what the implications of learning hidden premises are, we summarize our formal results as follows:

> **Proposition 2. (hidden premises):** Consider a strategist who learned one or more counter-theories and considers W to be stronger than E_1. Then the strategist will believe that E_i is false for all i. If the strategist considers all of the remaining premises of counter-theory i (i.e., all premises except E_i) to be at least as strong as S_i, then the belief "E_i is false" will become a premise for Z.

Proof: See Appendix A.

Learning by mentally integrating counter-theories helps the strategist to become clearer about his assumptions even if the objections are rejected, and it can help him decide which experiments to run. If the belief that E_i is false becomes a premise, then it will be a strategic belief. That is, it becomes a part of the strategist's *exclusive* explanation for the conjecture and not simply another belief that is relevant for thinking about the conjecture. As a consequence, given his other beliefs the strategist will be believe that that Z is true *only if* E_i is false.

We illustrate this with the second objection, which states that $T_{AB}\&T_{AC}$ is false: it cannot be both true that Tesla's high end model will have sufficient charge for a 450 km ride and acceleration faster than high-end combustion engine models. When this objection is rejected, the strategist believes that T_{MA} is false: a battery with enough charge for a 450 km ride will not make the vehicle too heavy for it to accelerate faster than competing ICE-powered cars. The belief that T_{MA} is false will become a strategic belief if T_{MB} (the belief that batteries will make Tesla's high-end electric cars heavier than competing ICE-powered cars) is at least as strong as the belief that $T_{AB}\&T_{AC}$ is true. Because T_{MB} is strong (large batteries do add a lot of weight to cars), the belief that T_{MA} is false becomes a strategic belief (a hidden premise): After rejecting the counter-theory, the strategist believes that given his assumptions only if T_{MA} is indeed false is it possible that T_Z will be true. In other words, the strategist now believes that the success of Tesla's strategy critically depends on the ability to make Tesla's high-end vehicles accelerate faster than competing ICE-powered cars despite them being heavier due to the weight of batteries. This conclusion is an outcome of the strategist's mental process of thinking about objections to his theory, and as we show next, it may have consequences for the focused experiment the strategist should run.

To summarize, even if objections are rejected the strategist can learn from counter-theories in two ways. He can learn that that given his assumptions the conjecture will become true *even if* a particular contingency is met (e.g., charging can be made convenient even if there is no dense charging infrastructure). And he can learn that the conjecture will become true *only if* a particular contingency is met (e.g., as batteries make electric vehicles heavier, customers will prefer the high-end model only if it can still accelerate faster than ICE-powered high-end models despite the added weight). Thus, learning from counter-theories enables the strategist to change his beliefs about important and even critical contingencies.

3.5 | Counter-theories and focused experimentation

Proposition 1 clarified the conditions under which the strategist has reasons to continue believing in the conjecture when exposed to one or more counter-theories, and Proposition 2 stated what is being learned from counter-theories if they are rejected. What is missing, however, are the implications for the initially mentioned attempt to refute one's belief in the conjecture by experiment. Above we defined focused experimentation as the process of testing weakest premises with the purpose of potentially refuting them. As we show next, when a counter-theory has been rejected (according to Proposition 1) and new beliefs have been learned (according to Proposition 2), the premises that should be targeted for focused experimentation can also change.

Recall that when an objection is rejected, the strategist will believe that E_i is false and, according to Proposition 2, this belief will become a *premise* for Z if the remaining beliefs of counter-theory i (each of the beliefs that composed the counter-theory except for E_i) are at least as strong as S_i (the premise in the theory that was contradicted by the counter-theory). Now there is one case in which the belief that E_i is false will become a target for a focused experiment. This happens when the counter-theory of which E_i was the weakest premise was targeted at a premise S_i of the theory that had W as its weakest premise and the belief that E_i is false becomes a premise for Z after rejecting the counter-theory (according to the second part of Proposition 2). In that case, the strength of the belief that E_i is false will be the same as the strength of W (if E_i would be added again, the strategist would have to give up W, as all other beliefs in the counter-theory i are at least as strong as W). That implies that the strategist is as willing to question whether E_i is false as he is willing to question W. Thus, it is a focused experiment to test both E_i and W. This leads to our third formal result:

> **Proposition 3. (focused experimentation)**: Consider a strategist with a theory who learned one or more counter-theories and considers W to be stronger than E_1. If a counter-theory i contradicts a premise S_i in the theory that has W as its weakest premise and the belief that E_i is false is a premise of Z, then it is a focused experiment to scrutinize both the belief that E_i is false and W.[15] If no counter-theory contradicts a premise S_i in the theory that has W as its weakest premise, it is a focused experiment to scrutinize W.

Proof: See Appendix A.

Proposition 3 states the condition under which a hidden premise becomes a target for focused experimentation. To illustrate Proposition 3, note that the second objection was targeted at T_A, and T_A had T_{ACA} as its weakest premise, which is also the weakest premise for T_Z (formally, T_{ACA} corresponds to W in the Tesla example). As we showed above, after rejecting the second objection the belief T_{MA} is false becomes a hidden premise. Therefore, according to Proposition 3, in addition to T_{ACA}, it will be a focused experiment to test the belief that T_{MA} is false. In other words, the strategist believes that for Tesla to be successful (for T_Z to become true) it is not enough to develop a battery that allows for a 450 km ride but this battery also needs to be below a critical weight so that an acceleration faster than established ICE-powered high-end cars can be reached. A focused experiment should thus not only focus on battery development but also on studying whether the development of an electric vehicle with such a battery is

---

[15]If there are multiple counter-theories that each contradict a premise T_k in the theory that is premised on W, and all $\neg E_k$s are premises for Z, then it is a focused experiment to scrutinize $\neg E_{k1} \wedge \neg E_{k2} \wedge ... \wedge W$.

feasible.[16] The outcome of such an experiment would inform the strategist not only about the feasibility of developing such a car but –because beliefs are logically linked– the strategist would also infer whether customers will perceive Tesla's high-end electric vehicle as equal or superior to competing ICE-powered high-end cars and, ultimately, whether he can maintain his belief in Z.

4 | ALTERNATIVE THEORIES

Counter-theories can help to test and improve the strategist's theory by raising objections. However, given the nature of uncertainty, there may also realistically be alternative, plausible views about how the conjecture may be reached (by assuming exclusivity we have so far ruled out that the strategist considers such alternatives). To illustrate, our Tesla example was based on the logic that success in the high-end segment would serve as a stepping stone for success in the midsize segment (encapsulated in T_C: customers in the midsize segment will have a preference for Tesla's electric vehicles if demand for Tesla's high-end cars is high, and T_D: that this is sufficient for Tesla to become one of the five largest players in the midsize segment). An alternative point of view for the focal conjecture concerning success in the midsize segment could be based on a logic of disruptive innovation (Christensen, 1997), where electric vehicles would initially be introduced in the low-end segment and then success in the low-end segment would spur sales among customers in the midsize segment.

Recall that a theory is an *exclusive* explanation that formulates a firm-specific point of view about the conditions under which a focal conjecture will be reached (Felin & Zenger, 2017). Therefore, alternative theories by definition contain elements that contradict the prior theory. In our framework, these viewpoints can be formulated as alternative theories that include a counter-theory and that contain elements of an alternative explanation for why the conjecture may be true. Any alternative theory at least questions the necessity of some assumptions in the original theory. Practically, however, the alternative theory can often be traced to questioning the core mechanism proposed in the original theory. In our example, the original theory is based on the idea of Tesla's initial success in the high-end segment while the alternative theory sketched above is based on the idea of Tesla's initial success in the low-end segment. Thus, the alternative theory contradicts either T_C or T_D. An argument against T_C is the following (see top part of Figure 6): The primary buying criterion for customers in the midsize segment is low prices (T_{CX}), which implies that high demand for Tesla's high-end cars does not necessarily lead to customers in the midsize segment preferring Tesla's electric vehicles (T_{CY}) and thus implies that T_C is false.[17] On the other hand, the belief that the primary buying criterion for customers in the midsize segment is low prices (T_{CX}) is also a premise for an alternative explanation based on initial success in the low-end segment (that is, the counter-theory that leads to the contradiction and the alternative explanation share the same premise; see Figure 6, bottom for this alternative theory). Specifically, the beliefs T_N and T_O formulate a logic by which an electric vehicle that is cheaper than comparable ICE-powered cars in the low-end segment but has acceptable

---

[16]Note that the first objection also contradicts a premise of the theory that has W as its weakest premise, as W is the weakest premise for T_A. However, we assume that the strategists considers the second strongest premise (T_{LAA}) in the first objection to be weaker than W. Thus, the first objection does not lead to a change of the focused experiment.

[17]Customers' decision to switch to electric cars could also be determined by a function of price and prestige, which could also be formulated as a premise. For simplicity, we focus on the either/or case in the example.
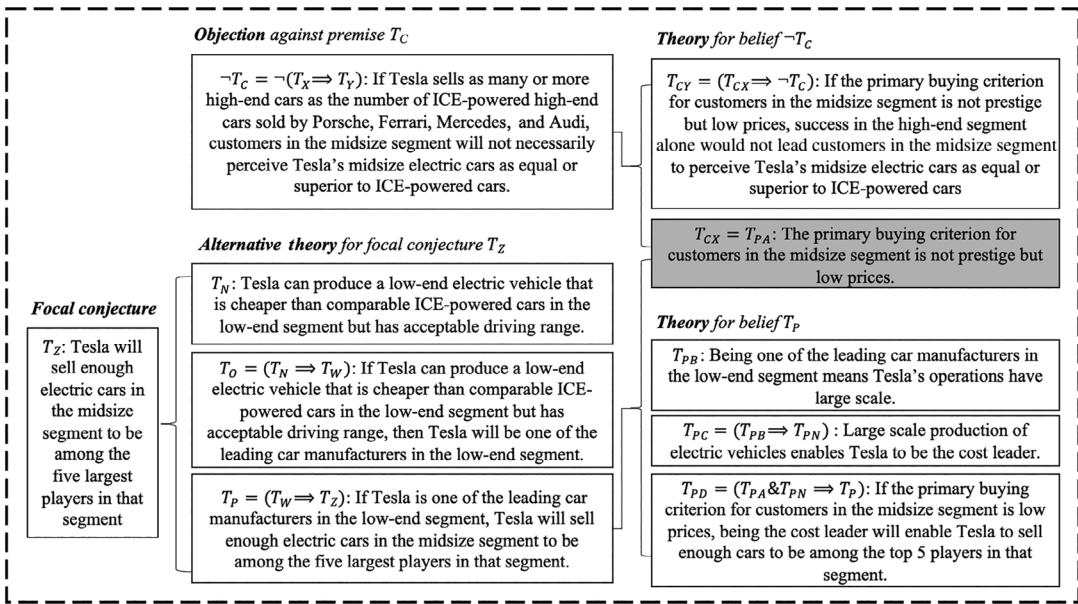
Objection *against premise* T_C

$\neg T_C = \neg (T_X \Longrightarrow T_Y)$: If Tesla sells as many or more high-end cars as the number of ICE-powered high-end cars sold by Porsche, Ferrari, Mercedes, and Audi, customers in the midsize segment will not necessarily perceive Tesla's midsize electric cars as equal or superior to ICE-powered cars.

Theory *for belief* $\neg T_C$

$T_{CY} = (T_{CX} \Longrightarrow \neg T_C)$: If the primary buying criterion for customers in the midsize segment is not prestige but low prices, success in the high-end segment alone would not lead customers in the midsize segment to perceive Tesla's midsize electric cars as equal or superior to ICE-powered cars

$T_{CX} = T_{PA}$: The primary buying criterion for customers in the midsize segment is not prestige but low prices.

Alternative theory *for focal conjecture* T_Z

T_N: Tesla can produce a low-end electric vehicle that is cheaper than comparable ICE-powered cars in the low-end segment but has acceptable driving range.

$T_O = (T_N \Longrightarrow T_W)$: If Tesla can produce a low-end electric vehicle that is cheaper than comparable ICE-powered cars in the low-end segment but has acceptable driving range, then Tesla will be one of the leading car manufacturers in the low-end segment.

$T_P = (T_W \Longrightarrow T_Z)$: If Tesla is one of the leading car manufacturers in the low-end segment, Tesla will sell enough electric cars in the midsize segment to be among the five largest players in that segment.

Theory *for belief* T_P

T_{PB}: Being one of the leading car manufacturers in the low-end segment means Tesla's operations have large scale.

$T_{PC} = (T_{PB} \Longrightarrow T_{PN})$: Large scale production of electric vehicles enables Tesla to be the cost leader.

$T_{PD} = (T_{PA} \& T_{PN} \Longrightarrow T_P)$: If the primary buying criterion for customers in the midsize segment is low prices, being the cost leader will enable Tesla to sell enough cars to be among the top 5 players in that segment.

Focal conjecture

T_Z: Tesla will sell enough electric cars in the midsize segment to be among the five largest players in that segment

FIGURE 6 Alternative theory: Disruptive, low-end strategy

driving range leads to Tesla becoming one of the leading car manufacturers in the low-end segment. If this is achieved, according to the alternative theory Tesla will become one of the top five players in the midsize segment (T_Z) because in addition to customers in the midsize segment preferring vehicles with lower prices ($T_{PA} = T_{CX}$) the low-end strategy also enables scale to be built (T_{PB}), thereby achieving cost leadership in the midsize segment (T_{PC}).

In our example, there is one key contradiction between prior theory and alternative theory, since the objection raised concerning the primary buying criterion for customers in the midsize segment being low prices (T_{CX}) leads to the belief that T_C is false, whereas T_C is believed to be true in the prior theory. Note that this difference in the two theories has repercussions on other beliefs, for example, what kind of capabilities Tesla needs to generate enough demand in the volume segment to be a top five player. The weakest premise of the objection against T_C is T_{CX}, as the strategist is sure that T_{CY} is true (namely that if it is true that customers in the midsize segment primarily care about price then the prior theory would not work), but is less sure whether customers in the midsize segment really do primarily care about price. Therefore the alternative theory would be rejected if T_{CX} would be weaker than T_C. We now examine what the strategist would then learn from hearing about the alternative theory.

Because the alternative theory includes an objection in the form of a counter-theory that we analyzed above, as a result of rejecting the alternative theory the strategist will believe that the weakest premise of the objection to the original theory posed by the alternative theory is false. In our example, the strategist now believes that customers in the midsize segment *do not* primarily switch to electric vehicles out of price considerations. In fact, in the example the belief that T_{CX} is false would even become a *premise* (and thus a strategic belief), because the belief that the top-down logic does not work if midsize customers primarily switch to electric vehicles out of price considerations is strong. Thus, the strategist will also need to pay special attention to the purchase criteria of customers in the midsize segment.

—WILEY⊥

If, on the other hand, the strategist would consider T_{CX} to be stronger than T_C, he may accept the alternative theory for T_Z. However, he would then need to expose the alternative theory to the same scrutiny and discipline as he did with the ex-ante theory. He would need to order the beliefs of the alternative explanation with respect to their strength and consider objections against the alternative theory. For example, the alternative theory ignores competition and assumes that large scale production of electric vehicles alone will make Tesla the cost leader (T_{PC}). Once objections and underlying counter-theories have been identified, the strategist could apply our three propositions again, to check if the alternative theory withstands objections, to learn from them (perhaps identify hidden premises as newly identified necessary assumptions for becoming the industry cost leader), and to run focused experiments to test weakest premises.

5 | DISCUSSION AND CONCLUSIONS

In this article, we introduced a framework with which strategists can systematically evaluate and improve the theories behind their envisioned future that is core to their strategy. We derived three propositions that help strategists to systematically learn from objections, to identify hidden premises, and to test the weakest assumptions behind their envisioned future. In the following, we discuss the implications of our framework and results.

5.1 | Theory-based learning under uncertainty

A key difference between our framework and other formal approaches to learning under uncertainty (e.g., learning by updating prior beliefs using Bayes' rule) is that in our framework strategists not only learn whether or not they should continue working toward their envisioned future based on making observations but they also learn from arguments formulated as counter-theories even if objections are rejected (e.g., by identifying hidden premises). This mental learning process has particular value under uncertainty, when the critical contingencies that will affect outcomes are not yet fully known. As a normative framework, the Bayesian learning approach instructs strategists to use evidence to update the probability of success of an idea or a way to implement it and abandon the idea or the implementation path if the posterior probability is below a threshold value (e.g., Camuffo et al., 2020; Kerr, Nanda, & Rhodes-Kropf, 2014). However, from an ex ante perspective, whether or not an implementation path will eventually work and enable a strategist to reach their strategic intent—such as being a number 5 player in the US car industry using electric instead of ICE-powered engines—depends on many contingencies, and the "true" mapping of currently available evidence to eventual outcomes is unknown. Theories suggest plausible mappings from evidence today to outcomes tomorrow. Using theories thus enables learning, as theories define possibly relevant contingencies (the space of variables) and suggest plausible relations between them (Griffiths & Tenenbaum, 2009). Theories as spaces of variables and plausible relations constitute priors, which are assumed to be given in prior formal work on entrepreneurial learning (Agrawal et al., 2021; Gans et al., 2019). In other words, while these studies assume existing priors, our results identify possibilities how priors can be improved from an ex ante perspective, namely by attempts to falsify them on a logical basis and by learning from counter-theories and experiments that are focused on weakest premises. Thus, our results show how firms, when faced

with uncertainty, can meaningfully learn about the best way to reach their strategic intent and increase the chances of success by forming and revising theories.

Our framework also helps to overcome some conceptual confusion in the prior literature. In the prior literature, a failure to abandon an idea in the face of contradictory evidence is often taken as indication of a cognitive bias, for example, due to overoptimism (e.g., Camerer & Lovallo, 1999; Kahneman & Lovallo, 1993). However, if the mapping from evidence today to outcomes tomorrow is not yet known, it is not even clear what contradictory evidence actually is! For instance, from the perspective of 2008 it is impossible to know in a statistical sense whether or not sales in the high-end segment for electric cars is conducive to sales in the midsize segment. So it is, from an ex ante perspective, not clear whether a lack of demand for Tesla's high-end model would indicate that there is no demand for Tesla's electric cars in the volume segment. In our approach, the idea is to try to gather contradictory evidence about weakest premises, because that would be enough to falsify theories. This does not necessarily mean that the conjecture is wrong but is first of all an invitation to continue thinking about conditions under which the conjecture will perhaps nonetheless materialize. Using our framework a strategist would only conclude that an idea might not be worth pursuing if he cannot find another strong theory for the conjecture (that is, a theory that has a strong weakest premise) and that would be stronger than the objections raised against it. Thus, our approach suggests that persistence should not be conflated with overoptimism but instead as continued search for a strong theory while being open for the possibility that one cannot find any plausible path toward making the conjecture true.

Finally, our approach closes an important gap in normative theories of entrepreneurial learning. Some assumptions (such as T_C or T_D in our Tesla example) cannot be tested without substantive investments. While prior contributions have suggested that this creates irreducible uncertainty (Gans et al., 2019), our approach suggests an indirect way of testing them as premises can be explained by sub-theories with a weakest premise that is testable at a low cost. While these sub-theories, of course, can be wrong, they can be improved by raising objections and learning from the counter-theories behind them. Thus, we propose a learning mechanism that, while not eliminating the uncertainty around major investments, at least provides strategists with a refined understanding of why they pursue major investments, and the conditions under which they will be successful.

5.2 | Linking theory-based learning and experimentation to the strategy process

Our work also contributes to a small but growing literature that argues for a normative approach to strategy formulation based on identifying processes and frameworks that are prescriptively designed, applied and taught in order to lead to better strategies and better decisions (e.g., Grandori, 1984; Nickerson & Argyres, 2018; Rindova & Martins, 2021). In this literature, identifying problems and designing and implementing solutions to solve these problems are the most central activities in the strategy process (Nickerson et al., 2012; Nickerson & Zenger, 2004). Our work complements and extends this literature by offering a normative framework for the systematic evaluation of theories (Popper, 1959) by raising objections and testing assumptions through experiments.

Formulating assumptions and testing them through experiments has been advocated by practitioner approaches like the lean startup (Ries, 2011). However, these approaches have been criticized for being myopic and focusing on incremental advances due to a lack of attention to

the underlying theories or vision entrepreneurs may have (Felin, Gambardella, Stern, & Zenger, 2020; Thiel, 2014). In our framework, on the other hand, theories and weakest premises provide clear direction for which experiments strategists should perform. More fundamentally, the logical linkages among beliefs allow strategists to make inferences about what the outcome of an experiment means for their other beliefs. Thus theories relate what is observable today to consequences in the future, which allows for making inferences from early-stage experiments about assumptions that cannot be tested without major investments. This means that our framework mitigates the paradox of entrepreneurship noted by Gans et al. (2019) by enabling strategists to indirectly test assumptions that are central to their strategy but that are at present untestable (such as Tesla's assumption that initial success in the high-end segment opens up a path for electric vehicles in the volume segment).

As a framework for the systematic evaluation of theories our approach is embedded in a wider strategy process that also includes other activities for finding, formulating and solving problems that go beyond our framework. Specifically, our mathematical results take conjectures, theories and counter-theories as inputs. We have suggested that theories can be created by starting from a conjecture that embodies a strategic intent (Prahalad & Hamel, 1994) or a shaping intention (Rindova & Martins, 2021) and identifying the premises as the necessary and jointly sufficient conditions under which the conjecture will materialize. The formulation of a conjecture and an underlying theory, however, is an at least partly creative process that requires imagination (e.g., Wiltbank, Dew, Read, & Sarasvathy, 2006). In this sense, our framework can be seen as a tool for making sure strategists couple imagination with discipline in their reasoning about an imagined future (cf. Weick, 1989). The theory creation process can also be supported by established frameworks like scenario planning (Schoemaker, 1993), Five Forces (Porter, 1980), or hypothesis trees (Davis, Keeling, Schreier, & Williams, 2007) that allow for both identifying potentially relevant issues and going deeper into particular topics and which are thus complementary to our framework.

As a tool for disciplined reasoning about a focal conjecture, our framework can be applied whenever a strategist wants to evaluate a given theory. This also means that it can be applied again whenever new objections must be considered, new alternative theories are to be evaluated, or the outcome of an experiment creates new facts to be incorporated into the theory, as long as the strategist starts with a complete theory that is formulated in line with our axioms. When a theory has been falsified (through either an objection that is accepted, or an experiment that falsifies a premise), the strategist can use the remaining beliefs as building blocks for formulating an alternative theory for reaching their strategic intent. Importantly, any new or changed theory also should be evaluated by checking if the theory withstands objections and by performing focused experiments and learning additional constraints from counter-theories.

5.3 | Further research, limitations and extensions, and conclusions

One avenue for further work is to link our model with game-theoretic models. For example, our model could serve as a microfoundational mechanism for strategic cognition in value capture theory, in particular when strategists differ in their views of a situation and deal with unknown unknowns (Bryan et al., 2021; Cappelli & Chatain, 2021). Here, different theories of different agents could provide a rationale for why agents envision different possibilities to create value together, or why they envision different sets of competitors. In addition, it appears worthwhile to study learning and revising theories not just from the perspective of an individual

strategist but also in the context of a group of strategists who have different and potentially overlapping or conflicting theories. This naturally leads to questions such as the conditions under which a group of strategists will agree on a conjecture, or under which conditions it is possible to persuade others of one's own theory (cf. Kaplan, 2008). In particular, our results can also form the basis for studies that examine how to strategically persuade others through arguments about the relation of statements rather than through providing signals or evidence, as in the literature on Bayesian persuasion (e.g., Kamenica & Gentzkow, 2011). Our results should also be useful to understand how arguments made by influential CEOs in speeches can alter the collective perception of critical contingencies that influence, for example, firm valuation.

Another avenue for further work is to link theory-based learning with statistical learning. Validating or falsifying theories using statistics from an ex-ante perspective is of course impossible: For some relations in a theory, data will only be available in the future, but theories can provide useful priors for statistical learning at a later stage (Ehrig & Foss, 2022; Ehrig & Schmidt, 2021; Griffiths & Tenenbaum, 2009). Moreover, strategists can employ a staged approach and use Bayesian learning, for example, to improve products in a test market (Zellweger & Zenger, 2021) or prototypes (Ehrig, Knudsen, & Rauh, 2022). Theories then inform strategists how to make inferences from the learning outcomes to the next stages, such as scaling to global markets or developing a market-ready product. In addition, our framework can also be linked to the process of learning about the value of real options (Trigeorgis & Reuer, 2017).

Finally, our work offers opportunities to investigate strategic reasoning skills empirically. We argued that our framework is normative and thus helps strategists who use it to come up with better theories and thus make better decisions. The aforementioned findings by Camuffo et al. (2020) also lend credence to the view that strategists and entrepreneurs fare better when they use a scientific approach to evaluate novel opportunities. However, it is an open question to what extent strategists in fact use such principles. There is anecdotal evidence for this: Elon Musk famously uses what he calls "thinking from first principles" (Meija, 2018), which is essentially using the core idea behind our framework: identifying the premises for one's conjecture and then testing them. Future work could study more systematically how theories are formulated and revised by strategists and entrepreneurs and how doing so is linked with success.

ORCID

Jens Schmidt https://orcid.org/0000-0002-3266-6240

REFERENCES

Agrawal, A., Gans, J. S., & Stern, S. (2021). Enabling entrepreneurial choice. *Management Science*, *67*(9), 5510–5524.

Baer, M., Dirks, K. T., & Nickerson, J. A. (2013). Microfoundations of strategic problem formulation. *Strategic Management Journal*, *34*(2), 197–214.

Bryan, K. A. D., Ryall, M., & Schipper, B. C. (2021). Value capture in the face of known and unknown unknowns. *Strategy Science*.

Camerer, C. F., & Lovallo, D. (1999). Overconfidence and excess entry: An experimental approach. *American Economic Review*, *89*, 306–319.

Camuffo, A., Cordova, A., Gambardella, A., & Spina, C. (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science*, *66*(2), 564–586.

Cappelli, V., & Chatain, O. (2021). Generalized biform games. Working paper.

Casadesus-Masanell, R., & Zhu, F. (2013). Business model innovation and competitive imitation: The case of sponsor-based business models. *Strategic Management Journal*, *34*(4), 464–482.

Christensen, C. M. (1997). *The Innovator's dilemma: When new technologies cause great firms to fail*. Boston, MA: Harvard Business School Press.

Davis, I., Keeling, D., Schreier, P., & Williams, A. (2007). The McKinsey approach to problem-solving. McKinsey staff paper: No. 66.

Durand, R., & Vaara, E. (2009). Causation, counterfactuals, and competitive advantage. *Strategic Management Journal*, *30*(12), 1245–1264.

Ehrig, T., & Schmidt, J. (2021). Making biased but better predictions: The trade-offs strategists face when they learn and use heuristics. *Strategic Organization*, *19*(2), 263–284.

Ehrig, T., & Foss, N. (2022). The strategist's imagination and rationality: Learning in open-ended worlds. Working Paper. Copenhagen Business School.

Ehrig, T., Knudsen, T. & Rauh, J. (2022). Finding needles in haystacks: A new perspective on the search for breakthrough innovations. Working paper. Copenhagen Business School.

Felin, T., Gambardella, A., Stern, S., & Zenger, T. (2020). Lean startup and the business model: Experimentation revisited. *Long Range Planning*, *53*(4), 101889.

Felin, T., & Zenger, T. (2009). Entrepreneurs as theorists: On the origins of collective beliefs and novel strategies. *Strategic Entrepreneurship Journal*, *3*, 127–146.

Felin, T., & Zenger, T. (2017). The theory-based view: Economic actors as theorists. *Strategy Science*, *2*(4), 258–271.

Gans, J. S., Stern, S., & Wu, J. (2019). Foundations of entrepreneurial strategy. *Strategic Management Journal*, *40*(5), 736–756.

Gardenfors, P. (1988). *Knowledge in flux: Modeling the dynamics of the epistemic states*. Cambridge, MA: MIT Press.

Gavetti, G., & Menon, A. (2016). Evolution cum agency: Toward a model of strategic foresight. *Strategy Science*, *1*(3), 207–233.

Grandori, A. (1984). A prescriptive contingency view of organizational decision making. *Administrative Science Quarterly*, *29*, 192–209.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review*, *116*(4), 661–716.

Hamel, G., & Prahalad, C. K. (1989). Strategic intent. *Harvard Business Review*, *67*, 63–76.

Harman, G. (1986). *Change in view*. Cambridge, MA: MIT Press.

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk. *Management Science*, *39*(1), 17–31.

Kamenica, E., & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, *101*(6), 2590–2615.

Kaplan, S. (2008). Framing contests: Strategy making under uncertainty. *Organization Science*, *19*(5), 729–752.

Kerr, W. R., Nanda, R., & Rhodes-Kropf, M. (2014). Entrepreneurship as experimentation. *Journal of Economic Perspectives*, *28*(3), 25–48.

Knight, F. H. (1921). *Risk, uncertainty and profit*. New York, NY: Houghton Mifflin Company.

Lewis, D. (1973). *Counterfactuals*. Cambridge, MA: Harvard University Press.

Martignoni, D., Menon, A., & Siggelkow, N. (2016). Consequences of Misspecified mental models: Contrasting effects and the role of cognitive fit. *Strategic Management Journal*, *39*(5), 1350–1369.

Meija, Z. (2018). Why Elon Musk wants his employees to use an ancient mental strategy called 'first principles'. CNBC.com. Retrieved from https://www.cnbc.com/2018/04/18/why-elon-musk-wants-his-employees-to-use-a-strategy-called-first-principles.html

Merriam-Webster. (2021). Retrieved from: https://www.merriam-webster.com/dictionary/conjecture

Musk, E. (2006). The secret tesla motors master plan (just between you and me). Retrieved from: https://www.tesla.com/blog/secret-tesla-motors-master-plan-just-between-you-and-me

Nickerson, J., & Argyres, N. (2018). Strategizing before strategic decision making. *Strategy Science*, *3*(4), 592–605.

Nickerson, J., Yen, C. J., & Mahoney, J. T. (2012). Exploring the problem-finding and problem-solving approach for designing organizations. *Academy of Management Perspectives*, *26*, 52–72.

Nickerson, J. A., & Zenger, T. R. (2004). A knowledge-based theory of the firm: The problem-solving perspective. *Organization Science*, *15*(6), 617–632.

Ocasio, W. (1997). Towards an attention-based view of the firm. *Strategic Management Journal*, *18*, 187–206.

Pillai, S. D., Goldfarb, B., & Kirsch, D. A. (2020). The origins of firm strategy: Learning by economic experimentation and strategic pivots in the early automobile industry. *Strategic Management Journal*, *41*(3), 369–399.

Popper, K. R. (1959). *The logic of scientific discovery*. Hutchison: London, UK.

Porter, M. E. (1980). *Competitive strategy*. New York, NY: The Free Press.

Prahalad, C. K., & Hamel, G. (1994). *Competing for the future*. Boston, MA: Harvard Business School Press.

Ries, E. (2011). *The lean startup*. New York, NY: Crown Publishing.

Rindova, V., & Martins, L. L. (2021). Shaping possibilities: A design science approach to developing novel strategies. *Academy of Management Review*.

Rumelt, R. (2011). *Good strategy/bad strategy*. New York, NY: Crown Business.

Schmidt, J. (2015). Cognition, resources, and opportunities: Managerial judgment, theories of success and the origin of novel strategies. *Advances in Strategic Management*, *32*, 549–575.

Schoemaker, P. J. H. (1993). Multiple scenario development: Its conceptual and behavioral foundation. *Strategic Management Journal*, *14*(3), 193–213.

Simon, H. A. (1962). The architecture of complexity. *Proceedings of the American Philosophical Society*, *106*, 467–482.

Spohn, W. (2012). *The laws of belief: Ranking theory and its philosophical applications*. Cambridge, UK: Oxford University Press.

Thiel, P. (2014). *Zero to one: Notes on startups, or how to build the future*. Crown Business: New York, NY.

Trigeorgis, L., & Reuer, J. J. (2017). Real options theory in strategic management. *Strategic Management Journal*, *38*(1), 42–63.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology*, *12*(3), 129–140.

Weick, K. E. (1989). Theory construction as disciplined imagination. *Academy of Management Review*, *14*(4), 516–531.

Wiltbank, R., Dew, N., Read, S., & Sarasvathy, S. D. (2006). What to do next? The case for non-predictive strategy. *Strategic Management Journal*, *27*(10), 981–998.

Zellweger, T. M., & Zenger, T. R. (2021). Entrepreneurs as scientists: A pragmatist approach to producing value out of uncertainty. *Academy of Management Review*.

—WILEY—

APPENDIX

PROOFS OF PROPOSITIONS
Proof of Proposition 1

By Axiom 1, the strategist will resolve all i contradictions caused by the counter-theories that each imply $\neg S_i$, respectively. Because we assumed that each counter-theory i only contradicts one premise of the conjecture S_i, to remove the contradictions it suffices to remove the respective weakest premises of each of the counter-theories (by Axiom 2). By our assumption, the weakest premise of $\neg S_1$ is stronger than the weakest premise of $\neg S_2$, and so on. Thus, if the weakest premise of $\neg S_1$, that is, E_1, is weaker than W, all the weakest premises of the beliefs $\neg S_i$ are weaker than W. In other words (using the notion introduced in the body text), if E_1 is weaker than W, any E_i is weaker than W. By Axiom 2, thus, all contradictions are resolved by giving up premises of the counter-theories, but not of the theory for Z, and the strategist maintains their belief in Z.

Proof of Proposition 2

As the strategist considers W to be stronger than E_1, he thus considers W to be stronger than any E_j. Thus, all contradictions between the theory and objections will be removed by giving up all E_js (as stated in Proposition 1). By Axiom 2, the strategist only discards as few beliefs as necessary to resolve contradictions. Thus, all beliefs from the counter-theories except for the E_js will be accepted by the strategist. As the beliefs of the strategist remain consistent and the strategist is assumed to form all logical implications of their beliefs (by Axiom 1), it follows that the strategist believes $\neg E_j$ is true for all j, or equivalently, that E_j is false for all j. This follows, as the strategist maintains their belief in the theory for Z *and* all beliefs that are premises for $\neg S_i$ except E_i. Adding E_i again would again create a theory for $\neg S_i$, as all other beliefs of the counter-theory are maintained. Thus, after learning the objections but giving up their weakest premises, the strategist believes $E_i \Rightarrow \neg S_i$. As the strategist also believes that S_i is true (all premises of the theory are maintained when W is considered to be stronger than E_1), E_i must be false.

Note that the strength of the belief $\neg E_i$ is equal to its weakest premise (Axiom 2). The strength of $\neg E_i$ is thus the weaker of the following two strengths: The strength of S_i, and the strength of the weakest of the remaining beliefs of the counter-theory. If the strategist considers all of the remaining premises of the counter-theory i (i.e., all premises except E_i) to be at least as strong as S_i, the latter strength is at least as high as the former, and thus the strength of $\neg E_i$ is equal to S_i. In this case, $\neg E_i$ is a premise for Z because given their belief strength ordering, the strategist would have to give up belief in Z if they learn that E_i is true. The resulting contradiction would be resolved by giving up both $\neg E_i$ and S_i (and thus also Z). If, however, the strategist considers at least one of the remaining premises of counter-theory i (i.e., one from the set of premises except E_i) to be weaker than S_i, then the strategist who learns E_i is true will resolve the resulting contradiction by giving up the weakest of the remaining premises of the counter-theory, but not S_i. Thus, the strategist can learn that E_i is true without giving up the belief in Z and $\neg E_i$ is thus not a premise for Z in this case.

Proof of Proposition 3

If a counter-theory implies that S_i is false while W is a premise for S_i, and the maintained premises of this counter-theory are at least as strong as W, $\neg E_i$ is a premise for Z (as stated in Proposition 2). Moreover, this premise has equal strength to W (as also stated in Proposition 2). If there is one such counter-theory i, by definition of a focused experiment, a focused experiment has to show that both $\neg E_i$ *and* W are true to confirm the strategist's belief in the conjecture. If there is more than one such counter-theory i, j, a focused experiment has to show that all $\neg E_i$, $\neg E_j$, and W are true. If no counter-theory contradicts a premise S_i in the theory that has W as its weakest premise, either no counter-theories exist or all counter-theories contradict premises S_j in the theory that are stronger than W. If the corresponding $\neg E_j$ are at least as strong as S_j they will thus not be weakest premises, as S_j is stronger than W. If the corresponding $\neg E_j$s are weaker than S_j, they will not become premises for Z. Thus, W remains the weakest premise of Z and it is a focused experiment to scrutinize W.