Symbiotic Theorization

Min Ding

Published online: 10 June 2014
© Springer Science+Business Media New York 2014

Abstract Grounded theory method (GTM) is the standard theorization approach used to develop theories grounded in data (TGD). In this paper, I propose symbiotic theorization (ST) as an alternate approach to developing theories grounded in data. Compared to GTM, ST elevates the role of subject matter owners (SMOs, the people whose behaviors a theoretician is studying) to theory aggregator to allow true emergence of theories grounded in data from their actions and interactions. This simultaneously reduces the role of theoretician to moderator, catalyst, and integrator of the theory. ST's interpretation and implementation of emergence brings it in line with the modern concept of emergence in complex system theory. ST also introduces a third type of participant, theoretical coders (TCer), who are well versed in various theoretic codes and are hired to serve as catalysts for the emergence of theoretical relationships. ST specifies an environment where subject matter owners and theoreticians are motivated to work in a symbiotic manner to generate a good theory that is sufficiently abstract and comprehensive, while explaining and predicting well.

Keyword Grounded theory · Emergence · Symbiotic theorization

1 Introduction

Following Hage's [11] widely adopted definition, *theory* in this paper is defined as a coherent framework consisting of a set of interrelated concepts that can explain and predict a phenomenon. I define *theorization* as the method used to develop a theory.

According to a major school of thought on theorization, theories grounded in data (TGD), theories should emerge from the data, instead of from the minds of theoreticians (see, for example, Merton [12]). Similar philosophy has also substantially influenced the development of marketing science, as Bass wrote "it is also possible for the observation to come first and for the explanation to follow. Science is a process in which data and theory interact" [1, p3]. A milestone in marketing related to this is the 1994 Wharton conference on empirical generalization that culminated in a special issue in *Marketing Science* in 1995. The empirical generalization approach is championed by Ehrenberg in marketing (see [5, 6]), and "when juxtaposed to the 'a priori' philosophy commonly associated with economics and other social sciences, they are likely to continue to have a profound impact on the thinking of marketing scientists" [2, G7].

This philosophy of theorization is best represented by the grounded theory (GT) literature, the origins of which can be traced to Glaser and Strauss [8]. In fact, grounded theorists now use GT to refer to both their theorization approach and their outputs (theories grounded in data). This usage, however, is conceptually incorrect. Glaser and Strauss [8] proposed one (successful) theorization for generating theories grounded in data; however, theirs is not the only method. In this paper, I thus use GTM to refer to the GT method originally proposed by Glaser and Strauss [8], and *theories grounded in data* to refer to theories that are grounded in data (and more generally, the philosophy that theories should be grounded in data, which existed before their 1967 work).

M. Ding (✉)
Smeal College of Business, Pennsylvania State University,
University Park, PA 16802-3007, USA
e-mail: minding@psu.edu
URL: <http://www.planetding.org>

M. Ding
School of Management, Fudan University, Shanghai 200433,
People's Republic of China

Following the philosophy of theories grounded in data, and taking the core idea of emergence in GTM one step further, I propose in this paper an alternate theorization for generating theories grounded in data, namely, symbiotic theorization. In the rest of the paper, I first provide a brief review of theory and extant ways to develop theories, and identify uncharted space on a theorization map. I then describe symbiotic theorization and explain its specific design before discussing several issues related to ST.

2 A Brief Review of Theory and Theorization

In this section, I first provide a brief overview of theories based on their scope. I then shift my focus to reviewing extant methods used in developing theory, i.e., theorization. Finally, I discuss an important gap in theorization and why it is fruitful to pursue this new method, providing both the foundation and motivation for the proposed symbiotic theorization in this paper.

2.1 Theories

Based on the scope they cover, theories can be roughly classified into three categories: grand theories, middle range theories, and everyday hypotheses (or lay theories). Middle range theories can be further divided into formal theories and substantive theories. I discuss each in this section.

Grand theories have the broadest scope and are represented by the work of influential sociologist Talcott Parsons (1902–1973). Although he did not personally like the label, action theory [14] is a particularly good example of a grand theory. A more neutral characterization for this type of theory would be general social theories. In grand theory tradition, one believes that overarching theoretical frameworks can be used to understand general social structures. Such theories may draw on a variety of social sciences such as psychology, political science, and philosophy, in addition to sociology. Grand theories tend to be highly abstract and are expected to cover the whole spectrum of social phenomena. Thus, they do not have explicit and testable propositions. Normally, one must extrapolate from grand theory to a specific substantive domain to create testable propositions unique to a specific context.

At the other end of the spectrum are the everyday hypotheses that normally are not considered to be true “theories,” as they tend to be disorganized, less abstract, have limited applicability, are often myopic, and quite possibly, wrong. Best practices exemplify everyday hypotheses. An expert (typically a subject matter expert) in a particular domain writes about the best practices in that domain based on personal experiences and/or case studies. This type of theory is very prevalent in management; successful CEOs often write books and dispense advice. Other examples include people writing about

how to raise kids, how to maintain a happy marriage, etc., based on their personal experiences. Best practices could also be verbal heuristics passed from person to person. Another example is lay theories, or beliefs held by people and used in their daily lives that have never been scientifically tested. Lay theories are often expressed in proverbs; for example, there is a Chinese saying, “To extend your life by a year take one less bite each meal.” Different cultures tend to have overlapping but not identical lay theories on the same subject.

Between these two extremes, Merton [12] proposed the idea of middle range theories, defined as “theories that lie between the minor but necessary working hypotheses that evolve in abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization, and social change” (p. 448). A good example of a middle range theory is the theory of reference groups [13]. The two key purposes of middle range theories are (1) to build theories on empirical phenomena and develop testable propositions, and (2) to incrementally build more general theories based on already established middle range theories to eventually reach the level of grand theory. The idea of middle range theory in sociology is very much inspired by how theories are developed in the natural sciences including biology, chemistry, and physics.

Middle range theories can be further divided into formal theories and substantive theories based on the level of abstraction. Substantive theories are formulated for specific subject domains (e.g., nursing, retailing service, business-to-business negotiation). Theories in marketing are substantive theories, although different marketing theories may be obtained using different approaches, as discussed in the next subsection on theorization. Formal theories are applicable to many substantive domains, and the theories (both the constructs and hypotheses) do not include anything referring to a specific subject domain. They typically cover a formal (conceptual) area of interest (e.g., reference group theory). Middle range theories are the dominant theories in contemporary social science and are also the focus of this paper.

2.2 Theorization

There are many different approaches to developing theories (i.e., theorization). In this subsection, I classify theorization methods based on the roles the theoretician (T) and the subject matter owners (SMOs) play in the process, which I depict visually on a theorization map (Fig. 1). On the theorization map, the theorization frontier separates the entire space into the feasible region of theorization (where subject matter owners and a theoretician can potentially create theories) and the infeasible region of theorization (where subject matter owners and a theoretician cannot create meaningful theories). On this map, existing theorizations can be grouped into four

Fig. 1 The theorization map

types, all along the theorization frontier. I discuss each type below, focusing on those related to theories grounded in data.

Type 1 Theorization This is the logical deductive theorization often used to construct grand theories and formal theories, but possibly used to create substantive theories as well. A theoretician may spend many years of mental effort to create a theory, with no or minimal input from subject matter owners. In some cases, data from subject matter owners may serve as inspiration for the theoretician. An analytical theory on competition, for example, may be derived by combining an existing formal theory (e.g., game theory) with a set of assumptions applicable to a specific context in a deductive manner. This type of theorization is the opposite of theories grounded in data, and I will not go into more detail as it is not the focus of this paper.

Type 2 Theorization Here, the theoretician plays the role of theory conjecturer, based on and verified by ample data supplied by subject matter owners. The logic behind this type of theorization is typically characterized as inductive then deductive for verification. This theorization is often used to develop substantive theories and, to a much lesser extent, formal theories. This type also includes conjecture from a formal theory (or a grand theory) to a substantive area, then verification of hypotheses using proper data. A theory in managerial (strategy) research with specific hypotheses, for example, may

be conjectured from a formal (but vague) theory (e.g., transaction theory), and then scholars will test the hypotheses to see whether the theory is sound. In other words, the conjecture can be either bottom up (from data) or top down (from grand or formal theories). While type 2 relies much more on data than type 1, it does not satisfy the criteria of theories grounded in data due to the use of conjecture.

A special note should be made regarding the existing three variations of GTM, namely, the Straussian approach [15], the classic Glaserian approach [7, 8], and the constructivist approach put forth by Charmaz [3]. Out of these three, the Straussian approach is a type 2 theorization. While Strauss and Corbin have labeled themselves as grounded theorists, their approach has evolved into an inductive method that is followed by deduction for verification. They ask theoreticians to follow a specific theoretical paradigm and seek to fill the paradigm from data in each application, while allowing conjectures. As such, the Straussian approach no longer strictly follows the theories grounded in data tradition, where no conjecture (deduction) is allowed and verification is not needed, even though some Straussian techniques can be traced to the 1967 joint work between Glaser and Strauss.

Type 3 Theorization Here the theoretician acts as either the discoverer or interpreter (depending on her epistemological inclination) of the theory and she is not allowed to engage in conjecture by any means. All theoretical components

(concepts and relationships) must emerge from the data supplied by subject matter owners, and should not be forced by theoretician.

In other words, a theoretician cannot have predetermined ideas and then try to find data to support his theoretical framework. In type 3 theorization, subject matter owners are data suppliers with two characteristics. First, they are passive suppliers; theoreticians need to go to them to obtain the kinds of data they want; second, only a subset of data they can provide will be useful to theory building (i.e., those sampled by the theoretician that can contribute to theoretical saturation).

The Glaserian version and the constructivist version of GTM belong to type 3 theorization. The Glaserian version takes a post-positivist perspective; Glaserians believe theory is there to be discovered from data, or in Glaser's words, "it will always emerge." Glaser argued that theories developed under the Glaserian approach require no verification; since such theories emerge from the data, by definition, they must be true. Glaserian theorization is purely inductive; deduction is only used sparsely for theoretical sampling purposes (unlike the Straussian approach). The constructivist view, which is the prevalent view in sociology these days, is that there are no objective data. Rather, the data and resulting theory reflect the theoretician's own interpretation of reality. While in the Glaserian approach, the theoretician acts as a theory discoverer, in Charmaz's constructivist approach, the theoretician acts as a theory interpreter. Despite this difference between the two, both argue for the theory to emerge from the data without forcing.

The core concept in type 3 theorization is emergence. I want to make two observations about this concept. First, according to Glaser [7], emergence does not mean the theoretician just sits back and waits for things to appear. Theoreticians are encouraged to learn as many theoretical codes as possible, and to try to see which theoretical code fits a particular set of data. Using this constant comparative method is supposed to allow the concepts and relationships to emerge. In other words, the process involves a lot of work by theoreticians, and emergence happens in the theoreticians' minds through analysis. Second, the word "emergence" refers to its original English meaning (i.e., the process or action of coming into view or becoming exposed after being concealed; Oxford Dictionary). In the context of theorization, it means if a theoretician examines the right data carefully and correctly, she will eventually have a Eureka moment and identify a theory. Note that this usage of the concept "emergence" differs from how the word is used in most academic literature these days, where it is associated with complex systems and refers to the fact that properties of a complex system emerge from very different properties of individuals within the system.

Type 4 Theorization Here the theoretician plays no role and the subject matter owners are the ones who generate the theory. Given the lack of theoretic training, subject matter owners typically intuit a hypothesis based on their experiences, which is then either adopted or forgotten by their peers over time. Everyday hypotheses (including lay theories and best practices) are created in type 4 theorization. Note that the creation of lay theories is more consistent with the academic definition of emergence discussed earlier, in that these theories truly emerge from individuals' actions and interactions, instead from someone's mind.

2.3 Uncharted Space Along the Theorization Frontier

A close examination of the theorization map (Fig. 1) reveals an uncharted space along the theorization frontier, between types 3 and 4 (marked by the circle with ? in the center). Note that this is an inherent blank space, instead of something created by merely adding more labels to the axis of theoreticians and subject matter owners. Theoreticians' roles in types 1, 2, and 3 are as *active* participants in theorization, even though the roles vary from creator to interpreter. By active, I mean no theoretical blocks (concepts, relationships) come into existence without the theoreticians. In type 4, however, the theoretician plays *no* role at all. So it is logical to identify a gap between an *active* role and *no* role in which the theoretician plays a *passive* role. As such, theoretical blocks come into existence even without her involvement, but her involvement somehow facilitates such emergence. For the subject matter owners, the opposite is true. Between the passive data provider role played in types 1, 2, and 3 and the sole theory intuitor role played in type 4, a logical gap exists where they play an active role in creating theoretical blocks in a more systematic manner with some guidance from professional theoreticians. Together, these two gaps constitute the uncharted space along the theorization frontier that is the motivation for the work described in this paper. Of course, there could be additional gaps between type 1 and type 2, and between type 2 and type 3, with more subtle differences. Such gaps, together with other blank space in the Feasible Region of Theorization, are candidates for additional new methods that I briefly discuss in the last section of this paper.

3 Symbiotic Theorization

In this section, I describe symbiotic theorization (ST), which I propose to fill this uncharted space. ST follows the theories grounded in data tradition, and is mainly used to generate substantive theories, and possibly formal theories. In other words, ST is an alternate theorization to GTM, but equivalent in status, both of which are designed to create theories

grounded in data. As such, I provide a comparison between ST and GTM at a conceptual level, after defining ST. At the end of this section, I discuss the criteria for designing the actual procedure for ST.

3.1 Definition

By *symbiotic theorization*, I refer to any theory-generating methods that satisfy the following criteria: (1) all theoretical blocks (concepts and relationships) emerge from the action and interaction of subject matter owners; (2) theoreticians (and any third parties) may only serve as moderators and catalysts of the emergence and integrators of the final theories; and (3) subject matter owners and theoreticians are incentive aligned to produce theories that are desirable for themselves.

In this context, a moderator is someone who maintains rules and ensures the process goes smoothly and in the right direction. A catalyst here, like a catalyst in the natural world, speeds up the reaction (in this case, the creation of theoretical blocks) but has no impact on what specific output is generated.

Unlike previous theorization approaches, ST has at least two types of individuals contributing to the final outcome and acting in their own best interests; in other words, it is a game. As such, ST is conceptually similar to the mechanism design literature (see [10] for an accessible introduction). While ST does not resort to the mathematical formulation of typical mechanism design and the incentive compatible condition, it adopts the idea of incentive alignment, a weaker and

nonmathematical form of incentive compatibility that has been used in marketing and economics literature [4].

3.2 Conceptual Comparison Between ST and GTM

Before I compare these two methods, I must highlight again that both methods are theorization approaches for developing theories grounded in data and both follow theories grounded in data philosophy strictly, so I will not repeat all the supporting arguments made in the long history of theories grounded in data theorization here. Instead, I focus on the difference between the two methods in this subsection (a more specific comparison on intermediate objectives, participants, and core components are provided in Table 1).

ST provides benefits over GTM by elevating the importance of emergence, which is the foundation of GTM (and theories grounded in data), and by following the tradition of generating emergent properties in complex systems. I discuss these benefits specifically below.

First, ST implements a definition of emergence that is consistent with extant literature on emergence in the study of complex systems that has become dominant in recent years. This relates ST to a state of the art intellectual movement and enables insights to be drawn from it to benefit theorization.

The core strength of GTM is the concept of allowing theory to emerge from data. However, despite heated discussions within the GTM domain, all three approaches of GTM result in theories marked by the imprints of the theoreticians. The Straussian approach requires the data to fit into a certain paradigm the theoretician has predetermined. Charmaz

Table 1 A comparison between GTM and ST

		GTM	ST
Intermediate objectives	Main concern	Open coding via CCA	Interaction among SMOs
	Core category	Open coding via CCA	Interaction among SMOs
	Concepts	Selective coding via CCA	Interaction among SMOs
	Relationships	Theoretical coding via CCA	Interaction among SMOs
	Integration	Integration (Sorting)	Integration
Participants	Subject matter owners (SMOs)	Data provider	Data provider, theory aggregator (theoretical blocks, i.e., concepts and relationships), theory evaluator
	Theoretician (T)	Theory discover or interpreter	Moderator, catalyst of emergence, theory integrator
	Theoretical coder (TCer)	N/A	Catalyst of theoretical relationships, possible catalyst of concepts
Core components	Emergence	From constant comparative analysis by T	From the actions and interactions of SMOs, assisted by T (and TCers)
	Theoretical sensitivity	Needed by T for all tasks	Helpful for theory integration
	Theoretical codes	T must learn and use	Helpful, but not required for T
	Theoretical sampling	Determined by T	Self-generated automatically by SMOs
	Memos	By T	By T and SMOs
	Sorting	By T	By T
	Integration	By T	By T, assisted by SMOs

CCA constant comparative analysis

follows the intellectual tradition of constructivism where theoreticians openly acknowledge a theory as a co-construction between the theoretician and data, and different scholars may generate different theories even with the same set of data. Glaser insists on the discovery of theory from data and abhors any imposition on data to yield theory. Nevertheless, he still asks theoreticians to learn as many theoretical codes as possible so that they can later test which one(s) should be used for the data at hand. As a result, the difference between Glaserian approach and Straussian approach may not be black and white. Another key aspect of all GTM approaches requires the theoretician to contribute actively to the creation of the theory, which is called theoretical sensitivity. In other words, a theory depends on who is developing it; the corollary of this is that no theory is 100 % based on data. In other words, the theory emerges through data *but in a theoretician's mind*.

A true emergence, based on the contemporary definition in the context of complex systems, should be a theory that appears from the action and interaction of individual subject matter owners, outside a theoretician's own mind. This would make it much closer to the modern definition of emergence used in the academic literature, i.e., emergence is “the arising of novel and coherent structures, patterns and properties during the process of self-organization in complex systems” [p. 49, 9]. This is the definition adopted and implemented in ST, thus fully realizing the potential of the fundamental motivation behind the philosophy of theories grounded in data.

Second, the emergence implemented in ST enables—but does not guarantee—the creation of theories that explain and predict phenomena better. While the notion of having subject matter owners develop theoretical blocks is a radical idea and many may cringe at the notion that subject matter owners can create theories that are as good (or even better) than those created by professional theoreticians, this concern is unwarranted. The key defining feature of a complex system is that it is capable of producing (emerging) properties that an individual member, on its own, is incapable of producing. As a matter of fact, this already happens in real world. Lay theories, albeit the simplest type of theory, are truly the result of emergence among relevant subject matter owners. With proper assistance from theoreticians employing ST in the right settings, subject matter owners will be able to produce more abstract and comprehensive theories (such as substantive theories) through a similar emergence process.

As a matter of fact, all theoretical blocks that emerge from the quasi-complex system created in ST are organic, and have gone through a process of constant evaluation and modification by subject matter owners. The intuition is simple. One trained theoretician is better at creating a theory than one subject matter owner, but many subject matter owners interacting with each other, moderated and catalyzed by theoreticians (and other third parties), will likely create better theories than one theoretician.

Third, ST substantially reduces the contamination (and eliminates forcing) of theoreticians in the creation of the theory. One guiding principle in GTM is that the theoretician should not force the theory (i.e., she should not have preconceived theoretical notions and then try to find the data to support them). As Glaser puts it, one can always find the data to complete a preconceived paradigm; however, doing so counters the philosophy of theories grounded in data. Although all three variations of GTM have different interpretations of forcing, no version recognizes forcing in its own approach. The constructivist version by Charmaz at least acknowledges the role of theoretician in building theories and posits that no theory can be free from the imprint of the theoretician who builds it. (I call this *contamination* in this paper.)

By relieving theoreticians of any active role in identifying theoretical blocks, ST eliminates even the interpretation aspect of theory building per Charmaz. The emergence of theoretical blocks happens in the open among subject matter owners, instead of in the mind of a theoretician, and forcing is structurally infeasible, even if it is desired. One may thus argue that ST can generate theories that are even more grounded in data than GTM.

Fourth, ST eliminates the need to decide which logical approach (inductive, deductive, abductive) to use in theory generation. Different theorization methods are based on different logical approaches (or combinations thereof). Even within GTM, there is an ongoing debate on exactly what kind of logical processes are, or should be, employed. Glaser insisted that GTM should only be considered an inductive approach, and frowned upon the incorporation of deduction in the Straussian approach, while Charmaz (and others) incorporated the concept of abduction explicitly in their framework, and argued that other variations also use abduction.

The emergence implemented in ST enables agnosticism about the logical process. The theoretical blocks come from the actions and interactions of subject matter owners within the system, not from any typical logical process employed by a theoretician. Like complex systems everywhere, it enables the potential emergence of new properties.

Fifth, ST reduces theoretician-based idiosyncrasies in developed theories and facilitates replicability. This is an important feature of ST with substantial ramifications. Since the theoretical blocks emerge independent of theoreticians (i.e., different theoreticians can interchangeably perform the same monitoring and catalyzing roles), it is then natural to conclude that very similar theories can be generated by two different theoreticians who have employed ST independent of each other. This is contrary to the position held by GTM scholars that theories grounded in data generated by GTM should not be considered replicable, and their insistence that their theories should not be judged by replicability. Such a strong position by GTM scholars, of course, is contrary to other scientific inquiries where replication is the touchstone of truth.

Finally, ST motivates subject matter owners and theoreticians to create theories that will most benefit them. There is always a tendency in theorization, as in other human endeavors, to produce an output that meets a minimum standard. In GTM, the theoretician is the judge, jury, and executioner. She decides when the theory is sufficient and whether the theory explains data and predicts well. Even when a given GTM work is peer reviewed, the reviewers typically do not see the raw data and thus have no idea how well the theory works. In ST, both subject matter owners and theoreticians are motivated to create a theory that they deem useful to them, where subject matter owners push each other to create better theoretical blocks, and the subject matter owners push theoreticians to help create something that is satisfactory to subject matter owners, and vice versa.

3.3 Key Design Decisions

The first design decision involves properly discretizing the process that corresponds to the generation of theories grounded in data, so that each stage corresponds to a specific task for subject matter owners. Since real-life emergence is not separated by exogenously imposed stages, a continuous process, if implemented in ST, will lead to inefficiency. In ST, I suggest a discretization scheme based on the key stages in GTM.

The second design decision for ST is that it must create an environment in which the subject matter owners can share their experiences and thoughts and react to each other, which may potentially enable abstract concepts and relationships to emerge. The environment should mimic a complex system in society, where many individuals freely interact with each other in an unstructured manner over a long duration. This environment could be the same for all stages of ST after discretization, or be specific for different stages.

Two necessary conditions for satisfying this decision are enabling technologies and virtual interaction which, fortunately, have become mainstream in recent years and thus make real-life implementation of ST feasible. Internet technology has made it possible to build a virtual platform that resembles a quasi-complex system with ease. The prevalence of forums and social media has also primed people to engage in open dialogues with people they have never met before in person.

The third design decision is to build *moderators* and *catalysts* into the system, and to determine who should play those roles. A complex system, if left on its own, can take a long time to reveal an emergent property, if at all. The moderator will keep the subject matter owners on the right track toward generating a theory by avoiding digressions, inappropriate interactions, and any other disruptions to the emergence process. He will also communicate the objectives to the subject matter owners and ensure the quality of the theory (i.e., sufficiently abstract, comprehensive). The catalyst will help accelerate the emergence. Like a catalyst in the natural world,

a catalyst in ST should never influence the nature of the output except by speeding up the process.

The fourth design decision involves building appropriate incentives into ST such that the incentives of subject matter owners and theoreticians are aligned to produce a good theory based on their own objectives. Specifically, subject matter owners want a theory that can explain and predict well and the theoretician further wants a theory that is sufficiently abstract and comprehensive.

4 Designing for Symbiotic Theorization

This section describes the details of one ST design that represents my best effort. Like Glaser and Strauss [8], I invite, and most certainly expect, alternate designs for ST to be proposed in the future. The conceptual discussion of ST in the previous section is what defines ST, not a particular implementation detail described in this section.

Following the guidelines laid out in the previous section, I first describe the discretization decisions made for this design. The rest is then organized following the typical description of a game: players and their payoffs, and the process (stages) of the game, including strategies and information.

4.1 Discretization of the ST Process

The objective of any theorization approach is to generate a theory that explains and predicts well. Many intermediate outputs that must be obtained, however, before such a theory can be generated. While most agree on what a final theory should accomplish, theoreticians disagree on which intermediate outputs should be created before the final theory is formulated; in other words, there are many different ways to arrive at the final destination.

In this design, I adopt the same intermediate outputs that have been very successfully used in GTM over the last half century; I mostly follow the nomenclature of Glaserian approach [7, 8]. Here, I reiterate the relevant terms before I describe the discretization of the process.

Main concern: The main problem in a substantive domain

Concept: An abstract term that captures something important; this is similar to the word “construct” used in other literature

Core category: A special type of concept that can explain the most variance related to the main concern

Category: A special type of concept that tends to be more abstract

Properties: A special type of concept that is not as abstract or comprehensive as a category, and instead is associated with a particular category

Hypotheses (or theoretical relationships): Relationships among the constructs and properties, especially those related to the core category and its properties

Theoretical sampling: The process of finding data that will facilitate theory development

In GTM, there are five intermediate objectives: (1) identify the main concern; (2) identify the core category; (3) identify all concepts related to core category, through theoretical sampling; (4) identify all hypotheses, through theoretical sampling; and (5) integrate the above material into a coherent and useful theory.

In ST, I discretize the entire process into four stages that correspond to the intermediate objectives used in GTM, where stages 1 and 2 correspond to the first two objectives, stages 3a and 3b correspond to the third and fourth objectives, and stage 4 corresponds to the fifth objective. The details in each stage are described later in the subsection on process, strategies, and

information. While these intermediate outputs are essentially identical to those in GTM, how each is obtained is substantially different under ST, as should become clear in the descriptions below (see Table 1, and also Figs. 2 and 3).

4.2 Players: Definition, Inherent Objectives, Role, and Payoff

There are three types of players (or participants) in this ST design: theoretician, subject matter owner, and theoretical coder. The roles and relationships among these three types of players are graphically represented in Fig. 4, which is followed by a definition, inherent objective, role, and payoff for each type of player.

Theoretician As used throughout this paper and elsewhere in the literature, a theoretician is a trained scholar who desires to develop a theory of whatever scope and using whatever theorization. In extant theorizations, the theoretician is the only active participant in the process (types 1–3). The required skillset of a theoretician under ST, however, is somewhat different from other approaches. The chief difference is that

Fig. 2 GTM procedure

Fig. 3 Symbiotic theorization procedure

a theoretician is no longer required to learn as many theoretical codes as possible. On the other hand, a theoretician must learn to become a skillful case teacher, in the sense that she needs to be able to guide a discussion and help achieve consensus.

The theoretician plays three roles in ST: monitor, catalyst, and integrator. As a monitor, she must ensure that the emergence process does not go astray, and that a productive discussion is continuously flowing. More broadly, the monitor recruits appropriate subject matter owners and theoretical coders to participate in her ST application. As a catalyst, she needs to accelerate the emergence of each intermediate outcome. For example, she may point out when the abstraction is insufficient and urge subject matter owners to move towards a more theoretical term. One thing the theoretician will not do is to contribute to any substantive aspect of the theory. In other words, she will not generate any theoretical blocks (concepts or theoretical relationships). Like her peers in GTM, the theoretician in ST also must integrate the theory, but the theoretical blocks used for this purpose come directly from subject matter owners as opposed to the theoretician's mind.

Based on the definition of her job, the theoretician is driven by her motivation to develop a good theory. This inherent motivation, however, may not be sufficient. Sometimes theoreticians settle for an output that is “good enough” due to laziness, sloppiness, unwillingness to exert additional effort, or a simple lack of knowledge. In ST, subject matter owners are asked to evaluate the final integrated theory, and I recommend that a theoretician does not declare success in ST unless a previously agreed upon subject matter owner rating is achieved.

Subject Matter Owners (Subject Matter Experts and Subject Matter Apprentices) Subject matter owners are the people whose behavior a theoretician is studying. I further divide them into two groups. Subject matter experts (SMEs) are people who are considered experts in that particular domain based on their own experiences. Subject matter apprentices (SMAs) are people who are trying to better understand this domain, and possibly to do things better; they are either new to the domain, or have not been able to do things successfully in

Fig. 4 A cross-section view of ST (stages 1–3)

the past. I suggest using a mixture of subject matter experts and subject matter apprentices in ST, with the actual proportion depending on the specific context.

Subject matter owners are the most important players in ST. They are the people who contribute to the emergence of the main concern, core category, and all theoretical blocks (concepts and theoretical relationships). This is qualitatively different from their role as the passive data providers in GTM.

Like theoreticians, subject matter owners have inherent objectives for which they will derive positive utility from building a good theory, although they are mostly concerned about whether and how well it will work. Inherent objectives (and rewards) may be sufficient to motivate some subject matter owners to engage in the ST process, and the theoretician can screen subject matter owners prior to ST to eliminate those whose inherent motivation is minimal. To further

encourage subject matter owners to contribute to theory emergence, I also include a peer evaluation at each stage (except stage 4), and each subject matter owner receives a financial reward that is determined by the peer evaluation results.

Theoretical Coder Subject matter owners have no training in theorization and thus emergence from subject matter owners may take a long time, or even be impossible. At the same time, I want to relieve theoreticians of the burden to learn and use a large set of theoretic codes. More importantly, I want to remove the possibility of the theoretician “forcing” data by suggesting theoretic codes directly in an attempt to catalyze the emergence. This ensures a clean emergence free of any imprints from theoretician who has a stake in the final output.

I thus propose the inclusion of professional theoretical coders (TCers) in ST. Theoretical coders are people who are

well versed in various theoretic codes and can be hired to participate in ST. The use of theoretical coders also enables theoreticians to focus more on guiding the development and integration of the theory.

The role of a TCer is to facilitate the emergence of the theory by serving as a catalyst who suggests possible theoretical codes to subject matter owners and helps them evaluate which ones might capture the data in terms of theoretical relationships. They may also help the emergence of concepts themselves. While both TCers and the theoretician play catalyst roles, TCers do so more directly. Since subject matter owners know that TCers are simply hired guns, they experience no demand effect and can freely dismiss their suggestions.

TCers have no stake in the final theory and thus have no inherent motivation other than the financial reward provided by the theoretician for their time. In addition, they receive bonus payments based on evaluations provided by subject matter owners, other TCers, and the theoretician. In cases where TCers are hired from a pool of professionals seeking work, future employment is also a large motivating force to work diligently in ST.

4.3 Process: Strategies and Information by Stage

Before I describe each stage in detail, I want to highlight three key components in GTM and their corresponding components in ST (also see Table 1). The first is constant comparative analysis, which is the essence of generating emergence in GTM. In ST, there is no explicit constant comparative analysis; instead, constant comparative analysis is achieved through the constant interaction (discussion, modification, and consensus building) of subject matter owners, supported by empirical data. The second is theoretical sampling, another defining characteristic of GTM, where the theoreticians are asked to sample data selectively for the sake of building a dense theory. In ST, theoretical sampling is achieved implicitly by subject matter owners. As they seek to contribute to the emergence of the theoretical blocks, they look back on their past experiences *selectively* to identify any pieces of data that will enrich the theory. The third is memoing. In ST, theoreticians are encouraged to write memos throughout the entire process and the comments subject matter owners make to each other on the platform can be considered memos as well. Given the unbiased role of TCers, I do not recommend asking TCers to write memos for the sake of theory building, except to make notes for their own use in identifying better ways to propose theoretical codes.

Next, I discuss each stage in detail (refer again to Figs. 3 and 4 and Table 1), including the actions each participant can take, what they know, and how they will be compensated specifically during that stage if financial incentives are provided.

Stage 1 This stage is designed to identify the main concern. The environment in stage 1 is a moderated discussion forum implemented in an online format. Theoreticians and subject matter owners participate in this stage; subject matter owners act as members of the discussion forum and the theoretician functions as the moderator. Any submissions to the forum by subject matter owners are sent to the theoretician first for approval, which allows the theoretician to screen the submissions and fulfill her role as the monitor. A participant can see all approved submitted comments made by other participants.

At the beginning of stage 1, subject matter owners are asked to suggest, discuss among themselves, modify, and reach consensus about the main concern for the given domain of interest. All subject matter owners are free to create original threads on a candidate for the main concern, or to comment in a thread started by another subject matter owner. They are, however, required to provide empirical evidence (a snippet/incident) to support both their original proposed main concern and comments on another subject matter owner's proposal. The theoretician will not approve a submission without empirical evidence. At the conclusion of stage 1, subject matter owners will evaluate their peers based on their contributions at this stage, and each will receive financial compensation based on this peer review.

The main concern emerges when a super majority (70–90 %) of subject matter owners converges on a concern. At no time may the theoretician comment on which candidate concern is more appropriate; however, she can remind the subject matter owners about the criteria of a main concern. True to emergence, the theoretician can decide to stop the ST process after stage 1 if she feels the emerged main concern is of no intellectual interest to her.

Stage 2 This stage is designed to identify the core category. It starts when a satisfactory outcome is reached in stage 1, and the theoretician informs the subject matter owners to move on to the next phase of ST. It is very similar to stage 1 described previously except for the following differences.

First, the objective is different. Subject matter owners are informed and tasked here with identifying an abstract concept that can explain most of the variations in the main concern (i.e., the core category). Second, the theoretician must challenge subject matter owners to raise the abstraction level of their suggested candidates for the core category to a satisfactory theoretical level. Unlike the concrete identification of main concern, subject matter owners are not used to expressing concepts in a highly abstract manner; thus, constant prodding from the theoretician is most helpful and constitutes a form of catalyzing. Third, many candidates for core category will be generated in stage 2. Those not chosen as the consensus core category are nevertheless very valuable concepts and will serve as the starting point for stage 3.

The core category emerges, similarly, when a super majority of subject matter owners converges. Again, the theoretician may choose to abandon the ST process if she concludes the emerged core category does not promise sufficient intellectual contribution.

Stage 3 This stage is designed to identify the concepts (stage 3a) and theoretical relationships (stage 3b) (i.e., the theoretical blocks) to be used to build the eventual theory. It starts when a satisfactory core category is identified in stage 2, and the theoretician informs the subject matter owners to move into the next phase of ST. It is very similar to stage 1 described previously except for the following differences.

First, this stage has two iterating substages, 3a and 3b, each with a different objective. When a sufficiently large number of concepts have been generated in stage 3a, the theoretician announces that ST will enter stage 3b with the goal of identifying theoretical relationships. When no new theoretical relationships seem to emerge during stage 3b, the theoretician will ask the participants to return to stage 3a to identify additional concepts, likely stimulated by the interactions and the generation of theoretical relationships in stage 3b. This process continues until no meaningful concepts and theoretical relationships emerge. Whereas one could consider the essence of the open coding procedure under GTM is incorporated in stages 1 and 2, in stage 3 the equivalent of the selective coding in GTM is incorporated.

Second, like stage 2, the theoretician must constantly challenge the subject matter owners to raise the abstraction level in stage 3a. The theoretician also must prod subject matter owners to identify concepts as comprehensively as possible in stage 3a, and the set of theoretical relationships among them as accurately and comprehensively as possible in stage 3b.

Third, unlike stages 1 and 2, stage 3 includes a new type of participant, TCers.

They serve as catalysts, playing a key role by accelerating the emergence of appropriate theoretical relationships (stage 3b). I do not have a strong position on whether TCers should be used in stage 3a as well; I leave the choice to the theoretician based her specific application. If TCers are used in 3a, they must be warned to ground their suggestions on empirical data revealed by the subject matter owners, which is limited in the ST context, since the theoretical sampling is done implicitly in subject matter owners' own minds. Note that TCers can contribute substantially more in stage 3b since all the concepts to which a theoretical code may relate will have already been identified in stage 3a. At the conclusion of stage 3, each TCer will be rated by subject matter owners, other TCers, and the theoretician based on contributions, and individual bonuses will be awarded based on the combined ratings.

The emergence in stage 3 is different from stages 1 and 2, in which a singular outcome is sought. In stage 3, the emergence is at the concept or theoretical relationship level, thus a

separate majority opinion on each concept or relationship is required.

Stage 4 This stage is designed to achieve integration of the theory. It starts when the theoretician is satisfied that no new concepts or theoretical relationships will emerge in continued iterations of stages 3a and 3b. Unlike the previous stages, the theoretician is the main player in stage 4, while subject matter owners play the role of sounding board.

At the end of stage 3, many concepts (the core category and its properties, and minor categories and their properties) and their relationships will have been created. These theoretical blocks, however, require two additional tasks that must be completed by the theoretician before they can be presented as a complete theory. First, there will be too many such building blocks. A side effect of an emergent platform is the generation of numerous theoretical blocks. Some are critical and relevant, others are of minor interest (e.g., those weakly or not centrally related to the core category), or those affecting only a small subgroup of people. As a result, the theoretician must identify the subset of such building blocks to be incorporated into the final theory. Second, the theoretician must weave these selected theoretical blocks into a coherent framework. These two tasks must be completed under the guiding principle to generate a succinct theory that can explain the vast majority of variance and help subject matter owners do their jobs better.

During the theory integration process, subject matter owner consultation is iterative. The theoretician will present a prototype theory to subject matter owners for feedback before revising it and presenting it to subject matter owners again for additional comments. This process continues until the theoretician is satisfied that no Pareto improvement can be made. At this time point, the theory will be put to a vote to the subject matter owners on how useful the resulting theory is in helping them understand their challenges and do their jobs better.

The subject matter owners' ratings are made available to the theoreticians anonymously only after all have been paid and dismissed. Under the doctrine of ST, the theoretician should disclose this rating during and after peer review, and in any publications. A theoretician may decide to go back and repeat stage 3 if she receives an unsatisfactory rating on the theory from the subject matter owners, possibly with a different set of subject matter owners and/or TCers, to see whether an alternate theory can be generated.

5 Discussion

In this paper, I described symbiotic theorization as a new approach to theory development. It fills a gap on the theorization map, and most importantly, it follows the philosophy of theories grounded in data by taking GTM one step further and

enabling the true emergence of theory from subject matter owners through a process that is consistent with the emergence in the complex systems theory literature. In this section, I discuss the costs and key challenges associated with running ST compared to GTM, additional designs, and the role of the theoretician.

The benefits of ST do not come without costs or challenges. The most significant challenge is that ST requires a different skillset from theoreticians than those typically used in GTM or other theorization approaches. The fixed cost of learning to run ST successfully varies depends on the theoretician, with those who are skilled case teachers and/or forum moderators incurring less cost. The second cost is the variable cost associated with running ST. Again, this is a different type of variable cost compared to GTM. Variable cost in GTM relates to data collection and the mental effort and time spent on data analysis. Variable cost in ST is associated with recruiting the appropriate subject matter owners and TCers and running the ST application over time. This variable cost also depends on the individual. Some theoreticians will incur less variable cost under ST, others more. The third cost is related to building the environment for ST. I believe a virtual platform that all participants can access via internet is most appropriate (compared to face to face) because it reduces cost while maintaining anonymity. While not all theoreticians are well versed in building such platforms, a third party entity may be used to build a sharable platform for others to use. Building such a platform does not require a lot of resources for someone who knows how to do it; it can be built using all free software, such as a combination of PHP and MySQL.

While I have laid out the conceptual framework of ST and one specific ST design in this paper, the current work can be extended in two directions. First, one can modify and improve the current ST design. Such improvements could be as minor as determining the optimal parameter values for ST, such as the number of subject matter owners needed, or the ratio between TCers and subject matter owners. Other modifications could be major, such as incorporating additional stages or revising the existing structure (or discretization). I describe one such possibility, tournament-based ST, as food for thought.

Tournament-based ST would involve multiple teams competing to generate the best theory for the domain. Each team would consist of subject matter owners and TCers, and be assisted by a theoretician. The tournament could start at stage 1, which means teams would be free to come up with their main concern; or it could start at stage 2, which means teams would be free to come up with the core category, but all teams would be working on the same main concern; or it could even start at stage 3, which means teams would all be working on the same main concern and core category. The theoretician could determine where the tournament should start (i.e., teams could diverge) based on her particular situation. There are three key benefits in running a tournament ST format. First,

it creates excitement and more incentive for subject matter owners, and may accelerate the emergence in each team. Second, it creates independent material that the theoretician may choose to integrate across different teams' outputs (for stage 4). Third, it can potentially provide evidence of independent convergence (similar to replicability) if different teams create very similar theories, and thus enhance one's trust in the final theory. On the flip side, tournaments require substantially more effort in running multiple teams and recruiting the right subject matter owners for each team.

Even more ambitiously, one may even try to develop new theorization that is different from ST. Note that all extant theorizations (including ST) are on the frontier of theorization. Filling the blank space further inside the feasible region of theorization requires breaking down existing schools of thought. For example, a theory conjecturer can work with subject matter owners who are theory aggregators. Unlike ST, theoreticians in such situations can freely suggest new concepts and theoretical relationships to subject matter owners, seek feedback, and decide whether to finally incorporate them into the finished theory. At the same time, they are also free to pass judgment on the concepts and theoretical relationships that emerge. Such theorization, however, would violate the doctrines of theories grounded in data. I leave this to other scholars to explore the pros and cons of such new theorization approaches.

Last but not the least, it is important to assess the role of the theoretician under ST. This is a sensitive issue if a theoretician chooses to use ST to generate theories for her domain of interest. A theoretician may feel uncomfortable using ST as she may think the intellectual contribution she can claim as a scholar will be limited if all the theoretical blocks come from the actions and interactions of the subject matter owners. Even the job of mastering theoretical codes can be outsourced under ST. I think this concern is unwarranted. The role of the theoretician under ST can be thought of as the role of a professor teaching a Harvard Business School case, except the professor does not have teaching notes and is not allowed to think about the case to form her own opinion before hand. The skills required to facilitate the emergence of a theory from the actions and interactions of subject matter owners and then weave them into a coherent and useful theory are not easy to acquire and should be every bit as valued as those used to develop theories under GTM.

References

1. Bass FM (1993) The future of research in marketing: marketing science. *J Mark Res* 30(1):1–6
2. Bass FM (1995) Empirical generalizations and marketing science: a personal view. *Mark Sci* 14(3):G6–G19

3. Charmaz K (2006) Constructing grounded theory: a practical guide through qualitative analysis. Sage, Thousand Oaks
4. Ding M (2007) An incentive-aligned mechanism for conjoint analysis. *J Mark Res* 44(2):214–223
5. Ehrenberg ASC (1968) The elements of lawlike relationships. *J R Stat Soc Ser A (General)* 131(part 3):280–302 and 15–29
6. Ehrenberg ASC (1995) Empirical generalisations, theory, and method. *Mark Sci* 14(3):G20–G28
7. Glaser BG (1978) Theoretical sensitivity. The Sociology Press, Mill Valley
8. Glaser B, Strauss A (1967) Discovery of grounded theory: strategies for qualitative research. The Sociology Press, Mill Valley
9. Goldstein J (1999) Emergence as a construct: history and issues. *Emerg Complex Organ* 1(1):49–72
10. Mas-Colell A, Whinston M, Green JR (1995) Microeconomic theory. Oxford University Press, Oxford
11. Hage J (1972) Techniques and problems in theory construction in sociology. John Wiley & Sons, New York
12. Merton RK (1949) Social theory and social structure. The Free Press, New York
13. Merton RK, Kitt A (1950) Contributions to the theory of reference group behavior. The Free Press, Glencoe
14. Parsons T (1937) The structure of social action. The Free Press, New York
15. Strauss A, Corbin J (1990) Basics of qualitative research: grounded theory process and techniques. Sage Publications, Thousand Oaks