Out of the frying pan and into the mixing bowl: Category grading and the adoption of novel practices

Peter Cebon  
*Melbourne Business School, University of Melbourne*

E. Geoffrey Love  
*University of Illinois at Urbana−Champaign, College of Business*

**Abstract**

Given that recent research has found that the meanings managers attribute to novel practices are crucial to understanding their adoption, we ask how institutional processes shape the way managers understand novel practices, and how that understanding affects adoption behaviour. In particular, we focus on how managers choose between component practices within a management practice (i.e. between members within a category). We argue, consistent with research in linguistics and anthropology, that categories are graded – that some members represent the category better than others. We go further to argue that the grading affects the value that participants attribute to category members. We build a model of the way processes of theorization and inter−organisational monitoring will grade categories in a field, and argue that, at any point in time, later adopters are more likely to have adopted central members of the category than earlier adopters. We test the resultant propositions by examining the adoption of Manufacturing Best Practice programs in Australia and New Zealand. Our data support our arguments regarding theorization, but not inter−organisational monitoring.

Published: 2006  
Out of the frying pan and into the mixing bowl: Category grading and the adoption of novel practices.

Abstract

Given that recent research has found that the meanings managers attribute to novel practices are crucial to understanding their adoption, we ask how institutional processes shape the way managers understand novel practices, and how that understanding affects adoption behaviour. In particular, we focus on how managers choose between component practices within a management practice (i.e. between members within a category). We argue, consistent with research in linguistics and anthropology, that categories are graded – that some members represent the category better than others. We go further to argue that the grading affects the value that participants attribute to category members. We build a model of the way processes of theorization and inter-organisational monitoring will grade categories in a field, and argue that, at any point in time, later adopters are more likely to have adopted central members of the category than earlier adopters. We test the resultant propositions by examining the adoption of Manufacturing Best Practice programs in Australia and New Zealand. Our data support our arguments regarding theorization, but not inter-organisational monitoring.
Introduction

Institutional research examining the adoption of novel practices has long focused on the way institutional pressures affect the perceived value of the practice and thus the likelihood of adoption (e.g. Baron, Dobbin, and Jennings, 1986; Burns and Wholey, 1993; Davis and Thompson, 1994; DiMaggio and Powell, 1983; Fligstein, 1985; Meyer and Rowan, 1977; Palmer, Jennings, and Zhou, 1993; Strang and Macy, 2001; Strang and Meyer, 1993; Tolbert, 1985; Tolbert and Zucker, 1983). These researchers have generally examined whether or not organizations adopt an administrative practice. The binary dependent variable (adoption or not) has meant such studies (often implicitly) treat the meaning of the practice, other than aspects related to its legitimacy (Meyer and Rowan, 1977), as relatively unproblematic. Recently, however, several researchers have turned their focus from whether organisations adopt novel practices to variations in how they adopt them. Novel practices are often complex, and contain many components, so managers face many choices about how they will adopt. This later research has found that the meanings that decision-makers construct of the practices shape those adoption choices (e.g. Cole, 1999; Czarniawska-Joerges and Joerges, 1996; Easton and Jarrell, 2000; Sahlin-Andersson, 1996; Zbaracki, 1998; Zilber, 2006). These researchers have found that managers spend time making sense of the innovation (Zbaracki, 1998), working out how it will fit within their organisation (Cole, 1999; Easton and Jarrell, 2000), and editing it (Sahlin-Andersson, 1996), translating it (Zilber, 2006), or configuring it (Zbaracki, 1998) to meet their understanding of their needs.

In this study, we extend research in this meaning-oriented tradition. Administrative practices such as Quality Circles, Total Quality Management, Manufacturing Best Practices, and Six Sigma all contain component practices – for example self-managed teams, work groups and supervisor-led teams. When managers adopt one of these administrative practices, they also choose to adopt some, or all, of its component practices. Not only do each of these component practices, being different, mean different things to the organizations adopting them, but their meanings are likely to differ depending on which administrative practice the organization adopts, and whether or not the administrative practice is
institutionalization. That is, how a manager thinks about self-managed teams vis-à-vis supervisor-led teams will change depending on whether their organization is implementing Quality Circles or Manufacturing Best Practices, and the degree to which Quality Circles or Manufacturing Best Practices are institutionalized. Our specific focus here is on how the meaning of each component practice to decision-makers varies vis-à-vis the other components, how those meanings change as an administrative practice becomes institutionalized, and how these affect the form (the set of component practices) their organizations adopt.

We approach this problem by building on an emerging body of institutional research in which the researchers assume that social actors perceive the focal social object as a member of a relevant category (e.g. Lounsbury and Rao, 2004; Zuckerman, 1999, 2000; Zuckerman and Kim, 2003; Zuckerman, Kim, Ukanwa, and von Rittmann, 2003). In our case, the administrative practice is the relevant category and the component practices are its members. Previous research in psychology, linguistics, and anthropology has shown that differences in meaning between members of a category are associated with the category structure, and, in particular the grading of the category. Category grading captures the idea that some members are better exemplars of a category than others (see Barsalou, 1992; Lakoff, 1987). To proceed, we first develop a model of the way in which institutional processes shape the grading of a category. We then use the model to explain how the field offers up meanings to members, and how those meanings influence adoption behaviour. We build the model by drawing on Tolbert and Zucker’s (1996) model of the institutionalization of novel organisational forms (of which organisational improvement programs are an example). We use Tolbert and Zucker’s two motors of objectification, theorization and inter-organisational monitoring, to construct two ideal-typical models of the way institutional factors shape category grading. The theorization motor treats managers as meaning-makers who construct their grading of the category on the basis of the theories they learn from actors in their environment. The inter-organisational monitoring motor treats managers as empiricists who grade categories on the basis of their observations of their environment, and leads to the prediction that the category grading will be determined
by bandwagon effects. In both cases managers’ preferences for a practice are shaped by its location on
the category gradient along with idiosyncratic variables such as organisational culture and technical merit.

We test these models by examining a 1994 cross-sectional sample of industrial firms within
Australia and New Zealand that were adopting an institutionalizing practice called Manufacturing Best
Practices. We measure the theorization-driven grading by consulting a set of experts on Manufacturing
Best Practices. We measure the inter-organizational monitoring driven grading with a measure based on
the popularity of practices. When we assess the pattern of component practices that firms adopted against
these category gradings, we find support for predictions generated by the theorization model, but not for
those generated by the inter-organisational monitoring model.

This study contributes to theory in two ways. First, we construct a theory of the social
construction of choice in which institutional processes shape the way actors interpret a practice, and those
interpretations shape the choices managers make with regard to that practice. This builds on recent work
in which institutional variables explain choices (e.g., Cole, 1999; Lounsbury, 2001; Zilber, 2002) rather
than idiosyncratic (firm-level) variables such as technical need and organisational culture. Category
grading thus provides insight into the constitutive nature of institutions - how actors’ preferences and
choices are shaped by events in their environment, how institutions enable action (e.g. Swidler, 1986),
and more broadly, into agency within institutional theory (DiMaggio, 1988). Second, this study
highlights the importance of considering categories (and by extension all innovations) in the context of
dense networks of actors, relationships, and meanings (see also Mohr, Forthcoming; Phillips, Lawrence,
and Hardy, 2004), rather than in isolation. Our argument suggests that these dense networks shape the
ways in which organizations can incorporate new practices (e.g. Cole, 1999; Zbaracki, 1998), the extent
to which they will become the locations for playing out old conflicts (Phillips, Lawrence, and Hardy,
2004), the contradictions that need to be managed, and the adjustments that need to be made to make the
new practices work effectively. That is, variations in the structure of categories across fields, or between
categories within a field, are not arbitrary, but are intimately related to the networks of meaning and
action that pervade the field and the organizations within it.
Theory & Hypotheses

Actors make sense of their environments, decide, and act. Our starting premise is that they do this by constructing internal representations of their external world, organised in a set of categories (e.g. Barsalou, 1999; Douglas, 1986). A category is “a number of objects that are considered equivalent” (Rosch, 1978 p. 30). As opposed to analytic categories like “the set of even numbers”, folk categories and conceptions of their membership are created through cultural processes (Barsalou, 1992; Lakoff, 1987). “Manufacturing Best Practices”, “Total Quality Management”, and “Six Sigma” are examples of folk categories, and practices such as “self managed teams”, “pareto charts” and “statistical process control” are examples of members of all of them. Institutional theory asserts that actors learn categories through their immersion in the field (Berger and Luckmann, 1967; Douglas, 1986; Zucker, 1991), and then act on the world as a consequence of that categorical knowledge (e.g. Zuckerman, 1999). That is, the categories to which social actors assign objects influence their behaviour with respect to those objects (Berger and Luckmann, 1967). In our case, this means that their propensity to adopt a component practice (e.g. self-managed teams) will depend on whether decision-makers see the component practice as being a member of the relevant category (e.g. Manufacturing Best Practices). We argue further that their propensity to adopt will also depend on how they see it being a member. They are more likely to adopt self-managed teams if they believe they are an important element of Manufacturing Best Practices.

In this section we develop our argument, namely that institutional processes shape the structure of institutional categories relevant to a novel practice, and particularly their grading, and that grading affects organisations’ preferences for component practices within the novel practice. We build the argument in three steps. First we present two metaphors for category structure. One of them – the frying pan metaphor (in which categories are not graded) – has dominated institutional thinking, while the other -- the mixing-bowl metaphor (in which categories are graded) -- has received limited attention. These metaphors differ in the key respect that the mixing bowl metaphor can incorporate institutional drivers of differences between component practices. The frying pan metaphor, in contrast, can only allow variables exogenous to institutional theory (such as culture or technical considerations) to determine differences.
between members of the category. Second, we argue that the category gradings suggested by the mixing
bowl metaphor can be created through processes of institutionalization and that the grading of the
category is a measure of the relative value of the component members. Finally, building on processes
during institutionalization that are emphasized in current research - theorization and inter-organisational
monitoring (Tolbert and Zucker, 1996) – we lay out the mechanisms by which institutionalization might
drive grading. We develop hypotheses around these two mechanisms.

**Category structure: The frying pan and mixing bowl metaphors**

Researchers adopt one of two metaphors to describe category structure (Lakoff, 1987). Much
research, including institutional research that focuses on the choice whether or not to adopt innovations, is
consistent with a metaphor we will call the frying pan. The metaphor (like a frying pan without a handle)
assumes that categories have flat bottoms and straight, near-vertical sides. That is, all members of the
category are equivalent and indistinguishable (being located somewhere on the featureless flat bottom)
and the boundaries of the category are sharply delineated (the vertical sides). Other researchers adopt a
metaphor we will call the mixing bowl. The mixing bowl slopes progressively from the center to the
edges. That is, the members of the category can be distinguished in terms of how well they represent the
category (with the ones located in the center being better representatives).

Researchers use the frying pan metaphor in three ways. The first way is to focus on the shape and
location of the sides of the pan. Some researchers assume the sides are vertical, but aim to explain their
location. For example, ceremonial adoption (Meyer and Rowan, 1977) can be interpreted as description
of the boundary of the set of practices which organisations can adopt acceptably. Similarly, Ingram and
Clay (2000) and DiMaggio and Powell (1983) invoke transactions costs and coercive, normative, and
mimetic pressures respectively to explain why some practices are inside the pan, and some are outside.
Other researchers do not assume the sides are vertical. Oliver (1991) can be interpreted as arguing that
the sides of the pan are not quite vertical. Even when organisations face coercive constraints on their
actions, they have some “wriggle room” to shape their behaviour. In a similar way, Zuckerman & Kim
(2003) show that the boundaries are not necessarily so clear-cut, and that agents often determine which category a social object ends up in. The second way researchers use the frying pan is to focus on its bottom. Authors in this tradition argue that anything inside the category (i.e. in the flat part of the pan) is eligible for invocation to action. Beyond that, they are all equivalent. So, for example, Swidler (1986) speaks of culture as if it is a set of tools which can be applied with equal felicity to all problems. Douglas (1986 p. 61) explicitly rejects variations between members within a category as an important explanatory variable. The third way researchers use the frying pan metaphor is to divide the social world into two categories – generally things that are legitimate and things that are not. If something is legitimate (i.e. inside the pan), then actors can enact it without fear of sanction and can increase their legitimacy; if it is not, then they cannot (Meyer and Rowan, 1977). This research then seeks to explain something interesting about the category containing legitimate objects, such as its evolution (e.g. Hoffman, 1997).

An alternative to conceptualizing the category as being shaped as a frying pan is to conceptualize it as a mixing bowl. Rather than having a flat bottom and near-vertical sides, a mixing bowl rises progressively from the centre to the sides. Such a metaphor is consistent with a large body of theory and empirical research in linguistics (e.g. Lakoff, 1987), psychology (e.g. Barsalou, 1992; Rosch, 1978; Rosch and Lloyd, 1978), and anthropology (e.g. Kempton, 1981). That research suggests that all folk categories have a structure which matches the bowl metaphor. That is, the categories are graded -- some members are better exemplars of the category than others (Lakoff, 1987: 12). For example, the category “red” is graded in that “fire engine red” is a better example of red then “crimson”. Similarly, statistical process control is a better example of a Total Quality Management practice than fishbone diagrams (Zbaracki, 1998), Scottish knitwear manufacturers can be better or worse examples of particular types (Porac, Thomas, Wilson, Paton, and Kanfer, 1995), and political action is a better example of an activity in a rape crisis center than counselling if you are a feminist, while the converse is true if you are a therapist (Zilber, 2002). If all folk categories are graded, then institutionalized categories, being built from cultural processes, must also be, and so it makes sense to build our theory around that assumption.
We can see the advantage of the mixing-bowl metaphor when we consider our research question: how do organisational actors choose between component practices within a more general practice, or more generally, how organisational actors choose between elements in a set. All that research using the frying pan metaphor tells us about this question is that decision-makers will reject the elements of the set which are outside the category and will contemplate elements inside it. Such approaches say nothing about choices between different locations on its floor. Consequently, institutional researchers have generally relied on explanatory variables drawn from outside institutional theory to explain choices between set elements. These include technical requirements (Meyer and Rowan, 1977), the culture of the adopting organisation (Czarniawska-Joerges, 1992), and managers’ receptiveness to rhetoric from other adopters (Zbaracki, 1998). While these processes are undoubtedly at work, exclusive reliance on exogenous variables to explain choices is problematic if one of the central claims of institutional theory is that institutionalization shapes actors’ cognition (Douglas, 1986; Friedland and Alford, 1991; Scott, 2001). To live up to that claim, it should be able explain choice within a set directly, using endogenous variables. An example of such an explanation is Cole’s (1999) explanation for why US firms were more likely to adopt a form of TQM aligned with Crosby’s model than Juran’s, Deming’s, or as practiced by the Japanese. Cole argued that Crosby’s model resonated better with U.S. business culture than the others. U.S. business culture is a field-level variable whose content is determined by institutional processes. We argue that U.S. business culture is linked to Crosby’s theory by the way that managers graded the category “TQM”.

**Grading and institutionalization**

If institutionalized categories are graded, then two things must happen for a category to become institutionalized. First, the category must be defined; it must acquire an identity and boundaries. Berger and Luckmann (1967), building on the work of Schutz (1962; 1967), describe this as a process of reciprocal typification -- the category “manager” can only exist in the presence of the category “worker”, and workers and managers must identify themselves and each other as such. Douglas (1986), in contrast,
argues categories acquire identity through ideational links to natural systems. Lakoff and Johnson emphasise the role of metaphor and metonymy in category creation (Lakoff and Johnson, 1980). While these explanations differ, all three share the idea that the identity of a category will be linked, in the minds of people, to the identities of other categories, within a network of meanings. My understanding of managers is linked to my understanding of workers, and my understanding of markets is linked to my understanding of jungles. Workers and jungles, are linked, in turn, to my understandings of a number of other categories such as unions and poisonous vines. Second, the category must become graded. We argue, consistent with the psychological literature (e.g. Barsalou, 1983) that the meanings to which people have been exposed via discourse will shape the grading of the category (see also Bourdieu and Wacquant, 1992; Phillips, Lawrence, and Hardy, 2000). We lay out the mechanisms more fully below, but to exemplify this argument, our theorization model leads us to expect that managers will place component practices which threaten their power or cause them anxiety further down the grade than those which do not. That is, the grading of the category “manufacturing best practices” is linked to the categories “power” and “anxiety management”, which, in turn, are linked to other meanings. Consequently, for any given individual, the grading of a given category is likely to reflect the meanings to which they have been exposed historically through discourse. Some of those meanings will reflect individual experience, say an event on a holiday yielding a preference for a particular cheese. Some will reflect group experiences. For instance, art school graduates may well believe that “abstract expressionism” is a better example of art than the general population (while not necessarily saying it is the best exemplar). Finally, some will represent field-level experiences, as with Cole’s TQM adopters.

“Fire engine red” is a better example of red than “crimson”. Does that mean it is more highly valued? The psychological and linguistic literature makes no claim. We argue, however, that if social processes drive the grading, then the grading of the category also influences the relative value which people attribute to the category members. Selznick (1957) has argued that institutionalization leads people to ascribe value to objects beyond their technical merit. Strang and Meyer (1993) have further argued that institutionalization leads to the privileging of certain social objects over others. If we
combine these three strands -- category grading, value beyond technical merit, and privileging -- we are led to conclude that to the extent institutional processes drive the grading of the category, the grading will be a measure of the relative value people attribute to category members. Our example of the art school graduates would seem to support this assertion. That is, not only would we expect art school graduates to believe that “abstract expressionism” is a better example of “art” than the general population, but we would also expect them to value it relatively more highly. Similarly, institutional pressures will push managers to prefer the central members of the category (i.e. high on the grade) over peripheral ones (i.e. low on the grade). As a direct extension of this, the more institutionalized the category is, the more we expect actors to prefer the central members.

We now consider how institutional processes shape category grading, in order to develop propositions about our domain of interest (the adoption of a set of component practices during diffusion and at least partial institutionalization) and our category of interest (Manufacturing Best Practice programs).

**Institutional determinants of category grading**

Tolbert and Zucker (1996) argue that institutionalization occurs in four stages – innovation, habitualization, objectification, and sedimentation (see also Berger and Luckmann, 1967). According to their model, organizations create and adopt new practices largely independently in the innovation and habitualization stages. Imitation might occur, but on a local basis rather than because of any field-level consensus regarding the new set of practices. In the third stage, objectification, some degree of consensus develops among organisational decision-makers concerning the value of a practice. There are two drivers of objectification – theorization, in which potential adopters learn a mental model of the innovation and convert that model to practice, and inter-organizational monitoring, in which they observe and copy others). In the fourth stage, sedimentation, the practice is taken-for-granted and perpetuated through time and across space (Tolbert and Zucker, 1996).
Tolbert and Zucker’s model of the first two stages is built on a frying pan metaphor, so we will discuss those stages in terms of a mixing bowl metaphor before moving to the stage of empirical interest – objectification. If organizations act independently, as stated in their model (above), then (by virtue of the independence) there are no field-level forces acting to grade the category and pattern behaviour, hence the frying pan. We argue, however, because novel practices do not enter the field independent of prior practices and pre-existing systems of belief, that a mixing bowl metaphor may give more insight. As we saw above, the grading of categories is determined, in part, by their relationships with other categories in the perceptual system of the perceiver. Consequently, because other relevant categories are institutionalized (such as beliefs about the prerogatives of managers, the roles of unions, and the problems organisations face), because of technical considerations (some practices are pragmatically more useful) and because of prior attempts and discourse to deal with performance gaps within the same management fashion niche (Abrahamson and Fairchild, 1999), we do not expect organisations to behave independently. Rather, we expect these prior factors to drive order and pattern behaviour. We also would expect these same drivers to influence the theorization process (Abrahamson and Fairchild, 1999). That is, while those practices, and their popularity, would not yet be associated with a particular theorized category, once the category is theorized, players in the field will have a good idea of how to interpret the theory, and have preferences for it and elements of it, before they even “open the book” or “attend the seminar” to learn about it.

The logical consequence of this shift to a mixing bowl metaphor for our interest in novel practice adoption is that we expect, when observing new sets of practices entering a field, that history will matter. Prior institutions will shape the way novel practices are interpreted, categorized and institutionalized (c.f. Cohen and Levinthal, 1990 at the organizational level; and Cole, 1999; Dobbin, 1994; Westney, 1987 at the field level). For example, Cole observed that U.S. managers struggled with Total Quality Management until they could put it into a form that was moderately consistent with prior field-wide cultural assumptions (i.e. the Crosby model). Once they achieved this, adoption proceeded more smoothly. That is not to say that novel practices must fit seamlessly into the iron cage of prior
institutions (DiMaggio and Powell, 1983). Because social systems are open, there are always competing meanings floating around, including those that arrive with the innovation. These novel and competing meanings can disrupt the prior meanings and prior categories and change them, just as TQM changed US business culture (Cole, 1999; Easton and Jarrell, 2000). Consequently, when we consider the role of categories in new practice adoption, we should assume that the categories are plastic, and not static. They can stay institutionalized but evolve – their membership and grading can change – as institutional processes play themselves out. For example, in Manufacturing Best Practice programs in Australia and New Zealand, “Lean Production” went from being a relatively unimportant member of the category in 1994 to being quite important by 2005.

In summary, we expect a category to express some structure and to exhibit some grading even during the pre-institutional phases of innovation and habitualization. We expect that grading will become more pronounced and more stable during the objectification and sedimentation phases. Tolbert and Zucker propose two basic drivers of objectification, namely theorization and inter-organisational monitoring. We discuss these in the next two sections and show how we expect them to affect the grading of the category of interest.

Theorization and mixing bowls

In theorization, a group of institutional entrepreneurs (DiMaggio, 1988) both present a problem to the field, and justify a particular innovative solution on logical or empirical grounds (Strang and Meyer, 1993). Such a process bounds the category by creating a basis for inclusion and exclusion of particular practices. It also grades the category by privileging certain elements as central and others as peripheral. Once the category is theorized, we expect the grading to affect managers’ preferences. In the next stage, as the practices become progressively objectified, and shared social meanings develop, actors will not only recognise the grading, but also give it normative significance. Consequently, the impact of grading on managers’ preferences is likely to be greater. The third stage, sedimentation, results in actors not only recognising the grading, and giving it normative significance, but also taking it for granted and seeing it
reflected in the environment with which they interact. Consequently, the impact of grading on preferences is likely to be greater still. Furthermore, Swanson & Ramiller (1997) argue that theorization occurs relatively early in the diffusion process, so we expect the marginal effect on grading to be greatest then. In summary, not only do we expect to see grading at all stages of the institutionalization process, but also we expect the impact of grading on preferences, and hence on behaviour, to increase as institutionalization progresses, with the greatest effect early on.¹

To understand how category grading induced by institutional processes affects the practice organizations adopt as diffusion progresses, we need to partial out the between-firm (i.e. field level) and within-firm (i.e. organization level) processes which occur over time. Between firms, the practice is institutionalizing and diffusing; within firms, the practice is being customized. Grading affects both processes. Below, we first address between-firms processes by assuming that the practices are institutionalizing as they diffuse over time, but that the firms adopt without customizing. Then we address within-firm processes by assuming that all firms customize over time, but that the level of institutionalization is constant throughout the diffusion process. When we combine the two effects, we will see that at any given point in time we expect later adopters to have adopted relatively more central practices and relatively fewer peripheral practices than early adopters.

Looking between firms, all adopters must trade-off between their idiosyncratic needs and pressures from the environment. If the practice (category) institutionalizes as it diffuses, we would expect early adopters to adopt component practices (members) much more in accordance with their idiosyncratic needs than later adopters, because early adopters face a more weakly graded category, and hence weaker relative pressure from the environment to adopt the central members. Later adopters will tip the balance

¹ Given that our argument presumes voluntary adoption by the focal organisation, it is worth discussing briefly the coercive case. In a coercive situation, the focal organisation acts at the behest of the coercive agent. The coercive agent can give the focal organisation choice as to how it acts, or it can prescribe behaviour. If the focal organisation has choice, it can act on the basis of either theorization or inter-organisational monitoring. If the coercive agent imposes behaviour on the focal organisation, it can either impose the consequences of its theorization, or the product of its inter-organizational monitoring, although in this case, the theory is less likely to be sensitive to the adopter’s context. Therefore, while the coercive case is much more complicated, theorization or inter-organisational monitoring (or both) are expected to still be at the root of behaviour.
towards those central members, because those adopters face an increasingly structured category, and hence stronger pressures to adopt the central members. At the end of the diffusion process, the early adopters will have adopted a larger proportion of component practices that are peripheral to the category than the late adopters.

Now, looking within firms, assume that all the diffusion occurs at some constant level of semi-institutionalization (Tolbert and Zucker, 1996). While institutional processes will affect adoption patterns at this stage, firms will also customize the form that they actually adopt (Czarniawska-Joerges and Sevâon, 1996). If institutional forces result in greater pressure to adopt the central practices, we expect a given firm will tend to adopt them first. As time passes, we expect it to drop some practices and replace them with new ones. We expect the firm to consider the most central of the remaining practices first, ceteris parabus, as it customizes its adoption to its idiosyncratic needs. The consequence is that the firm will be dropping central practices and adding progressively more peripheral ones, with the effect that it will move down the category gradient. Zbaracki (1998) observes just this. That is, rather than his research sites implementing Quality Management practices in accordance with their idiosyncratic needs, the sites considered the tools in a particular (and the same) order (Zbaracki, 1998 and personal communication, 2005). To a first approximation, they first considered the general TQM methods, then they considered the seven tools based on statistical process control methods (i.e. the Pareto diagram, the cause-and-effect diagram, the histogram, the control chart, the scatter diagram, graphs and check-sheets), third they considered the seven new tools (the relations diagram, the affinity diagram, the tree diagram, the matrix diagram, the matrix data-analysis diagram, the Process decision Program Chart, and the arrow diagram), and finally they considered intermediate statistics (such as design for experiments). We infer this to represent the grading of the category for two reasons. First, the training manuals that three of his firms used (the hotel, the defence contractor and the government agency) broadly support this argument (Zbaracki, personal communication, 2006). Second, Zbaracki’s defence contractor - the one organisation whose TQM program was subject to coercive pressure and surveillance (c.f. Westphal, Gulati, and

---

2 Details are available from the first author. Zbaracki only had the three manuals at the time of the communication.
Shortell, 1997) and thus would be expected to favour environmental pressures over idiosyncratic needs - implemented component practices with preferences approximately reflecting this ordering (See Zbaracki 1998, table 2, page 625). That is, the firms appear essentially to descend the gradient, from most central to most peripheral, seeing which tools they can match to their culture, technical needs, and capabilities, and then adopting them on that basis. The empirical consequence of this is that, at the end of the diffusion process, we expect the earlier adopters, having had more time to customize, to have adopted a larger proportion of component practices which are peripheral to the category than later adopters.

In combination, the two mechanisms suggest that at any point in the diffusion process, predictions based on institutionalization processes and customization processes will both lead to the same structural outcome. That is, the arguments above arguments suggest that the category grading will consistently and increasingly pressure adopters to choose more central component practices, and thus we hypothesize:

1. Organizations that adopt an administrative practice (the category) will adopt component practices which are central to the category’s grading more frequently, than they will adopt component practices which are peripheral.

2. At any point during diffusion, later adopters will have adopted a larger proportion of the central component practices than earlier adopters.

Furthermore, building on Swanson & Ramiller’s (1997) work:

3. The proportion of central practices adopted will vary curvilinearly with the time of adoption.

**Inter-organisational monitoring and mixing bowls**

In contrast to theorization, inter-organisational monitoring (at least in its ideal-typical form) does not involve the transmission of theorized meanings about a practice from the field to the organisation. Instead, it assumes actors in the focal organisation induce their preferences from their observations of their environment (first-hand observations, news stories, stock prices, etc.).

---

3 Note this hypothesis’s predictions are time-independent.
4 An alternative theoretical conception holds that inter-organizational monitoring is mindful, rather than purely empirical (e.g. Czarniawska-Joerges and Joerges, 1996; Sahlin-Andersson, 1996; Zbaracki, 1998). Under such a
consistent with this assumption: Inter-organisational monitoring leads to a preference for more popular practices (Tolbert and Zucker, 1996). Popular practices maximize network externalities (DiMaggio and Powell, 1983); minimize risk in uncertain environments (Abrahamson and Rosenkopf, 1993; Banerjee, 1992); and satisfy psychological needs for legitimacy and identity confirmation (Bandura, 1977). Furthermore, organisations have a greater probability of coming across the popular practices (Cohen, March, and Olsen, 1972); demonstrate reliability and accountability by following widely accepted practices (Hannan, 1977); and minimize risks associated with differentiation on a dimension which organisational actors consider strategically unimportant. Notwithstanding the motivation, the consequence is that the relevant category will be graded with the most popular members being central (i.e. most representative of the category) and the least popular being peripheral.

Westphal, Gulati, and Shortell (1997) in their analysis of TQM adoption by hospitals, presented two analyses of the way organisations selected between members of the category. One examined the extent to which organisations follow the theorized models of Deming, Juran, or Crosby. The other, which is of interest here, drew upon ideas of inter-organisational monitoring. They tested the idea that one of the determinants of adoption of a practice was the extent to which other organisations in the current or prior cohorts had adopted that practice. As institutionalization proceeded, they predicted that the level of conformity of hospitals – the extent to which they adopted configurations of practices similar to those of other hospitals – would increase. That is, they predicted and found that later adopters of TQM were more likely to adopt standardized forms of TQM than early adopters. In this study we test the same idea -- that later adopters will exhibit greater levels of conformity than early adopters. Hence, we hypothesize.

4. Later adopters will show higher conformity with current and prior cohorts than earlier adopters.

**Manufacturing Best Practices**

Our empirical tests concern the adoption of Manufacturing Best Practices prior to 1994 by manufacturers in Australia and New Zealand. Manufacturing best practices are a prominent example of conception, the relevant empirical propositions are hypotheses 1 - 3 above (rather than hypothesis 4 below) because the core activity that drives the adoption decision is a particular mode of theorization.
the broader concept of “best practices”. The core idea is that one can infer and then implement the best way to manage an operation or aspects of it (e.g. Hanson and Voss, 1995). Manufacturing best practices became an increasingly salient idea through the 1980s, driven, it seems, by the competitiveness “problem” faced by large U.S. firms, particularly in the automobile industry. The “problem” can be understood as a conflict between the institutionalized norm that managers be progressive (Abrahamson, 1996; Meyer, Boli, and Thomas, 1987), and the contradictory and widespread perception that U.S. firms’ management, production practices, and products were inferior to those of Japanese firms (Abrahamson and Fairchild, 1999; Cole, 1999). The initial solution to that “problem”, “Japanese management”, comprised a discrete set of practices such as quality circles and just-in-time logistics. “Total Quality Management” (TQM) emerged in the early 1980s and rapidly gained momentum as the idea that these practices needed to be integrated gained prominence (Abrahamson and Fairchild, 1999; Easton and Jarrell, 2000). With the rise of TQM, the idea spread that U.S. firms had become excessively internally focused and had lost touch with their environments (Venetucci, 1992). The search for best practices – initially from the Japanese, but later from best-in-class firms anywhere – was thus the next prominent solution to the competitiveness “problem.” (Hanson and Voss, 1995; Pilkington, 1998; Venetucci, 1992).

Manufacturing Best Practices provides an excellent venue for our study, for several reasons. First, it clearly represented a category in the minds of the adopters at the time. We analysed a database of business press articles and found that Manufacturing Best Practices was first theorized in the early 1980’s and then was referred to increasingly through the late 1980’s and beyond. Second, it was institutionalizing. By 1994, when our study was conducted, it was well established and prevalent – over 70% of firms in our sample stated they had adopted. A convenience sample of 10 manufacturing firms in 2005 indicated that Manufacturing Best Practices was well institutionalized, with the elements being largely unchanged from 1994 (the exception being that lean manufacturing had become quite central). Third, manufacturing in Australia and New Zealand represented a field. Among manufacturers in these two countries, there was extensive discourse about Manufacturing Best Practices, through education and training programs, professional associations, and agencies created to facilitate manufacturing
Finally, it has the dual virtues that coercive forces were relatively weak, so the risk of conflation of the preferences of coercive agents and adopting organisations is low, and the underlying concept was not overly prescriptive, so managers implemented with considerable discretion.

**Research methods**

*Australian Manufacturing Council Survey.* The data on manufacturing best practices adoption were obtained from a 1994 mail survey of Australian and New Zealand manufacturing firms conducted by the Australian Manufacturing Council (AMC) in conjunction with the Boston Consulting Group, the Australian Bureau of Statistics, and the Manufacturing Advisory Group (New Zealand) (Australian Manufacturing Council, 1994).

The survey was primarily designed to measure relationships between manufacturing practices and competitiveness, and has led to several published analyses in this vein (Australian Manufacturing Council, 1994; Challis, Samson, and Lawson, 2002; Samson and Ford, 2000; Samson and Terziovski, 1999; Terziovski, Samson, and Dow, 1997). The survey instrument was an extensive review of manufacturing strategy, practices and performance outcomes. The instrument was pilot tested at six sites and revised based on respondent reactions.

In addition to asking whether and when the responding firm embarked on a program aimed specifically at achieving “best practice”, the survey asked about a large number of manufacturing-related practices and technologies, the extent of their adoption, and their impact on the firm and its performance. These practices were assumed by the survey authors to be a superset of “best practices” (Samson, personal communication, December 2001). It also asked a number of questions relating to performance, and about the nature of the firm. Within the survey were nineteen practices for which we could obtain a relatively unambiguous measure of whether adoption had occurred. To increase confidence that adopters considered these practices to be component practices of a best practices program, we imposed an additional behavioural criterion that the practices should be used more frequently by adopters of

---

5 Further details and the analyses underlying these statements are available from the corresponding author.
Manufacturing Best Practices programs. We found, encouragingly, that 18 of the 19 practices met this criterion (using a t-test and p<0.05 criteria).\(^6\) Further, we asked our experts (see below under centrality survey) to comment on the list. This led us to eliminate three more.\(^7\) The fifteen remaining practices, which we refer to as “component” practices for expositional convenience, formed the basis for the dependent measures used in the study and are listed in Table 1:

| Table 1 about here |

**Sample.** The survey instrument was sent to a stratified random sample of the manufacturing sites that were registered with the Australian Bureau of Statistics or Statistics New Zealand and employed more than 20 people in 1993. The sample was stratified within each country using twelve industry codes (ASIC and NZIC) and three size categories (25-49, 50-100, and over 100). The stratification resulted in each cell containing at least 15 respondents.

The survey was mailed to 4,000 manufacturing site managers, 3,000 in Australia and 1,000 in New Zealand. 1289 responses were received within a ten-week period, with response rates of 32% for Australia and 38% for New Zealand. The AMC subsequently conducted a telephone survey of 108 non-respondents, who were asked a subset of questions with high predictive validity for the survey as a whole. No significant response bias was found (Australian Manufacturing Council, 1994). Of the 1289 responding sites, 1113 had no missing data and 801 of these (72%) indicated that the firms had adopted a best practices program by the time of the survey.

**Centrality Survey.** To operationalize the category grading of Manufacturing Best Practices as managers were likely to encounter it as they monitored their environment, we assessed the relative centrality of Manufacturing Best Practices component practices using a small snowball sample of expert

---

\(^6\) This practice had an adoption fraction of 95%, indicating that many organisations may have adopted independent of their Best Practices program.

\(^7\) An expert informed us that firms with computer-integrated manufacturing are more likely to have adopted best practices programs, not the other way around. In the case of production planning and control and warehouse management, an expert argued that a best practice program will influence how you perform those functions, but not whether you do them. His argument is supported by their very high adoption fractions (99% and 96% respectively) for these practices among adopters of best practices.
researchers into manufacturing competitiveness. We first wrote to academic researchers we had identified as experts in the area of manufacturing competitiveness, either through personal knowledge or through our reading. We sent them the list of eighteen practices described above. We asked these researchers to indicate “the relative centrality (i.e. closeness to the prototype) of the listed practices to the concept of manufacturing best practices, as you believe decision-makers in manufacturing firms would have understood them in 1994.” (emphasis in original) using five-point Likert scales. In addition we asked them for their comments and the names of other researchers who could respond for us. We wrote to eight people initially, and using first and second-order referrals, added eight more. We received five valid responses. Four others declined to participate. Three of those felt their expertise was too narrow (two had researched automobile manufacturing exclusively and one claimed to know only about human resources practices). One was known as an advocate for a particular approach and felt that people told him what they thought he wanted to hear. The remaining seven declined to respond to our e-mail.

Variables

Practice Centrality Score. For each of the fifteen component practices, this variable is the average of the centrality score assigned to that practice by the set of experts through the procedure described above. Table 1 also shows the practice centrality scores and population-wide adoption fractions (percentage of firms which have adopted manufacturing best practices which have also adopted a particular practice).

Practice Centrality Score Triangulation. As a way of triangulating on the category structure of Manufacturing Best practices, we constructed separate, theoretically-motivated rankings of the centrality of the fifteen practices. We did this based on the theory above and discursive models of institutionalization (Munir and Phillips, 2005; Phillips, Lawrence, and Hardy, 2004) and before receiving any responses from the experts. Using a three-point scale, we gave three points to Total Quality Management and Benchmarking, since they are the rhetorical lynchpins of Manufacturing Best Practices
(see Table 1)(Zbaracki, 1998). To differentiate other practices, we noted that two fundamental themes in the organizational literature are that managers are centrally interested in maintaining power and control (Perrow, 1986), and in avoiding anxiety (Schein, 1985). Thus we gave two points to component practices which meshed better with the prior network of meaning, being those which we assessed as being unlikely to threaten managers’ control or power or to create anxiety. We gave one point to those practices which we thought could be threatening or anxiety provoking (i.e. any practices that could be expected to involve active interaction with customers, suppliers, or employees).

We did not use our theoretical centrality measure to analyse adoption patterns themselves. Rather, it enabled us to analyze the relationship between a theoretically-derived centrality measure and the practice centrality score provided by our experts. We did so using ordinary-least square regression, with the experts’ measure as the dependent variable and the theoretically-derived measure as the independent variable. Given that there is only one independent variable, the results are not tabulated. Even with the small sample size (n=number of practices=15), there was a highly statistically significant relationship between the two variables (t=4.02, p=0.001). This result is consistent with the idea that the grading of the category is shaped by the network of meanings in which it is embedded, and particularly on a combination of variables that draw upon our understanding of the nature of rhetoric, on one hand, and the nature of managers on the other. While we used the expert’s scores in our analysis, their agreement with the theory-derived scale increased our confidence in the validity of experts’ scores.

Time of Adoption. The survey asked whether the site had embarked on a program aimed specifically at achieving “Best Practice” and if so, when such efforts commenced. Four categories were provided – before 1985, 1985-1988, 1989-1991, and post-1991 (the survey was mailed in January 1994). Approximately 70% of responding firms had adopted by the time of the survey, with the bulk of them in the last two categories⁸. This variable is coded somewhat differently in the two analyses we perform, in order to ease interpretation of interaction terms. The adoption curve is described below:

⁸ When modeling firm behavior, we assume that time of adoption of a “best practices” initiative represents the time the firm signed on to the crystallizing category in its field and started to act on it. At that time, the firm would have
Time of adoption squared. Square of the time of adoption variable. Used to measure curvilinear effects with time of adoption.

Not MBP Adopter. In some analyses, we included both adopting and non-adopting firms. These firms are identified by a dummy variable that (for ease of interpretation) takes value 1 for firms that did NOT adopt MBP and zero for those that did adopt.

Control Variables. Measures were included to control for characteristics of firms that might influence the dependent variables. Firm Size is the log of the number of full-time permanent employees at the time of the survey. Private is a dummy variable indicating that the firm is privately held (as opposed to a public company or government corporation). Independent is a dummy variable indicating a positive response to the first category of an item which asked the site an “independent firm,” an “operating unit of a large firm,” or a “subsidiary of parent or holding company.” Technology Level is calculated from a section of the survey that asked the extent to which several manufacturing technologies such as CNC machines, robots and materials-working lasers contributed to the firm’s competitiveness. Twelve such measures were available. Contribution was assessed using a five-point scale with the same anchors as seen earlier. The technology level was constructed as the mean level across the twelve measures (alpha .80). # of Unions is the number of unions reportedly represented at the site. Unionized is a dummy variable indicating that the number of unions is greater than or equal to one. Bargaining Unit is a dummy variable indicating whether or not the workplace operated as a single bargaining unit. Employee Growth is the log of the number of full-time permanent employees at the time of the survey divided by the number of such employees two years previously. Australian site is a dummy variable indicating that the respondent is physically located in Australia (as opposed to New Zealand). European

--- Figure 1 about here ---

started to adopt novel practices, influenced by the state of the category at the time. Such adoption is not a static process, as we have described above. As such, lateness of adoption does not refer to the time at which firms adopted particular component practices. Rather, it is measured in terms of the time at which they subscribed to a set of ideas (i.e., “MBP”) and started to introduce new practices to reflect them
Analysis and Results:

We performed two analyses to test the four hypotheses advanced above. The first tests hypotheses 1, 2 and 3 by analysing the relationship between relative centrality of the practices and their likelihood of adoption by the firm. The second, which tests hypothesis 4, closely follows Westphal, Gulati and Shortell (1997), and analyses the relationship between time of adoption and conformity to the prevailing form of MBP.

Analysis 1: Influence of centrality on likelihood of adoption (of component practices)

In the first analysis, our concern was with whether centrality influenced the likelihood of adoption of a particular practice by a particular firm. Accordingly, our unit of analysis and our dependent variable is adoption (or not) of a component practice, by a specific firm. Since there are 15 practices that we analyse, we constructed 15 corresponding data records for each firm. Thus for the 815 firms which adopted MBP, we have 12225 data records, and for the 1113 firms for which we had data, we have 16695 records. The 15 records for each firm differ only in an indicator variable showing adoption (or not) of the practice the record stands for, the associated practice centrality score, and related interaction terms.

Table 2 shows the means, standard deviations, and bivariate correlations for firms that had data available for the analysis. It is worth noting the positive but low correlation (0.11) between practice centrality and likelihood of adoption (which can also be seen in Table 1). This indicates that our experts were not simply nominating the most popular practices as the most central, and thus provides empirical support for the conceptual distinction between centrality as driven by theorization and as driven by popularity (through inter-organizational monitoring). We used logistic regression with the 0-1 indicator of adoption of a particular component practice by a particular firm as the dependent variable. We used clustering (by firm) and the Huber-White sandwich method to adjust standard errors for possible non-independence of records caused by duplicating records for each firm. Table 3 provides the results.
We examine hypothesis 1 in two ways. First, model 3A examines just the adopters of MBP and finds a statistically significant relationship between component practice centrality and the likelihood of adoption ($z = 13.1$, $p=0.001$). This supports hypothesis 1.

Second, we can go beyond just the adopters by seeing whether the centrality of a practice has a greater effect on the likelihood of adoption for firms that have adopted MBP than for firms that have not. We add this second test because, as we believe is the case for virtually all managerial innovations, firms can adopt any of the component practices without adopting MBP as package. In fact, for some practices, such as for housekeeping and preventative maintenance, adoption rates were over 90% among MBP adopters and non-adopters alike, which is significantly higher than the 72% adoption rate for MBP itself in the sample. If institutional processes are active, then the centrality of a practice should have more influence on the decision-making of adopters than non-adopters, because the adopters have been exposed more strongly to the theorization of MBP and because adoption signals their greater acceptance of that theorized model. Therefore, we see whether the effect of MBP as a graded category will also be seen in the relative effects of centrality on likelihood of adoption.

Accordingly, Model 3B considers both adopters and non-adopters of MBP, rather than just adopters of MBP as in Model 3A (n increases from 12225 data records and 815 firms, to 16695 data records and 1113 firms). We add a dummy variable (set to one if “Not adopt MBP”) as well as an interaction term: Not adapt MBP *Practice Centrality. The dummy variable is a control to account for the overall lower likelihood of component practice adoption by firms that did not adopt MBP. As would be expected, it is negative and highly significant ($z=-12.6$, $p<.001$). The interaction term is the term of interest. It is also negative and highly significant ($z=-6.3$, $p<.001$), indicating that practice centrality has a considerably weaker relationship to the adoption patterns of non-MBP adopters than it does for MBP adopters. Indeed, the -.24 coefficient for the interaction term is more than two-thirds the size of the .34 coefficient for the practice centrality term itself, indicating that the influence of practice centrality is
reduced by more than two-thirds for non-MBP adopters as compared to MBP adopters. These results are consistent with Hypothesis 1 and add to the support from Model 3A\(^9\).

Model 3C adds variables to model 3B in order to test hypothesis 2 that the component practices adopted by later adopters will be more affected by the category grading, and thus by the theorized centrality of component practices. The added variables are for adoption period (“Time of Adoption”) and an interaction term: Time of Adoption*Practice Centrality. In this analysis, the time periods on the survey are coded with earlier adopters having higher values: A value of 3 represents adoption before 1985, 2 adoption in 1985-88, 1 adoption in 1989-1991 and 0 adoption post-1991 (i.e., 1992, 1993, and early 1994). This coding eases interpretation because the latest adopters constitute the omitted class. The time of adoption variable is a control to account for differences in the overall lower likelihood of component practice adoption across adoption periods. As might be expected, it is positive, indicative that earlier adopters tended to have adopted more practices from among the 15 component practices we examine. The interaction term is the term of interest. It is negative and highly significant \((z=-3.03, p<0.002)\), consistent with hypothesis 2 that the component practices adopted by later adopters will be more affected by the category grading, and thus by the theorized centrality of component practices. The 0.41 coefficient for Practice Centrality represents that for the last adoption cohort (the omitted category).

\(^9\) We also ran additional analyses in which we controlled for additional characteristics of firms and practices. Specifically, in one set of additional analyses we controlled for the total number (out of 15) of component practices the firm adopted, and in a second set we controlled for the sample-wide fraction of MBP-adopting firms that adopted specific component practices. We did not present these additional analyses because while these control measures are conceptually desirable, they are problematic methodologically. This is because those measures are constructed from the dependent variable itself, so including them is akin to putting the same term on both sides of the equation. A more sophisticated and longitudinal approach (which our data do not allow) would be needed to properly implement such controls. Nevertheless, for the specific purpose of assessing the robustness of the analyses, we did run models with these controls individually included (not presented, available from authors). The z-scores for the controls were very high (>50), indicative of the specification issue we outlined. The results were quite robust. In both cases, the relative ordering of the effects of centrality was preserved (e.g., adopters were more influenced by centrality as in model 3b, and later adopters were most influenced among adopters as will be evident in model 3c) at high levels of significance \((p<0.01)\). The model that controlled for the number of practices adopted by the firm also preserved the strong main effect of centrality as in model 3a. The model that controlled for the fraction of adopting firms, however, showed no main effect for centrality and in some specifications had a negative effect. We believe this last difference is an artefact arising from the particular form of the relationship between the overall popularity of practices and their place in the MBP category grading. Specifically, some practices that are very popular overall (such as housekeeping and preventative maintenance that were mentioned previously) are of in the middle of the centrality scores for MBP. This popularity, which is unrelated to MBP, makes it impossible to draw conclusions regarding the effect of centrality when popularity is included.
The 0.09 coefficient for the Time of Adoption*Centrality term indicates that centrality affects each prior cohort by that much less, suggesting the effect is not only highly significant but of a quite substantial magnitude.

Model 3D uses a quadratic model to test hypothesis 3, that the effect of theorization on adoption will be greatest early in the diffusion process. The model adds (Time of Adoption)^2, and the interaction term (Time of Adoption)^2*Practice Centrality. The (Time of Adoption)^2 variable is the base term for the interaction and controls for differences in the overall likelihood of component practice adoption across adoption periods. Its coefficient is negative, indicating that the tendency for earlier adopters to adopt more practices is greater for the earlier cohorts. The interaction term is the term of interest. It is negative and highly significant (z=-2.92, p<0.004), consistent with hypothesis 3 that the effect of category grading on adoption behaviour will be strongest early in the diffusion process.

**Analysis 2: Influence of time of adoption on conformity**

We performed a second analysis to assess the influence of time of adoption on conformity and test hypothesis 4, that later adopters will demonstrate higher conformity with current and prior cohorts than earlier adopters. In this analysis, we follow the approach used by Westfall, Gulati and Shortell (1997) in their study of conformity among Total Quality Management practice adopters. Their analysis was at the level of the firm’s overall pattern of adoption. Though this differs from our first analysis, we chose to use this method both because it allows us to compare our results to theirs, and because data limitations (see footnote above) made us reluctant to try to use the first analysis’s approach in this case.

**Conformity.** Westfall, Gulati and Shortell (1997) constructed a conformity measure to capture how closely a firm matched the configuration of quality management practices that was prevalent in the field. They did this by examining which of a set of practices the firm adopted and which it did not adopt. For each firm and each practice, they measured conformity as the percentage of other firms whose response (adoption or non-adoption) matched that of the focal firm, from among firms that had adopted a best practice program in the same or earlier time period as the focal firm. They constructed their firm
level conformity score by summing this measure across the set of practices in their study. The matching procedure underlying the measure allows conformity to reflect adoption of widely-used practices, as well as non-adoption of less widely-used practices. We used this approach to build a firm-level conformity score based on the fifteen practices in Table 1.

We analyzed the relationship between time of adoption and conformity using ordinary-least square regressions, with the firm-level conformity score described above as the dependent variable, and time of adoption as the independent variable. In this analysis, the time periods on the survey are coded in ascending order, with 1 representing the before 1985 category, 2 representing 1985-88, 3 representing 1989-1991 and 4 representing post-1991.

The descriptive statistics for this analysis are provided in Table 4, and the results in Table 5. We used ordinary least-square regression in this analysis.

Model 5A shows that there is no statistically significant relationship between conformity and lateness of adoption (t=−1.56 n.s.), contrary to what hypothesis 4 would lead us to expect. 10

Discussion

Endogenous choice, enablement, and agency

Our finding (H1) that theorization-derived centrality predicts adoption of manufacturing best practices provides evidence that the grading of an institutionalizing category affects managers’ preferences for component practices within it. Further, our finding (H2) that such centrality has a stronger influence for later adopters provides evidence that institutional processes, and particularly theorization

---

10 Given this unexpected result, and the support for the non-linear model under hypothesis 3, we conducted some post-hoc analyses. Examination of the data shows that while the conformity increases among the first three cohorts, the last cohort has a significantly lower average conformity score than the penultimate cohort. This suggests the possibility of a curvilinear relationship, so we added a quadratic term for time of adoption. The model shows strong evidence for a quadratic relationship (t=3.64 p<0.001 for time of adoption, t=−3.96, p<0.001 for the quadratic term). This result remains counter to hypothesis 4. One explanation for this result is that we observe low conformity among the last cohort because members have only adopted partially. Analytically, the Westphal, Gulati, and Shortell measure gives equal weight to the practices that a firm adopts and the practices it does not adopt. Accordingly, we constructed a second measure in which we only considered the practices which a firm had adopted in the construction of the measure. This second analysis gave qualitatively identical results. Both analyses are available from the corresponding author.
processes, shape the structure of categories relevant to a novel practice, and particularly their grading. Our finding (H3) that the shift towards the central practices happens early in the diffusion process provides further evidence for the role of theorization processes. Finally, the lack of support for H4, that conformity-seeking predicts adoption of manufacturing best practices, provides evidence that inter-organisational monitoring does not have a statistically significant separate effect in our dataset. By implication, it also indicates that the results for H1-H3 are not artefacts of conformity-seeking behavior. Taken together, these findings suggest that it is useful for institutional theory to consider the grading of categories, in addition to their existence, in explanations of the adoption of novel practices in particular, and managerial action more generally.

Category grading has the potential to be a powerful analytic tool that extends the reach of institutional analysis, because it provides an approach to constructing theories in which institutional processes shape the way actors interpret a practice, and those interpretations shape the choices managers make with regard to that practice. That is, choice becomes endogenous. Category grading thus provides insight into the constitutive nature of institutions (Scott, 2001) - how actors’ preferences and choices are shaped by events in their environment. Under such a model, two mechanisms shape the way actors understand and value particular practices: the theories they learn, and the press of the popularity of certain practices and actions (Tolbert and Zucker, 1996).

Further, models that consider category grading and choices within the category extend our understanding of how institutions (here an institutionalizing practice) do not just constrain, but rather enable managerial action (e.g. Swidler, 1986). While researchers such as Swidler have proposed that actors make creative choices among culturally-available options (i.e., forms of a practice), the mixing bowl metaphor provides insight into how managers make such choices. In essence, our arguments and findings suggest that one effect of theorization and institutionalization is to create a ranked menu of (legitimate) options from which organizational actors can choose. This is also quite a different mechanism and emphasis from the institutional "pressures” that dominate most extant framings. In essence, such constraint-based framings usefully focus on the way the boundary of the category
determines legitimacy and so limits managers to choices within the category (the frying pan metaphor),
but again such framings provide little insight into which of the available options managers will choose.

More broadly, we believe that by considering category structure we open the door for much more
comprehensive models of agency within institutional theory. Institutional theory has been criticized for
focusing overly on constraint (DiMaggio, 1988). We suggest that this arises, in part, from the way
categories have been conceptualized within the theory. Prior conceptions of agency have adopted the
frying pan metaphor and have principally concerned themselves with agency in the context of between-
category issues such as the creation and determination of category boundaries (e.g. Douglas, 1986), the
allocation of objects to categories (e.g. Zuckerman and Kim, 2003), and resistance and compliance at the
boundaries of categories (e.g. Oliver, 1991). In contrast, we argue that agency is also important within
categories. The mixing bowl metaphor (i.e. category grading) enables us to consider this type of agency.
It allows us to consider the way that institutional factors (theorization and inter-organizational
monitoring) combine with non-institutional factors (such as technical merit, embeddedness in networks,
and absorptive capacity) to shape the preferences of actors.

**Networks of meaning within fields**

In addition to drawing researchers to consider the structure of categories, the theoretical
arguments and empirical evidence presented here also highlight the importance of considering the
relationships between categories.

As we noted above, Tolbert and Zucker (1996), in their model of institutionalization, make the
simplifying assumption that new practices and ideas enter the field autonomously. In particular, they
conceptualize the early stages of institutionalization as being relatively independent of “the common core
of knowledge and ideas that make an innovation attractive” (1996:181). Our model focuses attention on
institutionalization as the operation of social processes that shape that “common core of knowledge and
ideas” itself, rather than being something that operates alongside it. Our result that later adopters adopted
relatively more central practices suggests that institutionalization changed what Manufacturing Best
Practices meant to adopters. This is consistent with the notion that fields, rather than comprising independent categories, consist of dense networks of actors, relationships, and meanings (see also Mohr, Forthcoming). In particular, new categories (i.e. ideas and practices) must contend with prior categories (i.e. ideas, practices, and social relations)(Phillips, Lawrence, and Hardy, 2004) from the outset. These prior categories -- the prior systems of meaning of potential adopters and other actors -- shape the ways the new category will be interpreted and incorporated into the field. In the case of manufacturing best practices, we argued that novel practices impinge on the relationships between management and labor, and between companies and their customers and suppliers. We then considered the way micro-level (organizational behavior) variables, such as power and anxiety avoidance, are likely to shape those relationships. Our argument suggests that the prior configurations of those micro-level categories will shape the ways in which organizations can incorporate the new practices (e.g. Cole, 1999; Zbaracki, 1998), the extent to which they will become the locations for playing out old conflicts (Phillips, Lawrence, and Hardy, 2004), the contradictions that need to be managed, and the adjustments that need to be made to make the new practices work effectively. That is, variations in the structure of categories across fields, or between categories within a field, are not arbitrary, but are intimately related to the networks of meaning and action that pervade the field and the organizations within it.

Limitations and further research

Methodologically, there are a number of ways in which this study could be improved. In particular, it would be valuable to conduct a longitudinal study in which category structures and the practices being adopted are measured contemporaneously. This implies a need to measure category grading at the level of the field, category grading and preferences at the level of the organization, and adoption behavior. Furthermore, researchers could create more nuanced measures of category structure by using sophisticated tools available from anthropology and psychology (e.g. Barsalou, 1992; Kempton, 1981).
Second, this study suggests the value of a fully elaborated and tested model of the determinants of category structure. Our triangulation of expert and theoretically-derived centrality scores in the methods section suggests that the central members of the category will be those which make the category distinctive (those closely aligned with Total Quality Management and benchmarking), while those which a most peripheral will be those which mesh least with the existing category structure (those which threaten managerial power or induce anxiety). Notwithstanding, there is a place for a full-blown theoretical and empirical development of this issue.

Relatedly, the idea that fields can be considered as networks of meanings, actors, and practices, in which the categories structure each other needs to be developed considerably before it will be truly useful for institutional theory. For example, one interesting question is how category structures and specifically the strength of category gradings vary across the social space of the field. Research on this question would connect category structure to existing interest among institutional theorists about how institutional processes vary across actors and social space (e.g. Friedland and Alford, 1991; Strang and Tuma, 1993).

Finally, this study calls for much more careful theory development and empirical research concerning theorization, inter-organisational monitoring, and the relationship between them. Given that grading is driven by fundamentally different mechanisms in the two cases, but we can only end up with one category structure in the end, there is a place for careful thought and empirical research. In the case of inter-organisational monitoring, there is no a priori reason to expect all organisations to develop the same theory of the practices – unless of course theorization is really driving the process. And, in fact, if there is extensive buffering (Thompson, 1967), it is unlikely that they will move to the same theoretical model. If that is the case, is it reasonable to call these processes institutional? If not, what are they? If so, how do we reconcile it with the rest of institutional theory? In short, there is a need to understand these two processes in much more detail, to tie them coherently into institutional theory, and to understand the sorts of environments in which one, or the other, dominates.

More broadly, taking category structure and particularly category grading seriously raises a number of interesting theoretical questions. Here we highlight four which seem to be particularly
relevant. First, in this paper we have focused on the category structures for one set of actors in the field, namely the decision-makers in adopting firms. How do category structures vary with different institutional roles and locations in the field? Second, in this study we have assumed, consistent with most institutional theory, that category structures are stable and grounded in unconscious cognitive processes, and people’s basic assumptions about the nature of the world (see Lakoff, 1987; Lakoff and Johnson, 1980). A post-modern rendering of the same phenomena, in contrast, would treat category structure, and the relationships between categories, much more superficially. Which rendering best describes which aspects of the organizational world? Third, and relatedly, the comparison between institutional and post-modern models of category structure resonates strongly with Argyris’s distinction between espoused theories and theories-in-use (Argyris, 1990). That is, do people maintain multiple category structures (e.g. what I believe, what I know to be normatively correct, what I believe is possible within my circumstances, etc.)? If so, how do we incorporate these multiple structures into coherent theoretical models? Finally, in our study, there was a clear theoretical and empirical link between the psychological idea of representativeness and the sociological idea of value. That is, members of the category that are more representative are also more valuable. Is that always the case?

More broadly, in this paper, we have looked only at the relationship between category grading, institutionalization, and new practice adoption. Many other areas of institutional theory, including a more comprehensive examination of fields, organizational performance, ceremonial adoption, and institutional change would benefit from scrutiny under the same lens.
### Tables

<table>
<thead>
<tr>
<th>Practice</th>
<th>Centrality Score</th>
<th>Theoretical score</th>
<th>Adoption fraction (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Benchmarking</td>
<td>4.6</td>
<td>3</td>
<td>60</td>
</tr>
<tr>
<td>2. Taking the requirements of customers into account when designing new products and services</td>
<td>4.0</td>
<td>2</td>
<td>91</td>
</tr>
<tr>
<td>3. Designing products to match manufacturing and other capabilities</td>
<td>3.6</td>
<td>2</td>
<td>65</td>
</tr>
<tr>
<td>4. Systematically and regularly measuring customer satisfaction</td>
<td>3.6</td>
<td>1</td>
<td>65</td>
</tr>
<tr>
<td>5. Machine set-up time reduction</td>
<td>4.2</td>
<td>2</td>
<td>87</td>
</tr>
<tr>
<td>6. Statistical process control</td>
<td>4.0</td>
<td>2</td>
<td>86</td>
</tr>
<tr>
<td>7. Preventative maintenance</td>
<td>4.0</td>
<td>2</td>
<td>96</td>
</tr>
<tr>
<td>8. Housekeeping</td>
<td>3.6</td>
<td>2</td>
<td>98</td>
</tr>
<tr>
<td>9. Self managed and/or cellular work teams</td>
<td>3.8</td>
<td>1</td>
<td>77</td>
</tr>
<tr>
<td>10. Having an organization-wide training and development process including career path planning for all employees</td>
<td>3.4</td>
<td>1</td>
<td>43</td>
</tr>
<tr>
<td>11. Regularly and formally measuring employee satisfaction</td>
<td>3.2</td>
<td>1</td>
<td>39</td>
</tr>
<tr>
<td>12. Working closely with suppliers in product development</td>
<td>3.6</td>
<td>1</td>
<td>48</td>
</tr>
<tr>
<td>13. Total quality management</td>
<td>5.0</td>
<td>3</td>
<td>77</td>
</tr>
<tr>
<td>14. Just-in-time</td>
<td>4.8</td>
<td>2</td>
<td>62</td>
</tr>
<tr>
<td>15. Manufacturing resource planning (MRP, MRPII)</td>
<td>3.4</td>
<td>2</td>
<td>52</td>
</tr>
</tbody>
</table>

Number of observations: 801

**Table 1.** Component practices examined in this study, experts’ centrality scores, theoretical scores, and their adoption fraction.
```
<table>
<thead>
<tr>
<th>#</th>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Adoption (dummy)</td>
<td>.65</td>
<td>.48</td>
</tr>
<tr>
<td>2</td>
<td>Practice Centrality*</td>
<td>.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3</td>
<td>Not adopt MBP (dummy)</td>
<td>.28</td>
<td>.45</td>
</tr>
<tr>
<td>4</td>
<td>Not adopt MBP * Centrality</td>
<td>.00</td>
<td>.53</td>
</tr>
<tr>
<td>5</td>
<td>Time of adoption (0-3)</td>
<td>.55</td>
<td>.88</td>
</tr>
<tr>
<td>6</td>
<td>Adoption cohort * Centrality</td>
<td>.00</td>
<td>1.04</td>
</tr>
<tr>
<td>7</td>
<td>Firm Size</td>
<td>4.47</td>
<td>1.01</td>
</tr>
<tr>
<td>8</td>
<td>Private</td>
<td>.55</td>
<td>.50</td>
</tr>
<tr>
<td>9</td>
<td>Independent</td>
<td>.39</td>
<td>.49</td>
</tr>
<tr>
<td>10</td>
<td>Technology level</td>
<td>.79</td>
<td>.81</td>
</tr>
<tr>
<td>11</td>
<td># of unions</td>
<td>1.78</td>
<td>1.63</td>
</tr>
<tr>
<td>12</td>
<td>Unionized</td>
<td>.80</td>
<td>.40</td>
</tr>
<tr>
<td>13</td>
<td>Bargaining Unit</td>
<td>.53</td>
<td>.50</td>
</tr>
<tr>
<td>14</td>
<td>Employee Growth</td>
<td>.02</td>
<td>.16</td>
</tr>
<tr>
<td>15</td>
<td>Australian site</td>
<td>.74</td>
<td>.44</td>
</tr>
<tr>
<td>16</td>
<td>European parent</td>
<td>.12</td>
<td>.33</td>
</tr>
<tr>
<td>17</td>
<td>US parent</td>
<td>.07</td>
<td>.25</td>
</tr>
<tr>
<td>18</td>
<td>Japanese parent</td>
<td>.02</td>
<td>.13</td>
</tr>
</tbody>
</table>

Table 2: Means, standard deviations and bivariate correlations of variables used to test hypothesis 1 & 2. (N=16695)
(Industry control dummy variables (11) not shown.)
* Note that because each firm is associated with a separate data record for each of the 15 component practices, there is no correlation between any firm-level variable and the practice centrality score here. That is, the statistics here do not reflect on the average centrality of the practices the firm adopts. Rather, they are statistics related to the (always the same 15) practices each firm is considered to be at risk of adopting. The same comment applies for the two interaction terms for which the practice centrality score is a component.
```
<table>
<thead>
<tr>
<th>Dependent Variable -&gt;</th>
<th>Adoption of component practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model -&gt;</td>
<td>3a</td>
</tr>
<tr>
<td>Practice Centrality</td>
<td>.34*** (.03)</td>
</tr>
<tr>
<td>Not adopt MBP</td>
<td>-.58*** (.05)</td>
</tr>
<tr>
<td>Not adopt MBP * Centrality</td>
<td>-.24*** (.04)</td>
</tr>
<tr>
<td>Time of Adoption (0=late)</td>
<td></td>
</tr>
<tr>
<td>(Time of Adoption)^2</td>
<td></td>
</tr>
<tr>
<td>Time of Adopt * Centrality</td>
<td></td>
</tr>
<tr>
<td>(Time of Adopt)^2 * Centrality</td>
<td></td>
</tr>
<tr>
<td>Firm Size</td>
<td>.11*** (.03)</td>
</tr>
<tr>
<td>Private</td>
<td>-.01 (.06)</td>
</tr>
<tr>
<td>Independent</td>
<td>-.13† (.07)</td>
</tr>
<tr>
<td>Technology level</td>
<td>.32*** (.03)</td>
</tr>
<tr>
<td># of unions</td>
<td>-.04* (.02)</td>
</tr>
<tr>
<td>Unionized</td>
<td>.01 (.08)</td>
</tr>
<tr>
<td>Bargaining Unit</td>
<td>.08 (.05)</td>
</tr>
<tr>
<td>Employee Growth</td>
<td>-.04 (.17)</td>
</tr>
<tr>
<td>Australian site</td>
<td>-.07 (.06)</td>
</tr>
<tr>
<td>European parent</td>
<td>.01 (.08)</td>
</tr>
<tr>
<td>US parent</td>
<td>.10 (.10)</td>
</tr>
<tr>
<td>Japanese parent</td>
<td>.53** (.18)</td>
</tr>
<tr>
<td>Constant</td>
<td>.25 (.16)</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-7201.77</td>
</tr>
<tr>
<td>n</td>
<td>12225</td>
</tr>
</tbody>
</table>

**Table 1**

*** p<.001, ** p<.01, * p<.05, † p<0.10 two-tailed tests.

Industry control dummy variables (11) included in all models but not shown.

**Table 3: Logit model of influence of component practice centrality on adoption likelihood.**
| 1   | Conformity | 9.73 | 0.89 | 1.00 |
| 2   | Centrality | 3.97 | 0.10 | 0.27 | 1.00 |
| 3   | Time of adoption | 3.27 | 0.92 | -0.04 | 0.13 | 1.00 |
| 4   | Firm Size   | 4.55 | 1.09 | 0.24 | 0.13 | -0.01 | 1.00 |
| 5   | Private     | 0.49 | 0.50 | -0.14 | -0.12 | -0.11 | -0.28 | 1.00 |
| 6   | Independent | 0.34 | 0.48 | -0.20 | -0.14 | -0.13 | -0.36 | 0.61 | 1.00 |
| 7   | Technology level | 0.85 | 0.83 | 0.27 | 0.10 | -0.03 | 0.26 | 0.00 | -0.05 | 1.00 |
| 8   | # of unions  | 1.92 | 1.72 | 0.09 | 0.11 | 0.04 | 0.56 | -0.31 | -0.33 | 0.09 | 1.00 |
| 9   | Unionized   | 0.83 | 0.37 | 0.10 | 0.13 | 0.04 | 0.32 | -0.32 | -0.33 | 0.02 | 0.50 | 1.00 |
| 10  | Bargaining Unit | 0.57 | 0.50 | 0.13 | 0.04 | 0.03 | 0.14 | -0.14 | -0.21 | 0.06 | 0.03 | 0.29 | 1.00 |
| 11  | Employee Growth | 0.02 | 0.16 | -0.05 | -0.05 | 0.01 | -0.15 | 0.10 | 0.17 | -0.03 | -0.23 | -0.26 | -0.10 | 1.00 |
| 12  | Australian site | 0.72 | 0.45 | -0.03 | -0.01 | 0.07 | 0.01 | -0.06 | -0.11 | 0.07 | 0.02 | 0.01 | -0.04 | -0.08 | 1.00 |
| 13  | European parent | 0.12 | 0.33 | 0.01 | -0.01 | -0.01 | 0.08 | -0.15 | -0.26 | 0.04 | 0.09 | 0.09 | 0.06 | -0.12 | 0.08 | 1.00 |
| 14  | US parent   | 0.07 | 0.26 | 0.04 | 0.04 | -0.01 | 0.13 | -0.03 | -0.19 | 0.06 | 0.03 | 0.00 | 0.04 | -0.01 | 0.11 | -0.10 | 1.00 |
| 15  | Japanese parent | 0.02 | 0.16 | 0.11 | 0.00 | -0.08 | 0.17 | -0.01 | -0.08 | 0.09 | 0.12 | 0.07 | 0.07 | -0.06 | -0.08 | -0.06 | -0.05 | 1.00 |

Table 4: Means, standard deviations and bivariate correlations of variables used to test hypothesis 3. (N=801) (Industry control dummy variables (11) not shown.)
<table>
<thead>
<tr>
<th>Dependent Variable -&gt;</th>
<th>Conformity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model -&gt;</td>
<td>5a</td>
</tr>
<tr>
<td>Time of adoption</td>
<td>-0.051 (0.033)</td>
</tr>
<tr>
<td>Firm Size</td>
<td>0.126*** (0.035)</td>
</tr>
<tr>
<td>Private</td>
<td>-0.082 (0.077)</td>
</tr>
<tr>
<td>Independent</td>
<td>-0.255** (0.087)</td>
</tr>
<tr>
<td>Technology level</td>
<td>0.251*** (0.039)</td>
</tr>
<tr>
<td># of unions</td>
<td>-0.044† (0.023)</td>
</tr>
<tr>
<td>Unionized</td>
<td>0.003 (0.099)</td>
</tr>
<tr>
<td>Bargaining Unit</td>
<td>0.115† (0.064)</td>
</tr>
<tr>
<td>Employee Growth</td>
<td>-0.077 (0.199)</td>
</tr>
<tr>
<td>Australian site</td>
<td>-0.125 (0.069)</td>
</tr>
<tr>
<td>European parent</td>
<td>-0.153 (0.099)</td>
</tr>
<tr>
<td>US parent</td>
<td>-0.066 (0.121)</td>
</tr>
<tr>
<td>Japanese parent</td>
<td>0.130 (0.200)</td>
</tr>
<tr>
<td>Constant</td>
<td>9.537*** (0.232)</td>
</tr>
<tr>
<td>r²</td>
<td>0.17</td>
</tr>
<tr>
<td>n</td>
<td>801</td>
</tr>
</tbody>
</table>

*** p<.001, ** p<.01, * p<.05, † p<0.10 two-tailed tests. Industry control dummy variables (11) included in all models but not shown.

**Table 5: OLS model of the influence of time of adoption on conformity.**
Figure 1: Cumulative adoption of Best Practices initiatives among Australian & New Zealand manufacturing sites (n=1289)
References


